



Michigan Technological University
Create the Future Digital Commons @ Michigan Tech

Dissertations, Master's Theses and Master's
Reports - Open

Dissertations, Master's Theses and Master's
Reports

2014

ANALYSIS OF POPULUS TREMULOIDES CLONAL VARIATION AND DELINEATION IN THE OTTAWA NATIONAL FOREST

Kristina N. Flesher
Michigan Technological University

Follow this and additional works at: <https://digitalcommons.mtu.edu/etds>



Part of the [Ecology and Evolutionary Biology Commons](#)

Copyright 2014 Kristina N. Flesher

Recommended Citation

Flesher, Kristina N., "ANALYSIS OF POPULUS TREMULOIDES CLONAL VARIATION AND DELINEATION IN THE OTTAWA NATIONAL FOREST", Master's Thesis, Michigan Technological University, 2014.
<https://doi.org/10.37099/mtu.dc.etds/771>

Follow this and additional works at: <https://digitalcommons.mtu.edu/etds>



Part of the [Ecology and Evolutionary Biology Commons](#)

ANALYSIS OF *POPULUS TREMULOIDES* CLONAL VARIATION AND
DELINEATION IN THE OTTAWA NATIONAL FOREST

By

Kristina N. Flesher

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Applied Ecology

MICHIGAN TECHNOLOGICAL UNIVERSITY

2014

© 2014 Kristina N. Flesher

This thesis has been approved in partial fulfillment of the requirements for the Degree of
MASTER OF SCIENCE in Applied Ecology.

School of Forestry Resources and Environmental Science

Thesis Advisor: *Oliver Gailing*

Committee Member: *Martin Jurgensen*

Committee Member: *Evan Kane*

Committee Member: *Janice Glime*

School Dean: *Terry Sharik*

Table of Contents

<i>List of figures</i>	<i>vii</i>
<i>List of tables</i>	<i>ix</i>
<i>Acknowledgements</i>	<i>xi</i>
<i>Abstract</i>	<i>xii</i>
1 Introduction	1
1.1 <i>Phenology, morphology, and genecology</i>	<i>1</i>
1.2 <i>Phenotypic delineation</i>	<i>2</i>
1.3 <i>Genetic delineation</i>	<i>3</i>
1.3.1 Protein markers-	<i>4</i>
1.3.2 DNA markers-	<i>5</i>
1.3.2.1 -RAPDs-	<i>5</i>
1.3.2.2 -Microsatellites-	<i>6</i>
1.4 <i>Objectives, hypotheses, and study site</i>	<i>7</i>
2 Material and Methods	9
2.1 <i>Phenotypic clone identification</i>	<i>9</i>
2.2 <i>Clonal mapping</i>	<i>9</i>
2.3 <i>Stem selection</i>	<i>10</i>
2.4 <i>Phenotypic analysis</i>	<i>11</i>
2.4.1 Leaf measurements-	<i>11</i>
2.4.2 Visual analysis-	<i>12</i>
2.5 <i>Genotypic clone identification</i>	<i>13</i>
2.5.1 Sample selection and collection procedures-	<i>13</i>
2.5.2 Disruption and isolation of genomic DNA-	<i>13</i>
2.5.3 PCR procedures-	<i>14</i>
2.5.4 Capillary electrophoresis-	<i>14</i>
2.6 <i>Data analyses</i>	<i>14</i>
2.6.1 GenAlEx- genetic analyses-	<i>15</i>
2.6.1.1 -Probability of identification-	<i>15</i>
2.6.1.2 -Parameters of genetic variation-	<i>15</i>
2.6.1.3 -Clonal identification-	<i>16</i>
2.6.1.4 -Accuracy of clonal identification-	<i>17</i>
2.6.1.5 -Principal coordinate analysis-	<i>18</i>
2.6.1.6 -Relatedness measures-	<i>18</i>
2.6.2 XLSTAT- phenotypic analyses-	<i>19</i>
2.6.2.1 -Multicollinearity and correlation analysis-	<i>20</i>
2.6.2.2 -K-sample comparison of variance and ANOVA (with post-hoc test in SPSS)-	<i>20</i>
2.6.2.3 -Discriminant analysis-	<i>21</i>
2.6.2.4 -Classification and regression tree-	<i>21</i>
2.6.2.5 -Principal component analysis-	<i>22</i>
2.6.2.6 -Agglomerative hierarchal clustering-	<i>22</i>
2.6.3 PC-ORD- MRPP phenotypic analysis-	<i>23</i>

3	Results.....	25
3.1	<i>Genetic analysis</i>	25
3.1.1	Primer analysis-	25
3.1.2	Clonal analysis-	25
3.2	<i>Phenotypic analysis</i>	27
3.2.1	Phenotypic variable selection-	28
3.2.2	Phenotypic inter- and intraclonal agreement-	29
3.2.3	Accuracy of phenotypic variables for clonal assignment-	32
4	Discussion	36
4.1	<i>Genetic analysis</i>	36
4.2	<i>Phenotypic variable significance</i>	36
4.3	<i>Phenotypic inter- and intraclonal agreement</i>	36
4.4	<i>Phenotypic clonal assignment</i>	37
4.5	<i>Considerations for the application of phenotypic delineation techniques</i>	38
4.6	<i>Future research</i>	39
5	Conclusion	40
	Supplementary figures.....	42

List of figures

- Figure 2.1. Map of tagged *P. tremuloides* stems in Ottawa NF plot 28. Phenotypically delineated putative clone “3” and clone “Q” are represented together with the 120 exhaustively sampled stems. Distances from the plot center are illustrated in 5, 10, 15, and 20-meter circles. 10
- Figure 2.2. Leaf morphometric measurements applied to the selected *P. tremuloides* clonal colonies. Morphometric measurements included length measured in centimeters (A) and serration count (B). These were conducted with the upper epidermis facing upward. The two lengths were measured from the leaf base to the widest part of the leaf (length 1) and from the widest part of the leaf to the leaf tip (length 2). A leaf length ratio was calculated by dividing length1/length2 (A). Serrations were counted from the petiole to the tip on the west side of every leaf (B). 11
- Figure 2.3. Vegetative budbreak chart applied to selected *P. tremuloides* clonal colonies to analyze spring phenology. The scale ranges from "0" representing an unopened bud to stage "3" representing complete budbreak. 12
- Figure 2.4. A grey scale chart created to measure bark "brightness." This scale was based on a 20-color chart, which was reduced because the bark colors present among sampled *P. tremuloides* stems existed mostly in "grey scale." 12
- Figure 2.5. A color chart created to measure leaf senescence in nine stages (labeled 1-9) of *P. tremuloides* clonal colonies measured on September 17th. The chart represents the proportions of green, yellow, and brown visible in the canopy for each tree. Stage 1 represents a crown with all green leaves and no senescence, stage 5 represents a crown with all yellow leaves, and stage 9 represents a crown with all brown leaves. Stages 2, 3, and 4 represents 4:1, 1:1, and 1:4 proportions of green to yellow leaves, respectively. Stages 6, 7, and 8 represent 4:1, 1:1, and 1:4 proportions of yellow to brown leaves, respectively. 13
- Figure 3.1. The LTSP Ottawa NF plot 28 map of the selected *P. tremuloides* clonal colonies (BB, CC, E, G, M, and U). Distances are given in 5-meter increments from plot center. N, E, S, and W represent cardinal directions. 26
- Figure 3.2. A PCoA graphical representation of genetic relatedness among selected *P. tremuloides* clonal colonies (BB, CC, E, G, M, and U). Eigenvalues for each axis and the coordinates of each clone are given below the graph. 27
- Figure 3.3. Phenotypic comparison between selected *P. tremuloides* clonal colonies (BB, CC, E, G, M, and U). Phenotypic variables include Julian days to the start of vegetative budbreak (A), Julian days to the end of budbreak (B), the total number of Julian days between the start and end of budbreak (C), DBH (D), bark thickness (E), bark brightness

(F), leaf senescence (G), number of leaf serrations (H), and average leaf length ratio (I). Different letters above the std dev bars are representative of clonal means that are significantly different (Tahamane's T2, all 28

Figure 3.4. Multicollinearity statistics R^2 (A), tolerance (B), and variable inflation factor (VIF) (C), of phenotypic variables measured for *P. tremuloides*. Phenotypic variables: JDSBB=Julian days to start of budbreak; JDEBB=Julian days to end of budbreak; JDTBB=Julian days in total for budbreak; BT=bark thickness; BR=bark brightness; LS=leaf senescence; ALS=average number of leaf serrations; LR=average leaf length ratio. 29

Figure 3.5. Multicollinearity statistics R^2 (A), tolerance (B), and variable inflation factor (VIF) (C), of remaining phenotypic variables measured for *P. tremuloides* after removing variables found to be collinear. Phenotypic variables: JDEBB=Julian days to end of budbreak; BT=bark thickness; BR=bark brightness; LS=leaf senescence; ALS=average number of leaf serrations; LR=average leaf length ratio. 29

Figure 3.6. Discriminant analysis (A) for *P. tremuloides* genetically identified clonal colonies (BB, CC, E, G, M, and U) computed using the selected phenotypic variables. The correlation between the variables and factors F1 and F2 are shown in (B). JDEBB=Julian days to end of budbreak; BT=bark thickness; BR=bark brightness; LS=leaf senescence; ALS=average number of leaf serrations; LR=average leaf length ratio. 30

Figure 3.7. Principal component analysis (A) for *P. tremuloides* genetically identified clonal colonies (BB, CC, E, G, M, and U) computed using the selected phenotypic variables. The correlation between the variables and factors F1 and F2 are shown in (B). JDEBB=Julian days to end of budbreak; BT=bark thickness; BR=bark brightness; LS=leaf senescence; ALS=average number of leaf serrations; LR=average leaf length ratio. 32

Figure 3.8. Agglomerative hierarchal clustering (AHC) dendrogram of stem separation from the selected *P. tremuloides* clonal groups (BB, CC, E, G, M, and U) using phenotypic characters. The first letter or number (120, 3, or Q) represents the original phenotypic grouping, the second number is the sample number for each stem in the original grouping, and the final lettering represents the genetically identified clonal colony. This procedure makes no assumption about prior group assignment. 33

Supplementary figures

Suppl. Figure 1. Hand drawn *P. tremuloides* stem map of the Ottawa NF LTSP plot 28. The number of stems within 20 meters of plot center totals 585. Highlighted are the 181 stems selected for genetic and phenotypic analysis, which include putative clone "Q" (A), putative clone "3" (B), and the 120 exhaustive sampling (C)..... 42

Suppl. Figure 2. The 20-color chart used for initial bark color analysis. Colors were later converted into a "grey scale" format because most of the colors in sampled <i>P. tremuloides</i> were found within colors 1-5.	42
Suppl. Figure 3. LTSP Ottawa NF plot 28 map of all genetically identified <i>P. tremuloides</i> ramets and genets. Genets consisting of a single ramet are denoted with pink squares. Distances are given in 5-meter increments from plot center. N, E, S, and W represent cardinal directions.	43
Suppl. Figure 4. Classification and regression tree depicting <i>P. tremuloides</i> genetically identified clonal colonies (BB, CC, E, G, M, and U) separated phenotypically. Each branch separates stems based on one phenotypic variable with numbers representing the total number of ramets per genet per branch. BT=bark thickness; BR=bark brightness; LS=leaf senescence; ALS=average number of leaf serrations; LR=average leaf length ratio.	43

List of tables

Table 2.1. List of phenotypic measurements and the months they were recorded for the 120 exhaustive sampling, putative clone Q, and putative clone 3 <i>P. tremuloides</i> stems in the Ottawa NF LTSP plot 28.	11
Table 3.1. Table of the primers applied in this study. Included is information regarding the specific nature of the primers (SSR motif and annealing temperature) and statistical information regarding the choice of primer loci. The statistical data include the observed number of alleles (N_a), the effective number of alleles (N_e), the observed heterozygosity (H_o), the expected heterozygosity (H_e) probability of identification (PI), and probability of identification among siblings (PI_{sib}).	25
Table 3.2. The probability that <i>P. tremuloides</i> ramets in each clonal colony (BB, CC, E, G, M, and U) arose from sexual rather than asexual reproduction. Included is the area in meters each clone covers and the number of stems/m ² . N=number of ramets; P_{gen} =probability of the same genotypes belonging to different clones; P_{se} =probability of an observed genotype occurring more than once among identified genotypes.	26
Table 3.3. Pairwise genetic relatedness between selected <i>P. tremuloides</i> clonal colonies (BB, CC, E, G, M, and U). Results were obtained using the Lynch and Ritland (1999) relatedness measure in GenAlEx. Values near or less than zero indicate no genetic relatedness, values around 0.5 indicate half-siblings, and values around 1 indicate full-siblings. LR1 and LR2 represent the asymmetric estimator values and LRM represents the mean estimator value.	27

Table 3.4. Pearson's correlation matrix to detect excessive collinearity between phenotypic variables measured for *P. tremuloides*. JDSBB=Julian days to start of budbreak; JDEBB=Julian days to end of budbreak; JDTBB=Julian days in total for budbreak; BT=bark thickness; BR=bark brightness; LS=leaf senescence; ALS=average number of leaf serrations; LR= average leaf length ratio. 29

Table 3.5. Phenotypic pairwise comparison between selected *P. tremuloides* clonal colonies (BB, CC, E, G, M, and U) using a multi-response permutation procedure (MRPP). Given are the values for the t-statistic, intraclonal agreement value, and p-value for each pairwise comparison and clonal average. The higher the agreement value, the greater the phenotypic separation between clones. 31

Table 3.6. Results of the classification and regression dendrogram illustrating inter- and intraclonal agreement values (A-values) among selected *P. tremuloides* clonal colonies (BB, CC, E, G, M, or U). Each stem was analyzed and assigned to either its original clone or another clone based on phenotype. The results show the percentage of correct stem assignment in total and for each clone. The intraclonal distances for all stems in each clonal colony are also given and were calculated using a multi-response permutation procedure (MRPP). The delta shown is the observed delta with the expected delta=0.5.. 32

Table 3.7. Stem and leaf plot based on the results of the agglomerative hierarchal clustering (AHC). Phenotypic group assignment is illustrated for *P. tremuloides* ramets from selected clonal colonies (BB, CC, E, G, M, and U). The first letter or number (120, 3, or Q) represents the original grouping, the second number is the sample number of each stem in each original grouping, and the final letter represents the clonal colony..... 34

Table 3.8. Results of the classification and regression tree testing phenotypic inter- and intraclonal agreement (A-value) after an agglomerative hierarchal reclassification of *P. tremuloides* stems based on phenotypic characters. Each stem was analyzed and assigned to either its original phenotypic group or another phenotypic group. The results show the percentage of correct stem assignment in total and for each group. The intra-group distances for all stems in each phenotypic group are also given and were calculated using a multi-response permutation procedure (MRPP). The delta shown is the observed delta with the expected delta=0.5. 35

Acknowledgements

I would like to thank my advisor, Dr. Oliver Gailing, for his wisdom, guidance, and support on this project. He helped lead the project with his knowledge of genetics, assisted in making me an independent researcher, and further fueled my interest in forest genetics and molecular biology.

I would also like to thank my committee members, Dr. Martin Jurgensen, Dr. Evan Kane, and Dr. Janice Glime. Dr. Martin Jurgensen supported this project in numerous ways with his morale and scientific expertise, and brought the unique insight of a seasoned forester to this venture. Dr. Evan Kane contributed greatly with his exceptional knowledge of forests and soils, and I cannot thank him enough for opportunity to work with him as an assistant in his soil science class. Dr. Janice Glime strengthened this project with her profound botanical knowledge and I truly wish I had the opportunity to take her botany class.

There are several others that need to be acknowledged for their role in this project. First, I cannot thank Dr. Andrew Storer enough for being such a wonderful graduate student advisor and for finding me financial support in my final semester. His knowledge and guidance has truly benefitted myself and many other graduate students within our department. I also need to thank the MTU graduate school department. In particular, the Graduate Assistant Dean, Nancy Byers Sprague, for her role in helping me finish my degree by keeping up with my constant submission errors and going the extra mile to help me correct those blunders. I would also like to thank Dr. Debbie Dumroesse and Joanne Tirocke for their exceptional assistance in various aspects of this study. To my lab mates and colleagues, Sandra Owusu and Jennifer Riehl, who taught me how to perform most of my genetic analyses in both the lab and on the computer, thank you. I would not have known what I needed to do without both of these people guiding me along the way. This project would also not have been completed without hard working undergraduates. I would like to thank those undergraduates, Ted Maynard, Joseph Langley, and Bryan Reitter for their willingness to work holidays and weekends in some extremely unpleasant field conditions.

Additionally, I would like to thank the United States Forest Service, the MTU Ecosystem Science Center, and the MTU Biotechnology Resource Center for providing financial assistance for my graduate career, research, and poster conference.

Finally, I would like to thank my friends and family for their help and moral support. My mom, Karen Flesher and brother, Steven Githens, as well as my friend, Kyle Deane, assisted me with fieldwork when I most desperately needed help. In particular, it was my husband, Kevin Kruse, who volunteered numerous hours to help with fieldwork just so that his wife might have some spare time. I wish I could have assisted him with his graduate work in optical processing and polymer waveguides as much as he has helped me with my research.

Abstract

The technique of delineating *Populus tremuloides* (Michx.) clonal colonies based on morphology and phenology has been utilized in many studies and forestry applications since the 1950s. Recently, the availability and robustness of molecular markers has challenged the validity of such approaches for accurate clonal identification. However, genetically sampling an entire stand is largely impractical or impossible. For that reason, it is often necessary to delineate putative genet boundaries for a more selective approach when genetically analyzing a clonal population. Here I re-evaluated the usefulness of phenotypic delineation by: (1) genetically identifying clonal colonies using nuclear microsatellite markers, (2) assessing phenotypic inter- and intraclonal agreement, and (3) determining the accuracy of visible characters to correctly assign ramets to their respective genets. The long-term soil productivity study plot 28 was chosen for analysis and is located in the Ottawa National Forest, MI (46° 37'60.0" N, 89° 12'42.7" W). In total, 32 genets were identified from 181 stems using seven microsatellite markers. The average genet size was 5.5 ramets and six of the largest were selected for phenotypic analyses. Phenotypic analyses included budbreak timing, DBH, bark thickness, bark color or brightness, leaf senescence, leaf serrations, and leaf length ratio. All phenotypic characters, except for DBH, were useful for the analysis of inter- and intraclonal variation and phenotypic delineation. Generally, phenotypic expression was related to genotype with multiple response permutation procedure (MRPP) intraclonal distance values ranging from 0.148 and 0.427 and an observed MRPP delta value=0.221 when the expected delta=0.5. The phenotypic traits, though, overlapped significantly among some clones. When stems were assigned into phenotypic groups, six phenotypic groups were identified with each group containing a dominant genotype or clonal colony. All phenotypic groups contained stems from at least two clonal colonies and no clonal colony was entirely contained within one phenotypic group. These results demonstrate that phenotype varies with genotype and stand clonality can be determined using phenotypic characters, but phenotypic delineation is less precise. I therefore recommend that some genetic identification follow any phenotypic delineation. The amount of genetic identification required for clonal confirmation is likely to vary based on stand and environmental conditions. Further analysis, however, is needed to test these findings in other forest stands and populations.

1 Introduction

The technique of delineating *Populus tremuloides* (Michx.) clonal colonies based on morphology and phenology has been utilized in many studies and forestry applications since the 1950s (Barnes 1959). More recently, the availability and robustness of molecular markers, such as isozymes, random amplified polymorphic DNA (RAPDs), and nuclear microsatellites, have challenged the validity of such approaches for accurate clonal identification (Cheliak and Pitel 1984, Tuskan et al. 1996, Wyman et al. 2003). However, genetically sampling an entire stand is largely impractical or impossible. For that reason, it is often necessary to delineate putative genet boundaries for a more selective approach when genetically analyzing a clonal population. Here I re-evaluated the usefulness of phenotypic delineation by: (1) genetically identifying clonal colonies using nuclear microsatellite markers, (2) assessing phenotypic inter- and intraclonal agreement, and (3) determining the accuracy of visible characters to correctly assign ramets to their respective genets.

1.1 Phenology, morphology, and genecology

Populus tremuloides (trembling aspen), part of the Salicaceae family, is a temperate tree species that grows throughout the northern and central portions of North America (Perala 1990). White powdery bark, flattened petioles, and finely serrated and orbicular leaves characterize this species. While sexual reproduction does occur, vigorous vegetative reproduction is a prominent feature where suckers arise from a parental root system. This growth habit can form clonal colonies that are composed of genetically identical stems (Barnes 1966, 1969, Perala 1990). Clonal colonies are often referred to as genets. A genet, however, may refer to either a genetically distinct individual or a genetically distinct group that is composed of one or more stems, called ramets.

Because these clonal colonies, or genets, are genetically distinct, they may often be defined or delineated by certain phenological and morphological traits, such as the timing of vegetative budbreak, bud set, leaf senescence and leaf fall, leaf shape, leaf serration, bark color, diameter at breast height (DBH), stem height, and crown shape (Barnes 1969). Phenotypic expression of a genet is largely determined by the degree of genetic interaction with environmental conditions (St Clair et al. 2010). Certain phenotypic traits are more plastic than others and show greater intraclonal variation. Phenotypic traits that are more static within a genet are thus more accurate when delineating and identifying clonal colonies.

While many common garden studies have examined the effects of genetic control over phenology and morphology in the *Populus* species, very few publications exist that analyze these interactions specifically in *P. tremuloides*. St Clair et al. (2010) were one of the few who analyzed the genetic influence on phenotypic expression between genets in a *P. tremuloides* common garden study. Eighteen transplanted genets, totaling 417 trees and originating from the same location, were propagated in a common garden study over 28 years. While no intraclonal statistical differences in growth and physiology were

found, interclonal statistical differences were two- to sevenfold. Specifically, measured growth traits that included DBH, height, canopy fullness, and survival exhibited genetic correlations greater than 0.70 with broad-sense heritability estimates between 0.25 and 0.35. Thus, growth was demonstrated to be under moderate and continued genetic control. The effect of genotype on physiological characters that include the photosynthesis rate, stomatal conductance, and water potential, varied drastically within and between growing seasons even though genets were found to be significantly differentiated (St Clair et al. 2010).

Although publications analyzing the effect of genotype on phenology and morphology specifically pertaining to *P. tremuloides* are scarce, numerous common garden and plantation studies have been conducted on other species and hybrids within the Salicaceae family. Bradshaw and Stettler (1995) showed that genotype explained up to 98% of budbreak variation among several different clonal *Populus* hybrids, while another study using multiple different *Populus* hybrids found significant interclonal differences in budbreak and leaf fall (Pellis et al. 2004). In another publication on *P. nigra*, 76% to 86% of the variance for bud set in the onset-of-stage traits was attributed to genetic variation among clones (Rohde et al. 2011). Additionally, a common garden study using cuttings from natural populations of *P. tremula*, discovered a connection between nucleotide polymorphisms within the phytochrome B2 locus and differences in the timing of bud set (Ingvarsson et al. 2008). Finally, research conducted on *Salix* spp. in a clonal common garden study, demonstrated that genotype significantly influenced budbreak, leaf unfolding duration, growth cession, leaf abscission, and leafy period (Weih 2009).

In all of these common garden studies, phenotypic variation correlated strongly with genotype. However, interclonal variations, particularly in natural clonal populations, may overlap if genetic distances are low or if G×E interactions are high. While this may occur, the theory that *P. tremuloides* and *P. grandidentata* can form and exist in extensive clonal colonies was first conceived because of apparent phenological and morphological variations (Barnes 1966, 1969).

1.2 Phenotypic delineation

The use of phenology and morphology to delineate clonal populations of *P. tremuloides* has been widely accepted as a relatively accurate method in the forestry and ecological fields (Barnes 1966, 1969, Steneker 1973, Barnes 1975, Kemperman and Barnes 1976, Gom and Rood 1999). Phenological and morphological traits commonly used for clonal delineation include sex and flowering time, leaf senescence and leaf fall, leaf characters, stem and bark characters, and injury susceptibility (Barnes 1969). The theory behind this method of clonal delineation is that identical genotypes will exhibit the same or similar phenotypes, and therefore stems with similarly expressed traits should belong to the same genet.

Notable studies by (Barnes (1966), 1969)) and Kemperman and Barnes (1976) relied solely on phenotypic methods for the clonal delineation of *P. tremuloides* stands. In publications by (Barnes (1966), 1969)), clonal colonies were delineated when clones

were spatially separated and when clear differences in phenology and morphology existed between stems. These studies analyzed the range of inter- and intraclonal phenotypic variation, in part to determine proper techniques for clonal delineation in the field. The conflict in these papers is that many of the same phenotypic variables used to delineate the putative clones were also the variables analyzed to determine inter- and intraclonal variation. While the results presented by (Barnes (1966), 1969)) fit in the context of phenotypic expression being controlled by genotype, clonal colonies have been shown to exhibit a wide range of variation due to environmental influences. Similarly, two or more genets may exhibit phenotypic traits so uniform that they are indistinguishable from one another. Because phenotypic expression is most often influenced by both genotype and environment (i.e. climate and habitat), it is unclear whether the results presented by (Barnes (1966), 1969)) are conclusive. However, the analysis of enzyme variants (i.e. isozyme analysis), conducted on two to three stems in each putative clonal colony provided some evidence that these stands were each comprised of a single and unique genet (Barnes 1987).

Additionally, Kemperman and Barnes (1976) utilized similar techniques as (Barnes (1966), 1969)) to accurately delineate one of the largest organisms (by weight) in existence, an extensive *P. tremuloides* clone, dubbed “Pando.” Pando, genetically confirmed using nuclear microsatellite markers by DeWoody et al. (2008), consists of 47,000 ramets and inhabits an area of ~40 hectares. Even though 40 genets were found along the border of Pando (DeWoody et al. 2008), these results indicate that expressed traits can delineate clonal colonies with a high rate of accuracy. However, other recent studies genetically identifying clonal colonies have reported poor results when applying phenotypic techniques for clonal delineation (Cheliak and Pitel 1984, Tuskan et al. 1996, Wyman et al. 2003). These genetic studies are discussed in further detail within the sections pertaining to genetic delineation techniques (see sections 1.3.1 and 1.3.2.)

1.3 Genetic delineation

Genetic delineation and identification of clonal colonies relies on molecular markers to detect genetic and allelic variations among several different loci within a genome. A genetic locus in diploids, triploids, and tetraploids is the location on a specific chromosome set or group composed of two, three, or four related chromosomes, respectively (Conner and Hartl 2004, Rédei 2008). Alleles are variations in the genetic copies within the same chromosome set. While allelic variants may arise from somatic mutations caused by the environment or cell mitosis, allelic variants are most often a result of zygotic mutations formed during meiotic duplication and crossover (Conner and Hartl 2004, Rédei 2008).

Allelic variants at a specific locus may occur within an individual (maximum of two allelic variants in diploids) and may also be used to describe the genetic variation between individuals in a population (Conner and Hartl 2004, Rédei 2008). If an individual contains different allelic variants at the same genetic locus, it is heterozygous. If the opposite is true, that individual is homozygous. Co-dominant molecular markers, such as nuclear microsatellites and isozymes, are able to detect heterozygosity, whereas

dominant markers, such as RAPDs, can only detect specific genetic variants and not heterozygosity. When a population contains multiple allelic variants that are detectable by a molecular marker, the population is considered to be polymorphic at that specific locus. If no allelic variants are detectable, the population is said to be monomorphic at that specific locus (Conner and Hartl 2004, Rédei 2008).

In order for molecular markers to accurately identify a genet, the markers must be polymorphic within a population (Sunnucks 2000). Multiple genetic markers are used in most instances because several genets may contain the same genetic pattern at a few loci even though they may be genetically distinct. The ideal set of genetic markers is highly polymorphic (>6 alleles), co-dominant, and heterozygous in most individuals (Sunnucks 2000).

Here I review several studies that have used isozyme, RAPD, or microsatellite markers for clonal identification in combination with some form of phenotypic clonal delineation. Several of these studies (Cheliak and Pitel 1984, Tuskan et al. 1996, Wyman et al. 2003) were chosen because they are often cited as indications that phenological and morphological delineation techniques may be ineffective (Liu and Furnier 1993, Wyman et al. 2003, Suvanto and Latva-Karjanmaa 2005). However, other research provides evidence that many of these phenotypic characters can be very distinctive between genets in both common garden studies (Bradshaw and Stettler 1995, St Clair et al. 2010) and natural populations (DeWoody et al. 2008, Jelínková et al. 2013).

1.3.1 Protein markers-

The first molecular marker techniques utilized in genetic clonal delineation were not a direct analysis of DNA sequence variants, but analysis of enzyme variants, or isozymes (Brown and Moran 1981, Pasteur et al. 1988, Rédei 2008). Isozymes, often referred to as isoenzymes and allozymes, are groups of proteins that catalyze the same reaction, but differ by one or more amino acids in either length or composition. A single isozyme is typically composed of amino acid chains from several genic loci, but is also considered a co-dominant (Brown and Moran 1981, Pasteur et al. 1988, Rédei 2008). Mitton and Grant (1980), Cheliak and Pitel (1984), and Jelinski and Cheliak (1992) were among the first to genetically differentiate clonal populations of *P. tremuloides*. In these studies, isozymes were used to varying degrees in combination with phenotypic and/or geographical (i.e. distance, physical barriers, etc.) delineation techniques. Of these studies, Cheliak and Pitel (1984) were the first to test the accuracy of phenotypic delineation techniques using molecular markers.

According to Cheliak and Pitel (1984), 10 clones composed of 114 stems were “subjectively delineated” using stem phenology and morphology. Eight different polymorphic isozymes were used for clonal delineation that averaged around 1.65 alleles per isozyme locus. The genetic analysis identified 30 unique genets from the 10 putative clones. Two of these putative clones were accurately delineated and in another two putative clones, only two unique genotypes were found. All other putative clones contained 3-5 genets. The results presented by Cheliak and Pitel (1984) indicate that phenotypic delineation techniques were not adequate to identify all genets. However, it is difficult to assess the accuracy of such phenotypic delineation because very little information is presented on the applied delineation techniques.

Additionally, the analysis of expressed proteins is error prone because enzyme variation is often due to non-genic influences (i.e. environmental and post transcriptional/translational modifications) rather than allelic variants (Brown and Moran 1981, Poly 1997a, b, c, Buth and Murphy 1999). Furthermore, genic regions contain significantly fewer mutations than non-genic regions, making the detection of new genotypes difficult even when protein variation reflects genetic variation (Brown and Moran 1981, Buth and Murphy 1999). For these reasons, isozymes may not be able to identify all unique genets and may misclassify clonal groups when non-genic enzyme variation is misinterpreted as genic variation. This is not to say that the results by Cheliak and Pitel (1984) are inconclusive, but rather that a robust phenotypic analysis coupled with the use of variable DNA markers is needed for a more definitive conclusion.

1.3.2 DNA markers-

The majority of genetic marker techniques rely on the use of primers to identify and amplify DNA regions through polymerase chain reaction (PCR). Two of the predominant types of genetic markers that have been applied in molecular clonal delineation are RAPDs and microsatellites.

1.3.2.1 -RAPDs-

Before microsatellite markers became widely available in *P. tremuloides* and other species, RAPDs were quite commonly used in molecular ecology. RAPDs are relatively inexpensive markers that are easy to use (Williams et al. 1990, Hadrys et al. 1992). Several RAPD markers can be found from a single primer because primer-binding sites, or marker sites, are randomly distributed throughout a genome. The primers copy or amplify genomic fragments of varying lengths that are scored based on presence or absence, thus classifying RAPDs as dominant markers (Williams et al. 1990, Hadrys et al. 1992). These markers often show greater polymorphism than isozymes and can be 2.3 times more sensitive to genetic variation (Aravanopoulos 2010).

Tuskan et al. (1996) evaluated the post-fire seedling recruitment of *P. tremuloides* in Yellowstone National Park using RAPD markers for genetic identification. A total of 14 putative clonal colonies, ten mature stands and four seedling stands were characterized by leaf morphology and spatial separation. The number of sampled stems totaled 280. Using 14 primers, 163 polymorphic markers were found, with polymorphic loci averaging ~18% in adult stands and ~61% in seedling stands. Of the 280 sampled stems, 258 exhibited unique genotypes. All of the sampled seedlings and the majority of adult stems were genetically unique (Tuskan et al. 1996). While it is not unexpected that seedling stands should have a high level of genetic diversity, the genetic diversity of the adult stands is well above the number of genotypes expected in this area (DeWoody et al. 2009).

It should be noted that while the number of unique genotypes was high, the robustness of RAPDs is low. Analysis of RAPDs is difficult because of the inherent uncertainty of the genomic locations, allelic artifacts that are not representative of genetic variability, and frequent lack of consistent reproducibility of the RAPD amplification products (Jones et al. 1997, Bagley et al. 2001, Douhovnikoff and Dodd 2003). For example, Jones et al. (1997) tested the reproducibility of RAPDs in eight different

laboratories and found that, despite the ease of use, RAPDs had a reproducibility rate of only 75% in plants. Even when applying fastidious methodology, reproducibility rates can only be improved to 85% (Bagley et al. 2001). The erroneous nature of RAPD markers and the large number of the nearly identical genotypes found by Tuskan et al. (1996) suggests that clonal diversity was overestimated. Because of RAPD inaccuracies and because the measured leaf characters were not reported, inferences on the accuracy of phenotypic delineation cannot be made.

1.3.2.2 -Microsatellites-

Microsatellites, sometimes referred to as simple sequence repeats (SSRs), are a two to seven nucleobase pattern that is repeated several times within a given locus (Zane et al. 2002, Conner and Hartl 2004, Rédei 2008). The number of times the nucleobase pattern is repeated can vary between alleles and individuals. Therefore, it is the length of a microsatellite fragment, rather than the exact microsatellite sequence, that allows genetic differentiation among individuals at a specific locus. Microsatellite loci typically occur in non-genic regions and are hence more prone to zygotic mutations and accumulation of allelic variants. Multiple microsatellite primers, however, must be used because each primer is associated with one specific locus and because the same allele(s) may occur in multiple individuals. The number of primers needed for identification depends on the polymorphism within a population (Zane et al. 2002, Conner and Hartl 2004, Rédei 2008). Even so, microsatellite identification is a more reliable technique than isozymic or RAPD methods because the markers are co-dominant, highly polymorphic, and typically 100% reproducible (Jones et al. 1997, Zane et al. 2002, Conner and Hartl 2004, Rédei 2008).

Wyman et al. (2003) examined the genetic diversity of four *P. tremuloides* populations in Quebec, Canada, using four microsatellite markers for genetic identification. Three morphological characters were applied to phenotypically identify and delineate 60 putative clonal colonies (5 clones/stand) in three stands for each of the four populations. However, it is unclear which three characters were utilized, as the techniques are not exactly specified and are further described as “general morphological characters such as bark features, diameter at breast height (DBH), stem form, leaf shape and dentation, and crown shape (Wyman et al. 2003).” Approximately 14 to 15 stems/stand were selected for genetic identification, totaling 168-180 stems among the 60 putative clones. However, this is not certain because an exact figure is not provided (Wyman et al. 2003). Four microsatellite markers were used for molecular identification and contained an average of 7.43 alleles/locus. From these markers, it is stated that the number of genetically distinct individuals from the sampled 15 stems/stand, ranged from 1.62 to 2.2 times that of the phenotypic delineation (Wyman et al. 2003). Given that five putative clones were delineated in each stand, the average number of identified genets would range from 8.1 to 11 genets/stand. This, however, represents contradictory results to their stated ranges of 11 to 14 genets/stand out of the 14 to 15 stems/stand. This means that the number of unique genets/stand, when calculated from the stated ratio of unique genets to phenotypic clones per stand, is inconsistent to the stated number of unique genets/stand. With this information, the possible number of identified genotypes ranges

from 97.2 to 168 for an average of 8.1 to 14 genets/stand, respectively (Wyman et al. 2003).

Wyman et al. (2003) has been cited as an indication that phenotypic delineation is ineffective. These conclusions, however, are difficult to assess because the data presented are inconsistent, important information regarding clonality is missing, and phenotypic delineation techniques are not specified. Further information is required before the accuracy of phenotypic delineation can be determined.

In contrast, a more recent study by Jelínková et al. (2013) presents evidence that phenotypic delineation can accurately identify natural *P. tremuloides* clonal colonies. Here, 60 stems in three stands were sampled for genetic identification using microsatellite markers and leaf morphometric analysis using elliptical Fourier descriptors (EFDs). Collected leaves were from the same location on every tree and from the same developmental stage. The EFDs, applied within the program SHAPE v1.2 (Iwata and Ukai 2002), analyzed scanned leaves to define leaf shape using trigonometric functions, or harmonics. Among the three stands, 18 genets were identified using seven microsatellite markers. The EFDs were able to accurately delineate the two older stands, ages 38 and 42, containing five and two genets, respectively, but not the third and youngest stand, age 23, containing 11 genets. Other phenotypic identification techniques were applied and included budbreak, bark color, and bark texture. These were not included as part of the leaf morphometric delineation, but little success was reported in using spring phenology and bark assessments. Leaf or bud flush was reported to be the least reliable method of phenotypic delineation, but other research provides evidence contrary to this observation (Bradshaw and Stettler 1995, Pellis et al. 2004, Weih 2009). The accuracy of bark morphology as a clonal indicator is difficult to assess from this study because procedural applications are not fully described (Jelínková et al. 2013).

Jelínková et al. (2013) concluded that budbreak and bark characters were ineffective tools, but evidence contrary to this conclusion is presented in other publications, as previously mentioned (Bradshaw and Stettler 1995, Pellis et al. 2004, Weih 2009, Smith et al. 2011). However, the suitability of bark morphology as a clonal indicator is difficult to determine from their study because procedural applications are not fully described. Even so, two out of three stands could be accurately delineated based on leaf shape alone. Additional phenological and morphological traits may have more accurately delineated the third stand provided that the variables describing stem traits are based on categorical characterization or statistical assessment.

1.4 Objectives, hypotheses, and study site

The current research regarding clonality and phenotypic traits has generally supported the concept of phenotypic delineation. The application of those concepts, however, has had mixed results. While genotypic analysis is the most dependable way to determine clonality, it can be extremely difficult and costly to analyze a large stand of trees. In an attempt to elucidate and expand upon the current research regarding phenotypic delineation, a young *P. tremuloides* stand was examined within the Great Lakes long-term soil productivity (LTSP) study. LTSP plot 28 was chosen for analysis

and is located in the Ottawa National Forest, MI (46° 37'60.0" N, 89° 12'42.7" W). The study site was initiated and the stand was clear-cut in 1992. While all LTSP plots were set up with varying levels of compaction and organic matter removal, plot 28 contained no organic matter removal or compaction treatment. The fact that the history of this site is known and that significant soil alterations were not applied made it ideal to address the following three research objectives: (1) genetically identify clonal colonies using microsatellite markers, (2) assess phenotypic inter-and intraclonal variation, and (3) determine the accuracy of visible characters to correctly assign ramets to their respective genets as identified by microsatellite markers.

Based on these objectives, three general hypotheses were tested to assist in analyzing the accuracy of phenotypic delineation: (1) certain phenotypic traits will exhibit greater interclonal variation than others, (2) phenotypic expression will be significantly differentiated between genets, and (3) greater phenotypic variation will be exhibited between unrelated genets than between related genets.

2 Material and Methods

2.1 Phenotypic clone identification

Phenotypic clone identification of *P. tremuloides* stems in the Ottawa National Forest, MI LTSP study was conducted during the months of May and June 2012. Six out of 27 plots were reviewed for potential analysis. These included plots 6, 13, 17, 21, 22, and 28. In these plots, putative clonal colonies (genets) were initially identified using visual comparisons of bark texture and color. All stems similar for these characteristics were tentatively considered to be genetically identical and labeled using colored spray paint with either an identifying letter or number. Adjacent clonal groups were marked with different colored spray paint for further visual separation. Treatment plot 28, which is a total-tree harvest with no compaction treatment, was selected for analysis because it contained the largest number of stems and putative clones. In total, plot 28 contained 585 stems that were assigned to putative clonal colonies. It is important to note that the initial phenotypic identification measures only considered bark morphology, were highly subjective, roughly applied, and were not recorded. A more robust set of phenotypic measures (listed in Table 2.1) was developed for phenotypic analyses the following year once study plot 28 had been selected and the current literature on phenotypic delineation techniques could be fully reviewed.

2.2 Clonal mapping

After clones had been roughly delineated in the field, putative clones were mapped by hand using a pre-harvest plot map as a reference that was originally created in 1992 by the USDA Forest Service. The hand drawn map was used to select putative genets for genetic and phenotypic analysis (Suppl. Figure 1). Of the 585 stems, 181 stems were selected and labeled with individually numbered impression tags. The selected stems from the hand drawn map were then digitized using an overlying x, y coordinate grid so that the location of each stem could be accurately plotted in Microsoft Excel 2011 (Figure 2.1).

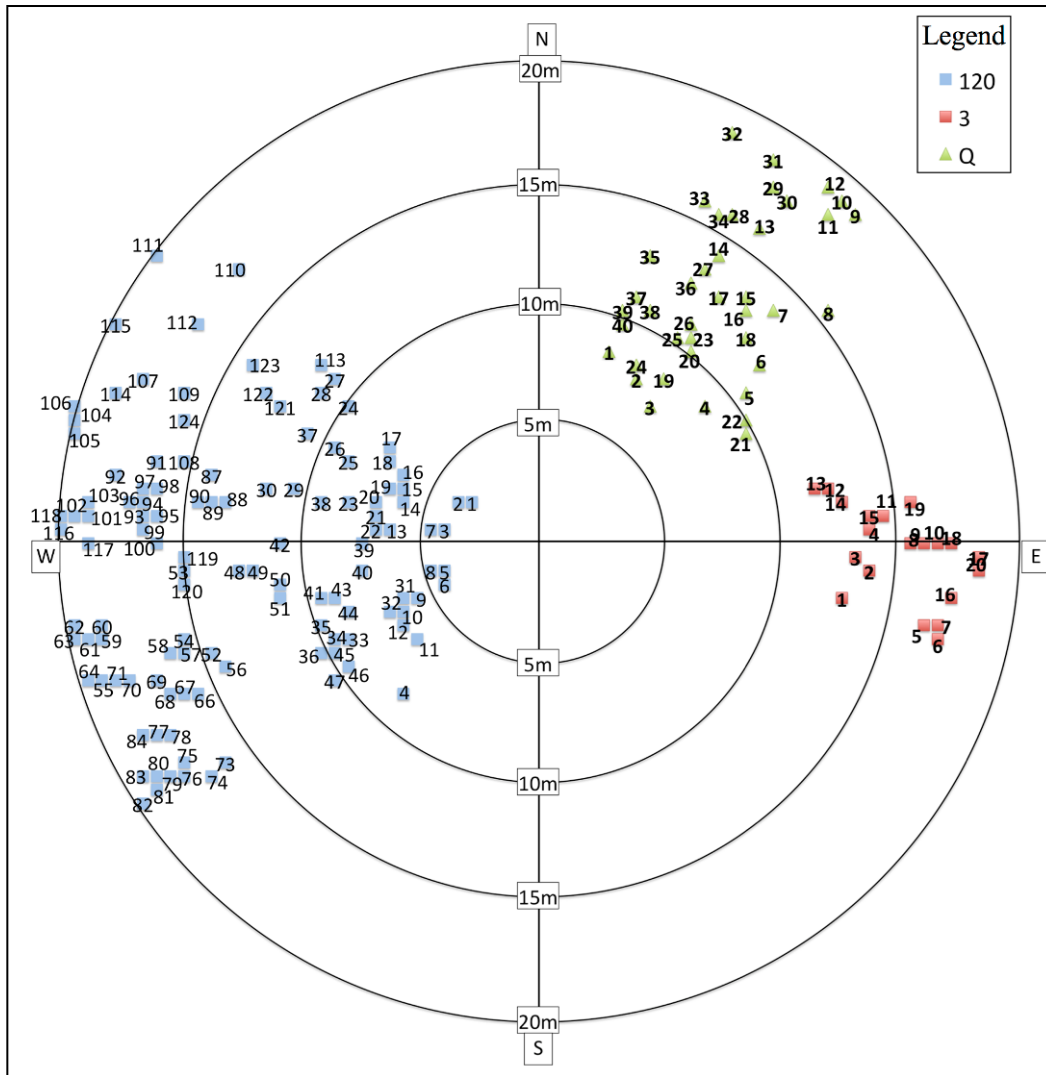


Figure 2.1. Map of tagged *P. tremuloides* stems in Ottawa NF plot 28. Phenotypically delineated putative clone “3” and clone “Q” are represented together with the 120 exhaustively sampled stems. Distances from the plot center are illustrated in 5, 10, 15, and 20-meter circles.

2.3 Stem selection

To compare clonal growth differences among genets under relatively similar environmental conditions, two adjacent morphologically differentiated clones (putative clone Q, 40 ramets; putative clone 3, 21 ramets) were tagged for assessment in August 2012 (Figure 2.1). For selection, both putative genets had to be contradistinctive in diameter and have a larger number of ramets than most other putative clones found within the plot. Additional analysis was conducted in the SW to NW quarter of plot 28, where 120 stems were exhaustively sampled to get an unbiased estimate of clonal size. The exhaustive sampling was also conducted to analyze inter- and intraclonal phenotypic variation in a relatively small area thought to be clonally diverse. Bole and leaf morphology and phenology were assessed in all sampled stems. All stems were genotyped using seven nuclear microsatellite loci.

2.4 Phenotypic analysis

Nine different morphological and phenological measurements were taken for the selected stems. Phenotypic measurements, in order of date collected, included average number of leaf serrations (ALS), average leaf length ratio (LR), Julian days till the start of vegetative budbreak (JDSBB), Julian days till the end of vegetative budbreak (JDEBB), the total number of Julian days between the start and end of budbreak (JDTBB), DBH, bark thickness (BT), bark brightness (BR), and leaf senescence (LS) (Table 2.1).

Table 2.1. List of phenotypic measurements and the months they were recorded for the 120 exhaustive sampling, putative clone Q, and putative clone 3 *P. tremuloides* stems in the Ottawa NF LTSP plot 28.

Phenotype	Trait	Date collected
Spring phenology	Julian days for total budbreak (JDTBB)	
	Julian days till start of budbreak (JDSBB)	May/June 2013
	Julian days till end of budbreak (JDEBB)	
Bole morphology	Diameter at breast height (DBH)	June, 2013
	Bark thickness (BT)	
	Bark brightness (BR)	August, 2013
Leaf phenology/morphology	Leaf senescence (LS)	September, 2013
	Average number of leaf serrations (ALS)	
	Average leaf length ratio (LR)	August, 2012

2.4.1 Leaf measurements-

Four leaves taken from the lowest branch of each tree collected for morphometric measurements in August 2012 were pressed and dried for at least two weeks, taped to printer paper and scanned into a pdf format. Two characters, LR and ALS, were measured for each leaf using Adobe Reader Version XI measurement tools. LR measured in centimeters (cm) using two length measures: length from the leaf base to the widest part of the leaf (length 1) and the widest part of the leaf to the leaf tip (length 2) (Figure 2.2). The total leaf length ratio for each leaf was calculated by dividing the length 1 by length 2 and then averaging that value from the four leaves for each stem. ALS was measured by counting the number of leaf serrations on the west side of every leaf then averaging across all leaves for the associated stem (Figure 2.2).

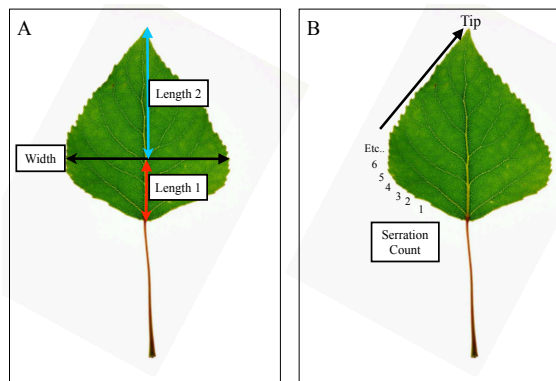


Figure 2.2. Leaf morphometric measurements applied to the selected *P. tremuloides* clonal colonies. Morphometric measurements included length measured in centimeters (A) and serration count (B). These were conducted with the upper epidermis facing upward. The two lengths were measured from the leaf base to the widest part of the leaf (length 1) and from the widest part of the leaf to the leaf tip (length 2). A leaf length ratio was calculated by dividing length1/length2 (A). Serrations were counted from the petiole to the tip on the west side of every leaf (B).

2.4.2 Visual analysis-

The timing of vegetative budbreak, bark color, DBH, and leaf senescence were collected for all tagged *P. tremuloides* stems in LTSP plot 28 during the summer of 2013. Timing of vegetative budbreak was scored once every one to two days in May and June using a four-stage reference chart with “0” representing no evidence of budbreak and “3” representing complete budbreak (Figure 2.3).

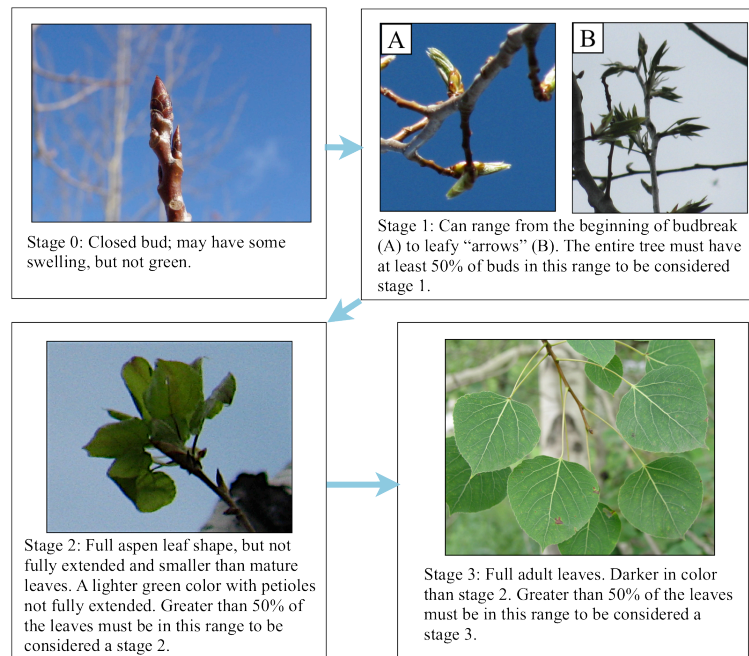


Figure 2.3. Vegetative budbreak chart applied to selected *P. tremuloides* clonal colonies to analyze spring phenology. The scale ranges from “0” representing an unopened bud to stage “3” representing complete budbreak

Bark and stem morphology were assessed between June and August 2013. DBH was measured using a Forestry Suppliers Inc. metric steel diameter tape in June. Tree bark was assessed for color and thickness in August. Bark brightness, or the deviation from white, was originally assessed as bark color using a colored chart created from 20 different paint swatches (Suppl. Figure 2). All bark colors were converted into bark brightness because the majority of the bark colors on all stems could be described in grey scale (Figure 2.4). To measure BT a Haglöf Barktax Bark Gauge, with increments in 1/10 of an inch, was used on the N, S, E, and W sides of each stem. Measurements were later converted to millimeters (mm).

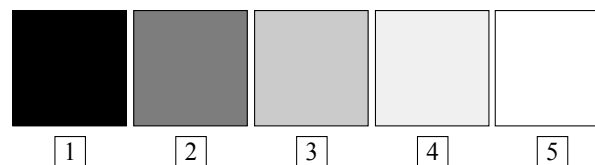


Figure 2.4. A grey scale chart created to measure bark “brightness.” This scale was based on a 20-color chart, which was reduced because the bark colors present among sampled *P. tremuloides* stems existed mostly in “grey scale.”

The final phenotypic assessment was leaf senescence (LS) and occurred on 17th September 2013. Senescence for each selected stem was estimated with a scalable color chart based on the percentage of green, yellow, and brown visible in canopy leaves (Figure 2.5). The color chart contained nine senescence categories, labeled 1-9, where a 1 represents no senescence (all leaves green), a 5 represents all yellow leaves, and a 9 represents all brown leaves. Senescence stages 2, 3, and 4 represent 4:1, 1:1, and 1:4 proportions of green to yellow leaves, respectively. Stages 6, 7, and 8 represent 4:1, 1:1, and 1:4 proportions of yellow to brown leaves, respectively.

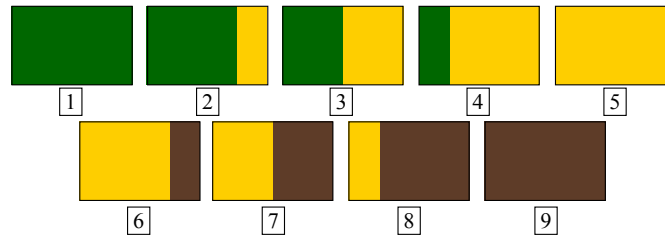


Figure 2.5. A color chart created to measure leaf senescence in nine stages (labeled 1-9) of *P. tremuloides* clonal colonies measured on September 17th. The chart represents the proportions of green, yellow, and brown visible in the canopy for each tree. Stage 1 represents a crown with all green leaves and no senescence, stage 5 represents a crown with all yellow leaves, and stage 9 represents a crown with all brown leaves. Stages 2, 3, and 4 represent 4:1, 1:1, and 1:4 proportions of green to yellow leaves, respectively. Stages 6, 7, and 8 represent 4:1, 1:1, and 1:4 proportions of yellow to brown leaves, respectively.

2.5 Genotypic clone identification

2.5.1 *Sample selection and collection procedures-*

Leaf material for DNA isolation was collected in August 2012 and July 2013. Limb cutters were used to cut the lowest branch of each tree, and up to 15 leaves were collected from each cut limb. Leaf samples were placed in Ziploc bags with labels corresponding to the sampled stem and stored at -20°C until DNA extraction could occur.

2.5.2 *Disruption and isolation of genomic DNA-*

A small piece of one leaf, approximately 5mm×5mm, was cut for each sample and placed in a tube with a tungsten carbide bead (Qiagen) for disruption. The rest of the leaf was stored in a 1.5 mL vial at -20°C. Samples were flash frozen in liquid nitrogen for 30 seconds before disruption in a Qiagen Retsch TissueLyser (MM300). All samples were disrupted at 20 hertz for 60 seconds, flash frozen in liquid nitrogen for another 30 seconds, then disrupted a second time.

Once disrupted, DNA was isolated from 207 trees using either Qiagen's DNeasy 96 Plant Isolation kit, or DNeasy Plant Mini Kit following the manufacture's protocol.

DNA quality and quantity were checked on 0.8% agarose gels at 100V for 40 min, stained with GelRed Nucleic Acid Stain (Phenix Research Products) and visualized under UV light. A 100-1000bp ladder (DNAL 1000-100 ready-to-load quantitative DNA ladder, Phenix Research Products) was used as a size marker.

2.5.3 PCR procedures-

Clones were identified using seven nuclear microsatellite loci, also known as nuclear simple sequence repeats (nSSRs). These were selected from a total of 20 different nSSRs originally developed for *P. nigra*, *P. tremuloides*, *P. deltoides*, and *P. trichocarpa* (Van Der Schoot et al. 2000, Smulders et al. 2001, DeWoody et al. 2009). Selection was based on amplification success, easy readability, polymorphism, and repeatability of results. The nSSR primers used for analyses were WPMS 14, WPMS 15, WPMS 16, and WPMS 20 originally developed for *P. nigra* (Van Der Schoot et al. 2000, Smulders et al. 2001, DeWoody et al. 2009), and GCPM_970-1, PMGC 576 and PMGC 2571 originally developed for *Populus* spp. (The International *Populus* Genome Consortium, http://www.ornl.gov/sci/ipgc/ssr_resource.htm). The following 15µL PCR mixture was prepared for each DNA sample: 3µL of HOTFIREpol master mix from Oak Biotechnologies (0.6 U of HOTFIREpol Taq polymerase, 10mM MgCl₂, 2mM of each dNTPs), 2µL (5 µM) of each forward and reverse primer, 6µL of ddiH₂O, and 2µL (~2ng) of DNA. All samples were amplified in the Peltier Thermal cycler GeneAmp PCR system 2700 (Applied Biosystems).

The following PCR program was used for primers WPMS14, WPMS 15, WPMS 16, WPMS 20, and GCPM_970-1: (1) initial denaturation at 94°C for 15min; (2) 35 cycles at 94°C for 30secs, annealing between 50-70°C for 40secs, and extension at 72°C for 40secs; (3) final extension at 72°C for 15min.

For primers PMGC-2571 and PMGC-576 the following PCR protocol was modified from De Woody et al. (2009): (1) initial denaturation at 95°C for 15 min; (2) 4 cycles at 94°C for 30sec, annealing starting at 58°C (PMGC 2571) or 55°C (PMGC-576) for 30sec, and extension at 72°C for 1min; (3) 4 cycles at 94°C for 30sec, annealing at 52°C for 30sec, and extension at 72°C; (4) 25 cycles at 94°C for 30sec, annealing at 50°C for 30 seconds, and extension at 72°C for 1min; (5) final extension at 72°C for 15min. Exact annealing temperatures for all primers are given in Table 3.1 and can be found within the results section (pg. 25).

2.5.4 Capillary electrophoresis-

Hi-Di™ formamide (Applied Biosystems) and 0.15µL of GeneScan 500 LIZ size standard (Applied Biosystems) was added to PCR products (1µL, 1:10 diluted) before separation on an ABI 3730xl DNA Genetic analyzer (Applied Biosystems). Fragments were scored using GeneMapper v.4.0 (Applied Biosystems).

2.6 Data analyses

Analytical data processes described here were chosen to identify and delineate clonal groups, identify related genets, and compare phenotypic variation within and among clonal colonies. The statistical programs used to accomplish these tasks were GenAlEx version 6.5 (Peakall and Smouse 2012), XLSTAT version 2013.5.4 (Famy and Aubry 2013), SPSS version 22 (IBM 2013), and PC-Ord version 5 (McCune and Mefford

2006). Formulas for the statistical procedures are described in instances where the procedure is uncommon in ecological studies, and to further distinguish between the three ordination methods utilized in this study. Genotypic analyses were computed in GenAlEx. Phenotypic intra- and interclonal analyses were computed in XLSTAT, SPSS, and PC-Ord.

2.6.1 *GenAlEx- genetic analyses-*

A total of eleven different statistical measures were calculated with GenAlEx. These included the probability of identification (PI), probability of identification among siblings (PI_{sib}), expected and observed heterozygosity (H_e and H_o , respectively), effective number of alleles (N_e), clonal identification through multilocus matches, genotypic probability (P_{gen}), probability of a second genotype encounter (P_{se}), principal coordinate analysis (PCoA), and the Lynch and Ritland (1999) pairwise relatedness measure.

2.6.1.1 -Probability of identification-

The PI and PI_{sib} statistical measures were used in conjunction to determine the optimal number of loci for accurate clonal identification (Taberlet and Luikart 1999, Waits et al. 2001, Peakall et al. 2006, Peakall and Smouse 2006). Both measures calculate the probability of the same allelic pattern occurring in two different genotypes at the given microsatellite loci by chance. However, PI measures identity among a general population, whereas PI_{sib} measures identity among potentially related individuals (i.e. siblings or parents and offspring). The following equations (Equations 1 and 2) are those used to calculate PI and PI_{sib} in GenAlEx, where p_i is the frequency of the i^{th} allele:

$$\text{Equation 1: } PI = 2(\sum p_i^2)^2 - \sum p_i^4$$

$$\text{Equation 2: } PI_{sib} = 0.25 + (0.5 \sum p_i^2) + [0.5(\sum p_i^2)^2] - (0.25 \sum p_i^4)$$

The chance threshold of two different genotypes sharing a similar allelic pattern for PI and PI_{sib} has to be less than 0.005. When this criterion was met, the number of markers for estimating population diversity and clonal identification was considered adequate (Taberlet and Luikart 1999, Waits et al. 2001, Peakall et al. 2006, Peakall and Smouse 2006).

2.6.1.2 -Parameters of genetic variation-

Genetic diversity between genotyped individuals was measured using H_e , H_o , and N_e statistical procedures. H_e and H_o are used in conjunction to estimate population diversity by determining the expected and observed frequencies of heterozygous individuals for a specific locus (Hartl and Clark 1997, Peakall and Smouse 2006). H_e is calculated (Equation 3) using squared allelic frequencies, p_i^2 , and H_o is calculated

(Equation 4) using the total number of observed heterozygotes for one locus divided by the population size (n):

$$\text{Equation 3: } H_e = 1 - \sum p_i^2$$

$$\text{Equation 4: } H_o = (\text{No. of Hetero})/n$$

If H_o is lower than H_e across all or most loci, inbreeding is prevalent within a population. Alternatively, if H_o is higher than H_e , genetic diversity is higher than what would be expected from random mating (Hartl and Clark 1997, Peakall and Smouse 2006).

N_e is calculated (Equation 5) for each locus using H_e and estimates the “ideal” allelic frequency (Kimura and Crow 1964, Brown and Weir 1983, Peakall and Smouse 2006):

$$\text{Equation 5: } N_e = 1/(1-H_e)$$

This estimate essentially represents the number of alleles at a specific locus that would be maintained if all individuals were homozygous. N_e allows for a more robust allelic comparison between loci than just the total number of alleles (N_a) present at each locus (Kimura and Crow 1964, Brown and Weir 1983, Peakall and Smouse 2006). These procedures were used to accurately assess both heterozygosity and allelic diversity among sampled individuals.

2.6.1.3 -Clonal identification-

Once a sufficient number of polymorphic loci were identified, all samples were analyzed for genetic matches and assigned either a letter, representing a genotype with more than one ramet per genet, or a number, indicating genets with only one ramet. Samples missing two or more loci due to amplification errors were removed from further analysis. Another consideration when conducting clonal assignment within GenAlEx is that the probability of allelic differentiation caused by somatic mutations is not computed into genet delineation. Because somatic mutations within microsatellites occur periodically (4×10^{-5} to 6×10^{-7} per locus per cellular division at certain *P. tremuloides* loci (Ally et al. 2008)), care must be taken when interpreting the results of such clonal assignment (Ellegren 2000, Ally et al. 2008).

There are two basic types of somatic mutations that may occur: insertions/deletions (indels) and homologous recombination (HR) (Lian et al. 2004, Boyko et al. 2006, Ally et al. 2008, Delacôte and Lopez 2008, Yao and Kovalchuk 2011). Indels may be caused by direct damage to the DNA itself, by the “slippage” of DNA polymerase during cellular mitosis, or by non-homologous end joining repair of a double stranded break (DSB) during mitosis, all of which can lead to an increase or decrease in

microsatellite length (Lian et al. 2004, Boyko et al. 2006, Ally et al. 2008, Delacôte and Lopez 2008, Yao and Kovalchuk 2011). In contrast, HR repair, which occurs during mitosis in the presence of a DSB, may result in a loss of heterozygosity for the given locus (Lian et al. 2004, Boyko et al. 2006, Ally et al. 2008, Delacôte and Lopez 2008, Yao and Kovalchuk 2011). HR resulting in homozygosity arises from a process called gene or genetic conversion where a damaged allele is replaced by its allelic variant at the same locus during mitosis. For these reasons, two genotypes differentiated by one allele at one locus, when all other loci contained the same allelic pattern, were considered to be part of the same genet and reassigned under the same clonal group. Additional somatic mutations may have been present at these loci, but not enough markers were used to further differentiate between somatic and zygotic mutations outside of one allelic variant (Lian et al. 2004, Boyko et al. 2006, Ally et al. 2008, Delacôte and Lopez 2008, Yao and Kovalchuk 2011). Six of the largest clonal colonies were selected for statistical analyses.

A map of all genetically identified genets was plotted in Microsoft Excel 2011 (Suppl. Figure 3). A second map was also created that contained only the six clonal colonies selected for statistical analyses (see Figure 3.1 in the results section). The area of each selected clonal colony was determined using the measurement tools available in Adobe Reader Version XI.

2.6.1.4 -Accuracy of clonal identification-

P_{gen} and P_{se} were used to determine the likelihood of genetically identical individuals developing through zygotic formation (Parks and Werth 1993, Mayes et al. 1998, Sydes and Peakall 1999). The first measure, P_{gen} , calculates the probability of ramets from the same genet containing identical allelic patterns arising from random mating. Calculations for this procedure are represented in Equation 6, where p_i is the frequency of each allele and h is the number of heterozygous loci:

$$\text{Equation 6: } P_{\text{gen}} = (\prod p_i) 2^h$$

The second measure, P_{se} , calculates the chance of an observed genet occurring twice or more among all identified genets as a result of random mating. P_{se} is calculated from obtained P_{gen} values (Equation 7) where n is number of identified genets:

$$\text{Equation 7: } P_{\text{se}} = 1 - (1 - P_{\text{gen}})^n$$

Both procedures are correlated to a population's heterozygosity and allelic diversity. If genetic diversity within a population is low, genet delineation through even the most robust molecular markers may be difficult to ascertain. Values for P_{gen} and P_{se} must be less than 0.05 for statistical confidence in clonal assignment (Parks and Werth 1993, Mayes et al. 1998, Sydes and Peakall 1999).

2.6.1.5 -Principal coordinate analysis-

The PCoA was computed using a genetic distance matrix created from given allelic data to graphically represent genetic relatedness in a two-dimensional plane (Grower 1966, Borcard and Legendre 2002, McCune and Grace 2002, Peakall and Smouse 2006, Famy and Aubry 2013). A PCoA was computed here in place of a principal component analysis because allelic data are non-Euclidean measures. Equation 8 was used to create a cross-products matrix δ from the distance matrix A , where non-scalable distances are converted to similarities that act like, or pretend to be, Euclidean distances. Here, a_{ij} represents the i_{th} row in the j_{th} column and “d” is the symmetric matrix $p \times p$, where p represents the variables of interest (in this case microsatellite DNA), which contain the distances between the p elements:

$$\text{Equation 8: } a_{ij} = (-d_{ij}^2) / 2$$

The cross-products matrix was centered (Equation 9) for the $\Delta 1$, where $\delta 1$ units are given as \bar{a}_i representing the mean of the i_{th} row, \bar{a}_j representing the mean of the j_{th} column, and \bar{a} representing the mean of all elements:

$$\text{Equation 9: } \delta 1_{ij} = a_{ij} - \bar{a}_i - \bar{a}_j - \bar{a}$$

The centered cross-products matrix was used to find eigenvalues (i.e. roots of the polynomials) where the eigenvalue is represented by lambda (λ) in Equation 10 (i.e. a “characteristic equation”) and where \mathbf{I} represents the identity matrix:

$$\text{Equation 10: } |\delta 1 - \lambda \mathbf{I}| = 0$$

Eigenvectors were calculated (Equation 11) in a similar manner where u represents the eigenvector that is part of the $p \times p$ eigenvector matrix:

$$\text{Equation 11: } [\delta 1 - \lambda \mathbf{I}]u = 0$$

From this process of variable reduction into eigenvectors, it becomes possible to visually identify clustering of related individuals (Grower 1966, Borcard and Legendre 2002, McCune and Grace 2002, Peakall and Smouse 2006, Famy and Aubry 2013). The PCoA was used in conjunction with calculations provided by the relatedness measures to analyze phenotypic variation in relation to genetic distance.

2.6.1.6 -Relatedness measures-

Of the three available pairwise relatedness measures available within GenAlEx (Queller and Goodnight (1989), Ritland (1996), and Lynch and Ritland (1999) pairwise

relatedness measures), the most appropriate for this data was the Lynch and Ritland (1999) pairwise relatedness measure. All estimates use the methods-of-moments (MM) to determine the probability of similar allelic patterns occurring by descent between two individuals. However, Lynch and Ritland (1999) is a more suitable estimate when the markers utilized are highly polymorphic (≥ 6 alleles per locus) and often heterozygous (Ritland 2000).

The pairwise relatedness measures were calculated using Equation 12 (Lynch and Ritland 1999, Ritland 2000):

$$\hat{r} = \frac{p_i(\delta_{jk} + \delta_{jl}) + p_j(\delta_{ik} + \delta_{il}) - 4p_i p_j}{(1 + \delta_{ij})(p_i + p_j) - 4p_i p_j}$$

Here the estimated relatedness (\hat{r}) is calculated for each locus by comparing alleles A_i and A_j from one individual to alleles A_k and A_l from the second individual. If an individual is homozygous at the particular locus $\delta_{ij}=1$, but when heterozygous $\delta_{ij}=0$. The same rule is followed when the two individuals are compared: if A_i and A_l are the same then $\delta_{il}=1$ and if they are different $\delta_{il}=0$. The frequencies of the i_{th} and j_{th} allele are represented by p_i and p_j , respectively. Each locus is given a weight, which is the inverse of the variance, and relatedness between the individuals is then averaged across all loci, giving a final relatedness measure that is then multiplied by two. The relatedness measure gives two asymmetric relatedness estimator values (LR1 and LR2) and the mean relatedness estimator value based on LR1 and LR2 (LRM). The pairwise relatedness value is represented on a -1 to 1 scale, where unrelated pairs=0, half siblings=0.5, and full siblings=1 (Lynch and Ritland 1999, Ritland 2000, Hardy 2003). Any pairwise values less than zero demonstrate that the two individuals are unrelated as either half siblings or full siblings (Hardy 2003).

Using the pairwise relatedness measures and PCoA, phenotypic distance and genetic distance between clonal colonies was compared.

2.6.2 XLSTAT- phenotypic analyses-

To process collected phenotypic data, eight statistical procedures were utilized in XLSTAT. These included tests of significance for each variable and for the deterministic value of significant variables in clonal delineation. Tests of variable significance included analysis of multicollinearity between variables, a Pearson's correlation between variables, k-sample comparisons of variance, and a one-way analysis of variance (ANOVA). Here it should be noted, that while the one-way ANOVA was conducted mainly in XLSTAT, a post-hoc test was conducted in SPSS but will be described under the XLSTAT ANOVA analysis for continuation purposes.

To find the deterministic value of the significant variables, two sets of statistical analyses were performed. The first set of analyses tested how well a stem's phenotypic value fits within each genetically identified clonal colony. These included discriminant analysis (DA) and a cluster and regression dendrogram.

The second set of analyses tested whether accurate clonal identification through phenotypic assessment is possible, given the assumption that the number and size of

clonal colonies is unknown. These tests included principal component analysis (PCA) and agglomerative hierarchical clustering (AHC).

2.6.2.1 -Multicollinearity and correlation analysis-

The R^2 value, tolerance ($1-R^2$), and variable inflation factor (VIF calculated as $1/(1-R^2)$) were calculated to identify multicollinearity between all phenotypic characteristics. The thresholds used here for expected multicollinearity were an R^2 value greater than 0.80, a tolerance less than 0.20, and a VIF greater than four (O'Brien 2007, Famy and Aubry 2013).

If the variables (characteristics) met or exceeded at least one of these thresholds, excessive multicollinearity was assumed and a Pearson's correlation coefficient was used to further identify linearly related variables. For any variables that might be collinear, all but one of those variables would be removed from further statistical analysis. The remaining characteristics were again tested for multicollinearity to confirm that collinearity had been reduced.

2.6.2.2 -K-sample comparison of variance and ANOVA (with post-hoc test in SPSS)-

A Levene's mean k-sample comparison of variances test was performed to test for interclonal homoscedasticity of the phenotypic variables (Brown and Forsythe 1974). The results of the Levene's statistic demonstrated that the interclonal variances showed a heteroscedastic rather than a homoscedastic trend.

Due to the heteroscedastic nature of the clonal colonies, both a Welch and Brown-Forsythe test were utilized to determine whether the variables were significantly different between clones. The Welch and Brown-Forsythe tests adjust the F-ratio by separating the sample variances or weighting the group variances, respectively, rather than using a pooled variance for a standard ANOVA (Clinch and Keselman 1982). In this way, these procedures are able to accommodate unequal variances. However, neither ANOVA procedure is able to determine which clones are significantly differentiated from one another, so a post-hoc pairwise comparison was conducted using the significant variables (Hilton and Armstrong 2006).

While post-hoc tests are available in XLSTAT, this program does not include post-hoc tests that accommodate unequal variances and sample sizes (Famy and Aubry 2013). For this reason, the post-hoc tests were conducted in SPSS to adjust for the deviations from standard ANOVA assumptions. SPSS offers four post-hocs that account for one or both ANOVA deviations, which include the Dunnett's C and T3, Games-Howell, and Tamhane's T2 test (Mooi and Sarstedt 2011). Of these, Tamhane's T2 fit this data the best for three reasons: 1) the degrees of freedom were not large enough for a Dunnett's T3 or C, 2) both clonal variance and size were unequal ruling out both Dunnett's C and T3, and 3) a more conservative procedure was needed to reduce Type I statistical error, ruling out Games-Howell (Tamhane 1977, Dunnett 1980, Dunnett and Tamhane 1992, De Muth 2006, Mooi and Sarstedt 2011).

The following statistical procedures require greater interclonal separation for the results to be statistically accurate. Therefore, in order to further utilize a variable, more

than one clonal colony had to be significantly differentiated under a Tamhane's T2 post-hoc test. If a variable did not meet these qualifications, it was removed from further analyses.

2.6.2.3 -Discriminant analysis-

A DA ordination was used to visually map clonal separation in a 2-D space based on the selected phenotypic characteristics. This differs from a PCA and PCoA in that it graphs data based on pre-defined groups, or, in this instance, pre-defined clonal colonies (Lachenbruch and Goldstein 1979, Klecka 1980, McCune and Grace 2002, Famy and Aubry 2013). A DA will also minimize within group distances, whereas a PCA/PCoA maximizes distances for all samples. Another important difference is that a DA cannot use a correlation matrix and must use a variance or covariance matrix.

Creating a covariance or variance matrix for each group, or clone, was the first step used in conducting the DA. The second step required the creation of a "pooled" matrix (δ_p) from all of the group matrices followed by calculating the interclonal variance for each variable (δ_g). The F-ratio (F) was maximized by finding \mathbf{u} (Equation 13), where \mathbf{u} is the eigenvector representative of a specific discriminant function:

$$\text{Equation 13: } F = (\mathbf{u}' \delta_g \mathbf{u}) / (\mathbf{u}' \delta_p \mathbf{u})$$

The eigenvalues were found using a characteristic equation as shown in Equation 14, which is similar to Equation 10 used for the PCoA:

$$\text{Equation 14: } |\delta_p^{-1} \delta_g - \lambda I| = 0$$

The eigenvectors were found with Equation 15, which is again similar to Equation 11 used to calculate the PCoA (Lachenbruch and Goldstein 1979, Klecka 1980, McCune and Grace 2002, Famy and Aubry 2013):

$$\text{Equation 15: } [\delta_p^{-1} \delta_g - \lambda I] \mathbf{u} = 0$$

While a DA can be used to define within group agreement (McCune and Grace 2002), it is only used here as a visual model for within group agreement because of the heteroscedasticity discovered through the Levene's mean test. A cluster and regression dendrogram, discussed next, and a multi-response permutation procedure (MRPP), conducted in PC-Ord, were used instead to define intraclonal agreement.

2.6.2.4 -Classification and regression tree-

A classification and regression tree was calculated using an exhaustive chi-squared automatic interaction detection procedure, otherwise known as an exhaustive

CHAID. This procedure performs multi-level splits while controlling for a type I error rate with a Bonferroni correction (Hill and Lewicki 2006). The basic procedures involved in an exhaustive CHAID included three steps: splitting, merging, and stopping. During the first step, splitting, a Tschuprow's T test is used due to the qualitative nature of this clonal data (Bergsma 2012). The Tschuprow's T (Equation 16) is a method similar to Pearson's chi-Squared test, where Φ^2 represents the Pearson's cumulative test statistic divided by n, and R and C come from the R×C contingency table:

$$T = \sqrt{\frac{\phi^2}{\sqrt{(R-1)(C-1)}}}$$

The CHAID selects the best split character or variable for that node using the lowest p-value provided by Tschuprow's T (Bergsma 2012, Famy and Aubry 2013). Merging of similar stems occurs when the largest value is greater than the alpha threshold ($\alpha=0.05$). The merging step is continuously repeated and only stops when p-values are less than or equal to the alpha level, or when only two categories are left (Bergsma 2012, Famy and Aubry 2013). Stopping follows when one or all of the stopping criteria are met (as listed in the XLSTAT user guide (Bergsma 2012, Famy and Aubry 2013)).

In using this method, inter- and intraclonal comparisons were easily made. With the CHAID used in combination with the visual aid provided by a DA and a total variable comparison between clones made by the MRPP (discussed later), a complete and robust analysis of clonal phenotypic characteristics was made.

2.6.2.5 -Principal component analysis-

A PCA ordination was created from a Pearson's correlation matrix to graphically summarize selected phenotypic variables in a two-dimensional space (Jolliffe 2002, McCune and Grace 2002, Famy and Aubry 2013). The PCA was based on a correlation matrix rather than a variance/covariance matrix because each variable was measured using a different criterion (i.e. days, cm, color, etc.). Once the correlation matrix was created, the PCA followed the same procedures as the PCoA for Equations Equation 10 and Equation 11.

As with a DA, clonal separation by phenotypic variables can be visually detected. Since a PCA does not include preconceived groups within its calculations, this type of ordination removes the bias that exists in a DA. Patterns, however, are often difficult to discern because a PCA also maximizes distances between points (Jolliffe 2002, McCune and Grace 2002, Famy and Aubry 2013). The agglomerative hierarchal clustering (AHC) method solves this problem by using a dendrogram to organize stems into phenotypic groups.

2.6.2.6 -Agglomerative hierarchal clustering-

AHC is commonly used in ecological studies for the creation of dendrogram groupings based on community data (McCune and Grace 2002). Here, however, the groups (genets/clonal colonies) were already established. Instead the goal was to

determine if stems could be separated into the correct clonal groups based on phenotypic characteristics, while removing prior bias concerning the number and size of clonal colonies (Steinbach et al. 2000, McCune and Grace 2002, Famy and Aubry 2013).

A Bray-Curtis dissimilarity matrix was first constructed before running the AHC in XLSTAT. The procedure selected to create the AHC dendrogram from the constructed matrix was a Bray-Curtis dissimilarity method with a flexible-linkage of -0.25. The dendrogram was scaled to 0.05 for defining the groups. Several other methods were tested, but this created the most accurate clustering of stems to their respected clonal colony.

2.6.3 PC-ORD- MRPP phenotypic analysis-

Only one statistical measure, a multi-response permutation procedure (MRPP) was conducted in PC-Ord (McCune and Grace 2002). The MRPP is a nonparametric measure not found in many statistical programs. It allows for a robust, interclonal comparison of all variables in a way that is similar to a multivariate analysis of variance (MANOVA), but avoids expectations involving sample distributions (McCune and Grace 2002).

The first step involved creating a Bray-Curtis dissimilarity matrix then calculating average individual intraclonal distances. The means of each clonal colony were weighted (C) based on size using a “natural weighting” (Equation 17), where n_i is the number of stems in each colony (Mielke Jr 1984, McCune and Grace 2002):

$$\text{Equation 17: } C_i = n_i / (\sum n_i)$$

A total intraclonal mean distance, or observed delta (δ), was calculated from the selected clones using C_i (Equation 18), where g represents the clones and X_i represents phenotypic distances between clones (McCune and Grace 2002):

The chance-corrected within-group agreement (A) was determined (Equation 19) using the expected delta (m_δ) and observed delta:

$$\text{Equation 19: } A = 1 - (\delta / m_\delta)$$

Both the observed delta statistic and A-value is the total variation for all clones based on individual intraclonal distances. The smaller the individual distance value, the greater the individual intraclonal agreement. For total group values, when the observed delta=0 and A=1, complete homogeneity exists among the groups. Greater heterogeneity than what can be explained by chance is exhibited when the observed delta is greater than the expected delta and when A is less than 0.

In addition to analyzing clonal homogeneity, an MRPP also conducts a pairwise comparison to determine which colonies are separate from one another. The A-value is additionally used to determine the pairwise phenotypic comparisons. In this instance, the A-value represents the intraclonal agreement when compared to another group. If the two groups are phenotypically similar and therefore overlapping, the pairwise A-value will be smaller. If the A-value is larger, the two groups are more separate and, thus, show greater intraclonal agreement. The pairwise interclonal distance value, or pairwise intraclonal A-value, should not be confused with the individual distance value. Contrary to the pairwise A-value, smaller individual distance values indicate greater intraclonal agreement when not compared to other clonal colonies. The A-value is also different from a simple one-way ANOVA in that it uses all variables to make pairwise comparisons rather than one variable at a time (McCune and Grace 2002).

3 Results

3.1 Genetic analysis

3.1.1 Primer analysis-

The seven nuclear microsatellite markers selected demonstrated an overall satisfactory rate of genetic delineation with a PI and PI_{sib} of 3.6×10^{-7} and 3.3×10^{-3} , respectively (Table 3.1). The average number of alleles (N_a) across all loci was 10 with an average number of effective alleles (N_e) being 3.97 for all discovered genotypes. Total observed heterozygosity (0.70) was slightly higher than the expected heterozygosity (0.68).

Table 3.1. Table of the primers applied in this study. Included is information regarding the specific nature of the primers (SSR motif and annealing temperature) and statistical information regarding the choice of primer loci. The statistical data include the observed number of alleles (N_a), the effective number of alleles (N_e), the observed heterozygosity (H_o), the expected heterozygosity (H_e) probability of identification (PI), and probability of identification among siblings (PI_{sib}).

Locus	SSR motif	Annealing Temp (°C)	Allelic range (bp)	PI	PI_{sib}	N_a	N_e	H_o	H_e
WPMS-14	(CGT) ₂₈₋₃	66.8	197-233	1.4E-01	4.4E-01	9	3.34	0.61	0.70
WPMS-15	(CCT) ₁₄₋₃	60	190-202	2.0E-01	4.9E-01	6	2.58	0.61	0.61
WPMS-16	(GCT) ₈ (ATCCTC) ₅	60	170-199	2.7E-01	5.5E-01	7	2.16	0.66	0.54
WPMS-20	(TTCTGG) ₈	60	206-230	2.5E-01	5.5E-01	7	2.13	0.66	0.53
GCPM_970-1	(TGC) ₁₀	53.8	116-133	1.9E-01	4.8E-01	8	2.78	0.84	0.64
PMGC-2571	GA	Initial- 58	84-117	4.0E-02	3.4E-01	15	6.39	0.92	0.84
		Second- 52							
		Third- 50							
PMGC-576	GA	Initial- 55	149-208	2.6E-02	3.2E-01	18	8.38	0.59	0.88
		Second- 52							
		Third- 50							
Across loci	-	-	-	3.6E-07	3.3E-03	10	3.97	0.70	0.68

3.1.2 Clonal analysis-

In total, 32 unique genets were identified from the 181 sampled stems (Suppl. Figure 3) using microsatellite markers. The average size of a genet was 5.5 ramets with 11 genets containing only one verified ramet. The other 21 genets were found to exist in (clonal) colonies of two or more ramets. Colonies of these genets could have been larger, but are sized only from stems that have been genetically identified.

The six largest genetically identified clonal colonies used in further analyses were BB, CC, E, G, M, and U containing 13, 11, 25, 11, 32, and 16 ramets, respectively (Table 3.2, Figure 3.1). Clonal colonies BB, CC, G, M, and U were all found in the NW to SW quarter of plot 28 within the 120 exhaustive sampling and clonal colony E found in the N to NE quadrant within putative clone Q. No clonal colony from putative clone 3 was large enough to use. Spatially, clone M was the largest covering 163.8 m² and clone CC the smallest covering only 18.9 m² (Table 3.2, Figure 3.1). Clone CC contained the most stems/m², with 0.58 stems/ m², and clone G the least stems/m², with 0.12 stems/m². Including the spatial overlap among clonal colonies, the total area covered by the selected clonal colonies was 467.93 m². Without including spatial overlapping, the total area

covered by selected clonal colonies was 249.44 m². The sampling area encompassed 448.32 m² of the 1256 m² plot.

Table 3.2. The probability that *P. tremuloides* ramets in each clonal colony (BB, CC, E, G, M, and U) arose from sexual rather than asexual reproduction. Included is the area in meters each clone covers and the number of stems/m². N=number of ramets; P_{gen}=probability of the same genotypes belonging to different clones; P_{se}=probability of an observed genotype occurring more than once among identified genotypes.

Clone	N	Area (m ²)	Stems/m ²	P _{gen}	P _{se}
BB	13	29.83	0.44	1.22E-06	1.32E-04
CC	11	18.91	0.58	1.22E-08	1.32E-06
E	25	78.40	0.32	1.46E-06	1.57E-04
G	11	94.63	0.12	1.33E-07	1.44E-05
M	32	163.80	0.20	3.21E-05	3.46E-03
U	16	82.36	0.19	3.29E-07	3.55E-05

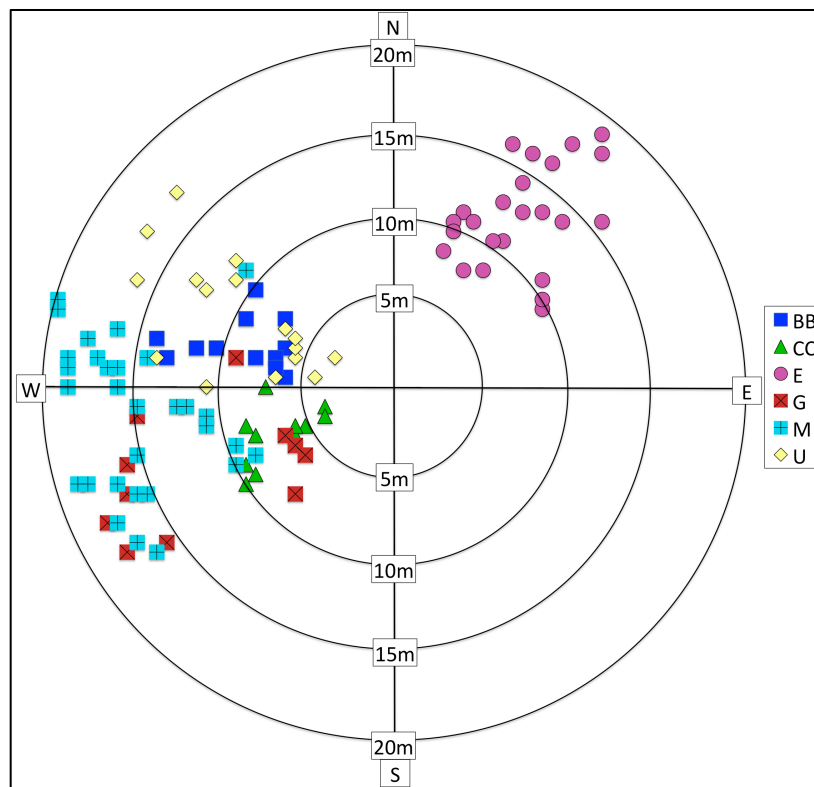


Figure 3.1. The LTSP Ottawa NF plot 28 map of the selected *P. tremuloides* clonal colonies (BB, CC, E, G, M, and U). Distances are given in 5-meter increments from plot center. N, E, S, and W represent cardinal directions.

The P_{gen} and P_{se} values for the selected clones ranged from 3.21×10⁻⁵ to 1.22×10⁻⁸ and 3.46×10⁻³ to 1.32×10⁻⁶, respectively, confirming that the colonies were formed through root-sprouting, or asexual reproduction, rather than sexual reproduction (Table 3.2). When the six clonal colonies were analyzed for relatedness using both a PCoA (Figure 3.2) and the Lynch and Ritland (1999) pairwise relatedness measure (Table 3.3), none of the clones were closely related. Therefore, further analyses between phenotypic traits and relatedness among clones were not made.

Table 3.3. Pairwise genetic relatedness between selected *P. tremuloides* clonal colonies (BB, CC, E, G, M, and U). Results were obtained using the Lynch and Ritland (1999) relatedness measure in GenAlEx. Values near or less than zero indicate no genetic relatedness, values around 0.5 indicate half-siblings, and values around 1 indicate full-siblings. LR1 and LR2 represent the asymmetric estimator values and LRM represents the mean estimator value.

Clone 1	Clone 2	LR1	LR2	LRM
BB	CC	-0.044	-0.079	-0.062
BB	E	-0.448	-0.256	-0.352
BB	G	-0.177	-0.226	-0.202
BB	M	-0.246	-0.181	-0.214
BB	U	0.000	0.000	0.000
CC	E	-0.241	-0.249	-0.245
CC	G	-0.410	-0.290	-0.350
CC	M	-0.158	-0.209	-0.183
CC	U	-0.268	-0.467	-0.368
E	G	-0.291	-0.212	-0.252
E	M	-0.078	-0.100	-0.089
E	U	-0.347	-0.205	-0.276
G	M	-0.256	-0.241	-0.249
G	U	0.135	0.167	0.151
M	U	-0.353	-0.269	-0.311

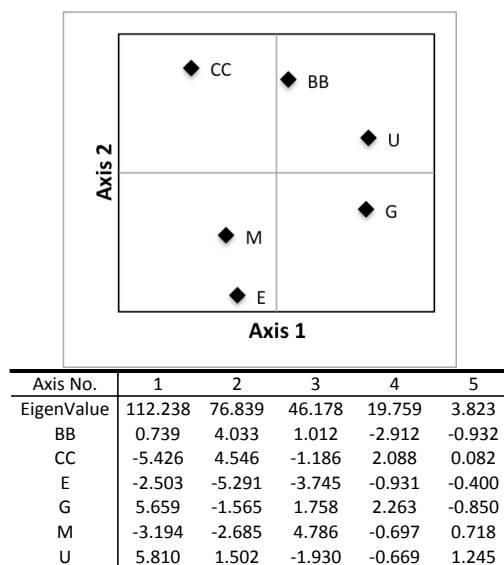


Figure 3.2. A PCoA graphical representation of genetic relatedness among selected *P. tremuloides* clonal colonies (BB, CC, E, G, M, and U). Eigenvalues for each axis and the coordinates of each clone are given below the graph.

3.2 Phenotypic analysis

Several stems within the selected clonal groups were removed from additional analyses because of missing phenotypic data. The clonal ramet count for phenotypic analysis was: BB=12, CC=11, E=25, G=11, M=31, and U=14. Phenotypic variables are organized throughout the result section by commonality of characteristics (budbreak, bark, and leaf characters) and are ordered accordingly: Julian days till start of budbreak (JDSBB), Julian days till end of budbreak (JDEBB), Julian days for total budbreak (JDTBB), diameter at breast height (DBH), bark thickness (BT), bark brightness (BR), leaf senescence (LS), average number of leaf serrations (ALS), and average leaf length ratio (LR).

3.2.1 Phenotypic variable selection-

When the phenotypic variables were tested for significance using both the Welch and Browns-Forsythe ANOVA, all variables were significant ($p < 0.05$) for at least one clonal group. Although Tamhane's T2 pairwise ANOVA post-hoc confirmed this, only one clone, clone U, displayed a small but significant differentiation in DBH (Figure 3.3). DBH was removed from further analysis. All other phenotypic variables were significantly different in a least two clonal colonies with JDSBB and JDEBB demonstrating the greatest interclonal variation.

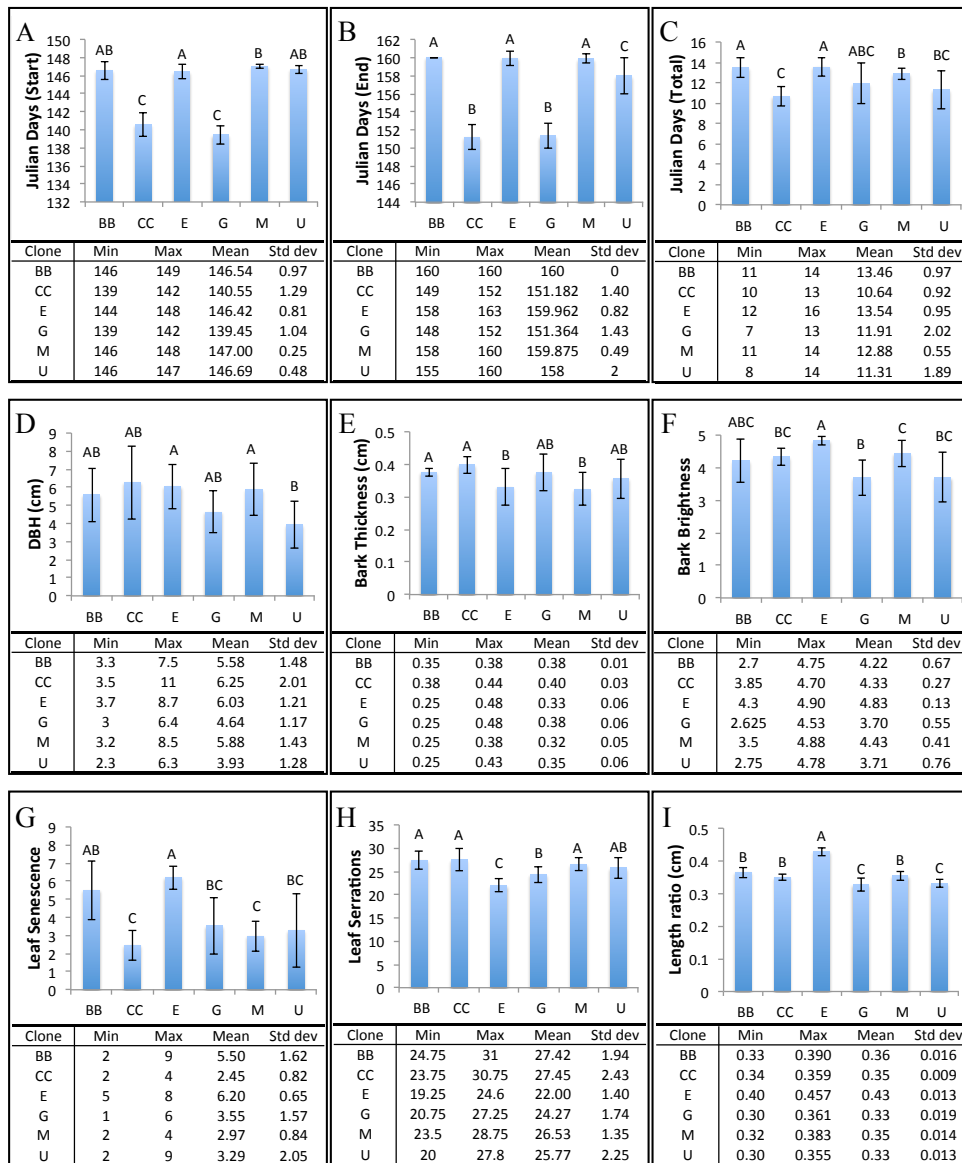


Figure 3.3. Phenotypic comparison between selected *P. tremuloides* clonal colonies (BB, CC, E, G, M, and U). Phenotypic variables include Julian days to the start of vegetative budbreak (A), Julian days to the end of budbreak (B), the total number of Julian days between the start and end of budbreak (C), DBH (D), bark thickness (E), bark brightness (F), leaf senescence (G), number of leaf serrations (H), and average leaf length ratio (I). Different letters above the std dev bars are representative of clonal means that are significantly different (Tamhane's T2, all p -values < 0.05).

The phenotypic variables were tested for multicollinearity and three, JDTBB, JDSBB, and JDEBB exhibited excessive collinearity with R^2 values of 1.0 (Figure 3.4). A Pearson's correlation confirmed a strong relationship among these variables with a correlation value as high as 0.916 (Table 3.4). Because JDEBB had the strongest correlation with both JDTBB and JDSBB (0.660 and 0.916, respectively), JDEBB was kept and the others removed from analysis. The remaining variables were checked again for multicollinearity and collinearity was successfully reduced with R^2 , tolerance, and variable inflation factor (VIF) values falling within acceptable ranges of $R^2 < 0.8$, tolerance > 0.20 , and $VIF < 4$ (Figure 3.5) (O'Brien 2007, Famy and Aubry 2013).

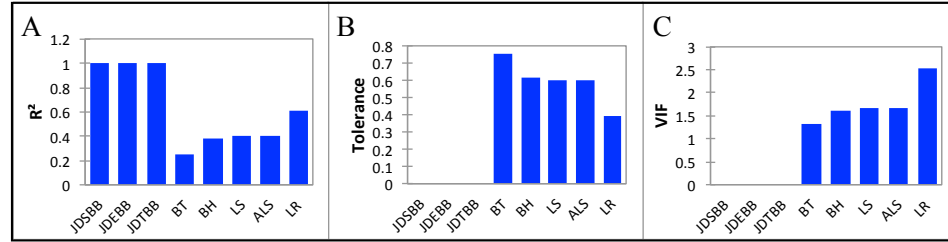


Figure 3.4. Multicollinearity statistics R^2 (A), tolerance (B), and variable inflation factor (VIF) (C), of phenotypic variables measured for *P. tremuloides*. Phenotypic variables: JDSBB=Julian days to start of budbreak; JDEBB=Julian days to end of budbreak; JDTBB=Julian days in total for budbreak; BT=bark thickness; BR=bark brightness; LS=leaf senescence; ALS=average number of leaf serrations; LR=average leaf length ratio.

Table 3.4. Pearson's correlation matrix to detect excessive collinearity between phenotypic variables measured for *P. tremuloides*. JDSBB=Julian days to start of budbreak; JDEBB=Julian days to end of budbreak; JDTBB=Julian days in total for budbreak; BT=bark thickness; BR=bark brightness; LS=leaf senescence; ALS=average number of leaf serrations; LR= average leaf length ratio.

Variables	JDSBB	JDEBB	JDTBB	BT	BH	LS	ALS	LR
JDSBB	1.000	0.916	0.303	-0.364	0.245	0.216	-0.017	0.338
JDEBB	0.916	1.000	0.660	-0.412	0.375	0.322	-0.134	0.435
JDTBB	0.303	0.660	1.000	-0.296	0.432	0.362	-0.287	0.400
BT	-0.364	-0.412	-0.296	1.000	-0.096	0.007	0.192	-0.201
BH	0.245	0.375	0.432	-0.096	1.000	0.316	-0.201	0.554
LS	0.216	0.322	0.362	0.007	0.316	1.000	-0.429	0.568
ALS	-0.017	-0.134	-0.287	0.192	-0.201	-0.429	1.000	-0.545
LR	0.338	0.435	0.400	-0.201	0.554	0.568	-0.545	1.000

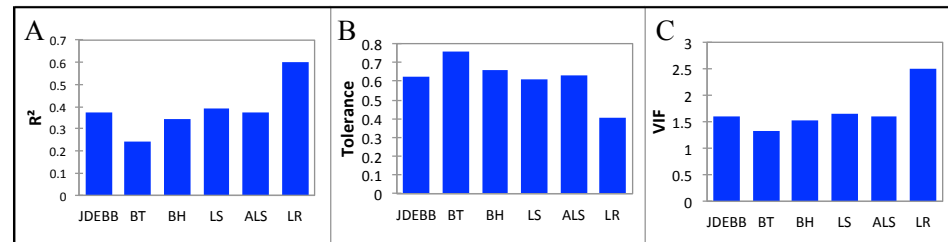


Figure 3.5. Multicollinearity statistics R^2 (A), tolerance (B), and variable inflation factor (VIF) (C), of remaining phenotypic variables measured for *P. tremuloides* after removing variables found to be collinear. Phenotypic variables: JDEBB=Julian days to end of budbreak; BT=bark thickness; BR=bark brightness; LS=leaf senescence; ALS=average number of leaf serrations; LR=average leaf length ratio.

3.2.2 Phenotypic inter- and intracolonial agreement-

The DA accounted for approximately 94% of the phenotypic variance in axis F1 and axis F2 (Figure 3.6). The variable with the strongest positive correlation was

JDEBB, with a positive correlation of ~95% on axis F1 and ~27% on axis F2. BT had the strongest negative correlation, with a negative correlation of ~37% on axis F1 and ~2% on axis F2. As illustrated in Figure 3.6, the ramets exhibited an intraclonal grouping pattern with overlap between clones BB, M, and U and between clones CC and G. In contrast, clone E did not overlap phenotypically with any other clone. Clone U showed the largest phenotypic variation.

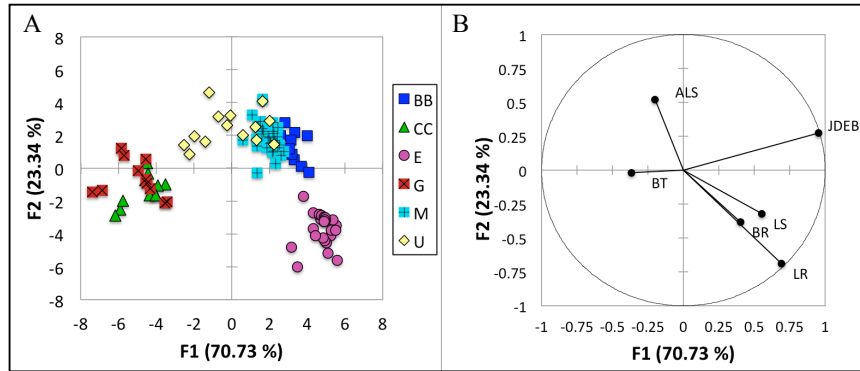


Figure 3.6. Discriminant analysis (A) for *P. tremuloides* genetically identified clonal colonies (BB, CC, E, G, M, and U) computed using the selected phenotypic variables. The correlation between the variables and factors F1 and F2 are shown in (B). JDEBB=Julian days to end of budbreak; BT=bark thickness; BR=bark brightness; LS=leaf senescence; ALS=average number of leaf serrations; LR=average leaf length ratio.

All MRPP pairwise clonal comparisons had pairwise chance-corrected intraclonal agreement (A) values greater than 0, which indicated phenotypic separation among Table 3.5. The greatest pairwise distance occurred between clones BB and CC (A=0.48) and the smallest between clones M and U (0.09). In correspondence with the DA, the MRPP averaged pairwise agreement values were highest in clone E (0.35) and pairwise agreement values were smallest among clones BB, M, and U, and clones CC and G. Clone U, with an averaged A=0.25, had the smallest averaged agreement value.

Table 3.5. Phenotypic pairwise comparison between selected *P. tremuloides* clonal colonies (BB, CC, E, G, M, and U) using a multi-response permutation procedure (MRPP). Given are the values for the t-statistic, intracolonal agreement value, and p-value for each pairwise comparison and clonal average. The higher the agreement value, the greater the phenotypic separation between clones.

Clone 1	Clone 2	T-statistic	Agreement	P-value
BB	CC	-13.810	0.476	<0.001
BB	E	-17.689	0.260	<0.001
BB	G	-14.206	0.471	<0.001
BB	M	-10.618	0.145	<0.001
BB	U	-7.484	0.161	<0.001
CC	E	-21.484	0.360	<0.001
CC	G	-4.590	0.136	0.002
CC	M	-22.935	0.326	<0.001
CC	U	-12.612	0.352	<0.001
E	G	-21.385	0.354	<0.001
E	M	-34.208	0.463	<0.001
E	U	-20.098	0.291	<0.001
G	M	-22.883	0.319	<0.001
G	U	-12.216	0.364	<0.001
M	U	-7.710	0.089	<0.001
Pairwise Average		T-statistic	Agreement	P-value
BB		-12.761	0.303	<0.001
CC		-15.086	0.330	<0.001
E		-22.973	0.346	<0.001
G		-15.056	0.329	<0.001
M		-19.671	0.277	<0.001
U		-12.024	0.252	<0.001

In total, 94.23% of all stems were correctly assigned to their clonal colony using the selected phenotypic variables under the CHAID classification and regression statistical procedure (Table 3.6).

Clone G at 81.82% and clone U at 85.71% showed the lowest correct stem assignment, whereas BB, CC, and E had a stem assignment accuracy of 100%. Clone M was in between with an accuracy of 93.55%. Stem reassignment occurred between clones BB and M, M and U, and CC and G with no reassignment to clone E. The MRPP individual distance values for each clone showed that the widest phenotypic intracolonal distance occurred in clone U (0.427) and the smallest phenotypic distances in clone E (0.116) and M (0.148). Clones BB, CC, and G had MRPP individual distance values of 0.240, 0.289, and 0.315, respectively. Overall, the MRPP values (observed delta=0.221, expected delta=0.5; A-value=0.558) demonstrated that ramets had a greater phenotypic association to their clonal colony than what could be expected by chance (Table 3.6). Both the CHAID and MRPP stem distance results are in agreement with the results from the DA and pairwise MRPP. A complete stem assignment decision tree can be seen in Suppl. Figure 4.

Table 3.6. Results of the classification and regression dendrogram illustrating inter- and intraclonal agreement values (A-values) among selected *P. tremuloides* clonal colonies (BB, CC, E, G, M, or U). Each stem was analyzed and assigned to either its original clone or another clone based on phenotype. The results show the percentage of correct stem assignment in total and for each clone. The intraclonal distances for all stems in each clonal colony are also given and were calculated using a multi-response permutation procedure (MRPP). The delta shown is the observed delta with the expected delta=0.5.

From To	BB	CC	E	G	M	U	Total	% correct	MRPP-Distance
BB	12	0	0	0	0	0	12	100.00%	0.240
CC	0	11	0	0	0	0	11	100.00%	0.289
E	0	0	25	0	0	0	25	100.00%	0.116
G	0	2	0	9	0	0	11	81.82%	0.315
M	1	0	0	0	29	1	31	93.55%	0.148
U	0	0	0	0	2	12	14	85.71%	0.427
Total	13	13	25	9	31	13	104	94.23	Delta-0.221 A-value-0.558

3.2.3 Accuracy of phenotypic variables for clonal assignment-

The PCA accounted for approximately 64% of the total variance, with 44.72% of the variance explained on axis F1 and 19.24% explained on axis F2 (Figure 3.7). Phenotypic variable LR showed the highest positive correlation with a positive correlation value of ~88% on axis F1 and ~15% on axis F2. Phenotypic variable ALS was the most negatively correlated variable with negative correlation values of ~64% on axis F1 and ~29% on axis F2. Variables LS, BR, and LR were all positively correlated on both axes. Clone E showed the most isolation, graphically, from the other clones. Clone U demonstrated significant phenotypic overlapping of clones BB, CC, G, and M. Clones BB and M overlapped somewhat and clones CC and G showed pronounced overlap with one another. A single ramet from clone G was found next a ramet from clone M, but clones G and M did not phenotypically cross in any other instance (Figure 3.7).

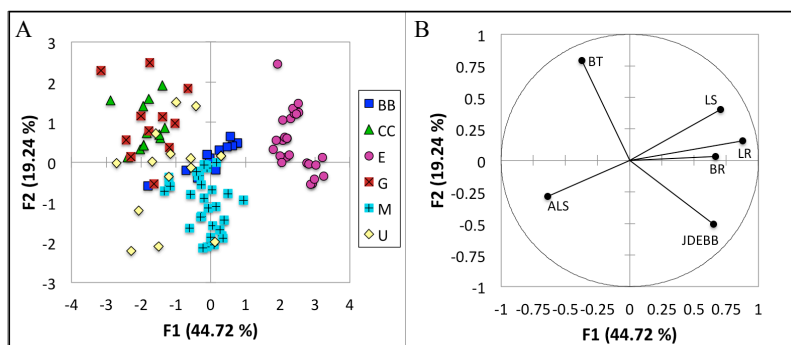


Figure 3.7. Principal component analysis (A) for *P. tremuloides* genetically identified clonal colonies (BB, CC, E, G, M, and U) computed using the selected phenotypic variables. The correlation between the variables and factors F1 and F2 are shown in (B). JDEBB=Julian days to end of budbreak; BT=bark thickness; BR=bark brightness; LS=leaf senescence; ALS=average number of leaf serrations; LR=average leaf length ratio.

The AHC statistical procedure separated stems into six different phenotypic groups based on the Bray-Curtis flexible linkage (-0.25) method (Figure 3.8, Table 3.7). All phenotypic groups, labeled 1-6, contained more than one clonal colony, but each was primarily composed of or associated with one clone: group 1 with clone BB, group 2 with clone M, group 3 with clone CC, group 4 with clone G, group 5 with clone E, and group

6 with clone U. Phenotypic group 5 was genetically the “purest” group and contained only one ramet from clonal colony U. Phenotypic group 2 was the most genetically diverse and contained ramets from clonal colonies BB, E, M and U. Overlap among clonal colonies within the phenotypic groups occurred between clones BB, E, and M and between clones CC and GG. Ramets from clonal colony U were found in almost every phenotypic group except group 3. Phenotypic group 6, associated with clone U, contained only six of the 14 ramets from clone U and two ramets from clone M (Figure 3.8, Table 3.7).

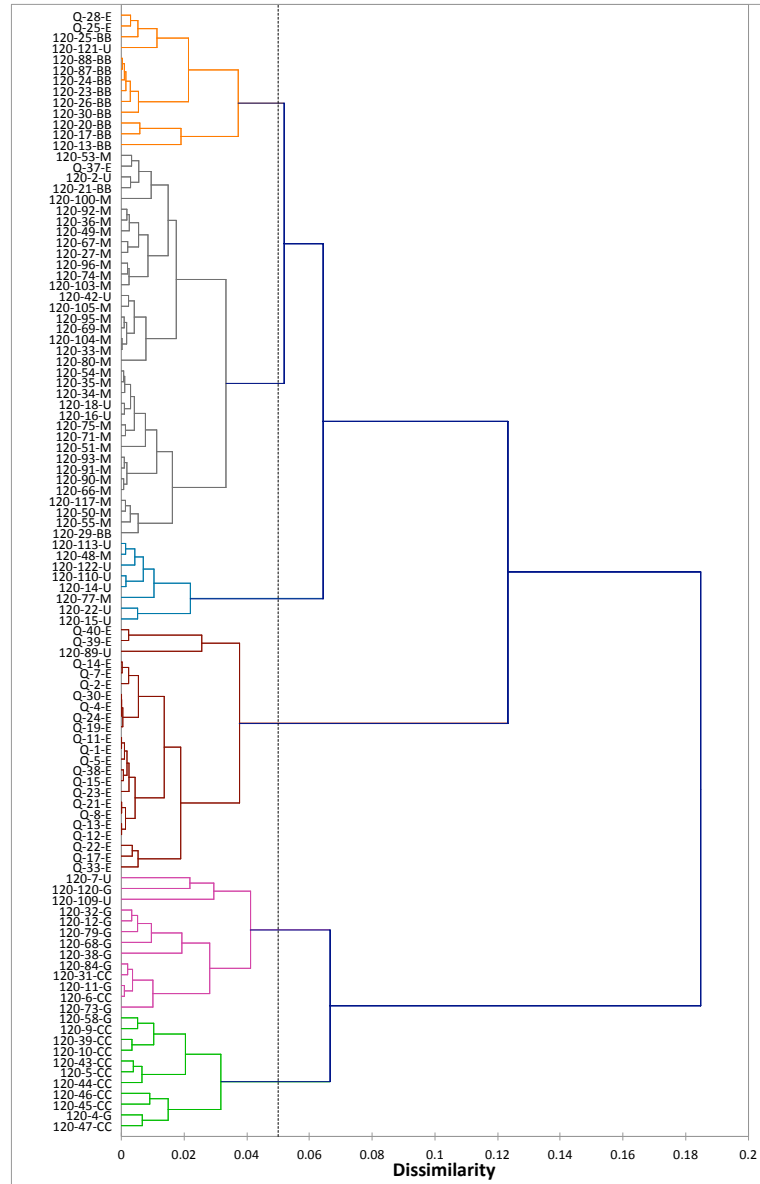


Figure 3.8. Agglomerative hierarchical clustering (AHC) dendrogram of stem separation from the selected *P. tremuloides* clonal groups (BB, CC, E, G, M, and U) using phenotypic characters. The first letter or number (120, 3, or Q) represents the original phenotypic grouping, the second number is the sample number for each stem in the original grouping, and the final lettering represents the genetically identified clonal colony. This procedure makes no assumption about prior group assignment.

Table 3.7. Stem and leaf plot based on the results of the agglomerative hierarchical clustering (AHC). Phenotypic group assignment is illustrated for *P. tremuloides* ramets from selected clonal colonies (BB, CC, E, G, M, and U). The first letter or number (120, 3, or Q) represents the original grouping, the second number is the sample number of each stem in each original grouping, and the final letter represents the clonal colony.

Phenotypic group	1	2	3	4	5	6
Stems	13	36	11	13	23	8
Within-class variance	0.004	0.002	0.005	0.004	0.002	0.002
Min distance	0.030	0.017	0.032	0.027	0.012	0.020
Max distance	0.128	0.088	0.094	0.103	0.098	0.075
Mean distance	0.056	0.039	0.062	0.055	0.036	0.039
	120-13-BB	120-21-BB	120-5-CC	120-6-CC	Q-1-E	120-48-M
	120-17-BB	120-29-BB	120-9-CC	120-31-CC	Q-2-E	120-77-M
	120-20-BB	Q-37-E	120-10-CC	120-11-G	Q-4-E	120-14-U
	120-23-BB	120-27-M	120-39-CC	120-12-G	Q-5-E	120-15-U
	120-24-BB	120-33-M	120-43-CC	120-32-G	Q-7-E	120-22-U
	120-25-BB	120-34-M	120-44-CC	120-38-G	Q-8-E	120-110-U
	120-26-BB	120-35-M	120-45-CC	120-68-G	Q-11-E	120-113-U
	120-30-BB	120-36-M	120-46-CC	120-73-G	Q-12-E	120-122-U
	120-87-BB	120-49-M	120-47-CC	120-79-G	Q-13-E	
	120-88-BB	120-50-M	120-4-G	120-84-G	Q-14-E	
	Q-25-E	120-51-M	120-58-G	120-120-G	Q-15-E	
	Q-28-E	120-53-M		120-7-U	Q-17-E	
	120-121-U	120-54-M		120-109-U	Q-19-E	
		120-55-M			Q-21-E	
		120-66-M			Q-22-E	
		120-67-M			Q-23-E	
		120-69-M			Q-24-E	
		120-71-M			Q-30-E	
		120-74-M			Q-33-E	
		120-75-M			Q-38-E	
		120-80-M			Q-39-E	
		120-90-M			Q-40-E	
		120-91-M			120-89-U	
		120-92-M				
		120-93-M				
		120-95-M				
		120-96-M				
		120-100-M				
		120-103-M				
		120-104-M				
		120-105-M				
		120-117-M				
		120-2-U				
		120-16-U				
		120-18-U				
		120-42-U				

The CHAID classification and regression statistical procedure was able to correctly classify 99.04% of the stems under the AHC phenotypic groups (Table 3.8), a 4.81% higher correct assignment than that of the genetic groups (Table 3.6). Only one stem from phenotypic group 2 was reassigned to group 1. When comparing the MRPP individual distance values between clonal colonies (Table 3.6) and the associated phenotypic groups (Table 3.8), individual intraclonal distance values were higher in the clones BB (0.240), M (0.148), CC (0.289), and U (0.427) than the associated phenotypic groups 1 (0.232), 2 (0.145), 3 (0.283), and 6 (0.147). Clonal colonies G (0.315) and E (0.116) had smaller individual distance values than the associated phenotypic groups 4 (0.325) and 5 (0.129). The MRPP observed delta for the phenotypic groups was 0.189 (expected delta=0.5) (Table 3.8), which is lower than the observed clonal delta (0.221) (Table 3.6), and the agreement value was 0.621, which is higher than the clonal agreement value (0.558).

Table 3.8. Results of the classification and regression tree testing phenotypic inter- and intraclonal agreement (A-value) after an agglomerative hierarchal reclassification of *P. tremuloides* stems based on phenotypic characters. Each stem was analyzed and assigned to either its original phenotypic group or another phenotypic group. The results show the percentage of correct stem assignment in total and for each group. The intra-group distances for all stems in each phenotypic group are also given and were calculated using a multi-response permutation procedure (MRPP). The delta shown is the observed delta with the expected delta=0.5.

To \ From							Total	% correct	MRPP-Distance
	1	2	3	4	5	6			
1	13	0	0	0	0	0	13	100.00%	0.232
2	1	35	0	0	0	0	36	97.22%	0.145
3	0	0	11	0	0	0	11	100.00%	0.283
4	0	0	0	13	0	0	13	100.00%	0.325
5	0	0	0	0	23	0	23	100.00%	0.129
6	0	0	0	0	0	8	8	100.00%	0.147
Total	14	35	11	13	23	8	104	99.04%	Delta-0.189 A-value-0.621

4 Discussion

4.1 Genetic analysis

Overall, the seven nuclear microsatellite markers were adequate for proper genet identification. The observed heterozygosity was slightly higher than the expected heterozygosity, which demonstrated a high allelic diversity with the stand. This is not unexpected of a clonal stand, as heterozygosity tends to remain high when genotypes are able to persist in a stand through vegetative reproduction (Balloux et al. 2003). Additionally, the number of genetically distinct individuals was higher than what would be expected of a purely clonal stand with little sexual reproduction (Perala 1990). It is common in the Great Lakes region, particularly in Upper Michigan, to find genetically diverse *P. tremuloides* stands because the requirements for germination (Perala 1990), such as moist and exposed mineral soil, are usually met. The high genetic diversity and location of this stand indicates that sexual reproduction is prevalent. Finally, the six clonal colonies identified were determined to have arisen from vegetative rather than sexual means, showing that all ramets with the same genotype were from the same genet.

4.2 Phenotypic variable significance

The present study was designed to analyze inter- and intraclonal phenotypic variation among genetically identified *P. tremuloides* clonal colonies. In doing so, I was able to evaluate the accuracy of using phenotypic variables or characters for clonal delineation. Out of the nine variables, the three analyses of budbreak, JDSBB, JDEBB, and JDTBB, demonstrated the greatest interclonal discrimination. Furthermore, it was JDEBB that was the most reliable of the budbreak characters for clonal discernment even though the budbreak variables were collinear. The second most reliable character was LR. Leaf length ratio had a similar variance among the selected clones, making LR a reliable character for clonal delineation.

In contrast, DBH was the least reliable character for determining clonality because only one clone (clone U) was significantly different from at least one other clone. The other phenotypic characters, BT, BR, LS, and ALS, exhibited varying degrees of clonal discrimination and intraclonal variation. No single phenotypic variable was able to accurately distinguish more than one clonal colony. However, interclonal discrimination was more definitive when applying multivariate analysis.

4.3 Phenotypic inter- and intraclonal agreement

There was a general trend for the combined phenotypic characters, JDEBB, BT, BR, LS, ALS, and LR, to vary with genotype. The CHAID classification and regression was able to assign ramets to the correct genet using phenotype with an accuracy rate of 94.23%. However, phenotype did overlap between clones BB and M, clones CC and G, and clones M and U with some overlap between BB and U. Contrastingly, clone E was phenotypically separated from the other five clones in both the DA and MRPP pairwise analysis. This may have been caused by microsite variations, as clone E was spatially

separated from the other five clonal colonies. Even so, spatial separation and microsite variation cannot entirely explain the phenotypic clonal separation. If inter- and intraclonal phenotypic distances were based entirely on microsite variation (i.e. soil conditions and light availability), one would expect to see greater phenotypic variation in clonal colonies that covered more area or contained more ramets. The data presented here suggest the opposite as the largest clone both in area and in ramet number, clone M, had the second lowest intraclonal phenotypic distance (0.148). Clone E had the lowest intraclonal phenotypic distance (0.116).

Furthermore, if microsite variation had a greater influence than genotype, clones that spatially overlapped would be expected to share more phenotypic similarities (Gilmore et al. 1968, Keith et al. 1997). While clones BB and M and clones CC and G spatially overlap to some extent, clones G and M shared more of the same spatial area than clones BB and M or clones CC and G. Despite this, clones G and M do not phenotypically overlap within the DA. Clones G and M additionally exhibited a greater phenotypic pairwise MRPP distance than that of clones BB and M or clones CC and G. This pattern, where spatial distance is less correlated to phenotypic distance, is also present between CC and M when comparing clones BB and M, and between BB, and U when comparing clones M and U.

Overall, clonal colonies were phenotypically separate and intraclonal distances were less than what would be expected by chance. The inter- and intraclonal variation analyses, however, used predefined groups to determine clonal separation. In a stand where the number and size of clonal colonies is uncertain, clonal delineation is not as straight forward. Genets will often have ramets that are phenotypic outliers or two genets may be very phenotypically similar. This may cause an over or under estimation of the true number and size of clonal colonies. It was therefore important to test phenotypic clonal assignment using statistical parameters that phenotypically separated stems without prior group bias.

4.4 Phenotypic clonal assignment

Both the PCA and AHC separated stems based on phenotypic characters. The PCA exhibited greater interclonal similarities than the DA, but that was expected as PCAs maximize distances between data points and DAs use prior established groups for comparison. Intraclonal agreement was present in the PCA, but if the genetic clonal labels are not considered, only two or three phenotypic groups can be graphically separated. It was difficult to determine how stems should be grouped based only on the PCA because clone U overlapped four other clones. Furthermore, the PCA did not represent all of the possible variation so phenotypic grouping could only be generalized. The AHC analysis was an easier and more robust measure for assigning phenotypic groups than the PCA.

The AHC analysis accurately determined that there were six distinct groups. Additionally, the majority of stems in each phenotypic group were comprised of a single genet. Phenotypic separation, however, was not completely representative of genotypic separation. No single genotype was fully represented by one phenotypic group and all phenotypic groups contained at least one other genotype. The genotypic overlap among

phenotypic groups was similar to that in PCA, with the exception of clone E. Clone E was phenotypically separated from the other clonal colonies in the PCA, DA, and CHAID regression analysis.

It is sometimes the case that the AHC will wrongly assign an individual to a particular group. For this reason another MRPP and CHAID analysis was conducted to test the robustness of the AHC phenotypic groups. Overall, the MRPP phenotypic intragroup distances were smaller than intraclonal distances and the CHAID analysis assigned 99.04% of the stems to the correct phenotypic group. Both the MRPP and CHAID analysis demonstrated that the AHC grouping was a fairly accurate representation of phenotypic separation.

These results have demonstrated that phenotypic characters vary among clonal colonies under natural conditions. Evidence has been provided that stand clonality may be determined using expressed traits, but that clonal delineation may not always be 100% accurate. From these data and the data presented in other studies (Suvanto and Latva-Karjanmaa 2005, Jelínková et al. 2013), it is clear that phenotypic delineation is not as robust as genetic identification. However, phenotypic delineation may be used to complement genetic analysis. Phenotypic delineation would be particularly useful when genetically identifying an entire stand is impractical, or when determining which and how many stems should be genetically analyzed. Nevertheless, several considerations must be made when clonal delineation involves the combination of phenotypic and genetic analysis.

4.5 Considerations for the application of phenotypic delineation techniques

Based on this data and findings presented by other natural population studies for *P. tremuloides* (Barnes 1966, 1969, Namroud et al. 2005, DeWoody et al. 2008, Mock et al. 2008, Jelínková et al. 2009, 2013), several considerations are discussed regarding the application of phenotypic techniques for clonal delineation. First, the accuracy of phenotypic delineation most likely varies with stand density, stem and stand age, and the frequency of successful germination. Younger stands generally exhibit a higher stem density and greater genetic diversity than older stands (Barnes 1966, Tuskan et al. 1996, Pellis et al. 2004, Jelínková et al. 2013). This is because competition over resources and an increase in the number of ramets/genet remove less viable genets, decreasing genet diversity over time (Weiner 2004). The number of ramets/genet in a stand is further influenced by the rate of germination and seedling establishment (Barnes 1969, Romme et al. 2005, DeWoody et al. 2009). Populations are more genetically diverse in areas where soil is sufficiently moist during crucial seedling developmental periods and where bare mineral soil is often exposed due to frequent stand disturbances (Perala 1990, Namroud et al. 2005). Areas with greater seedling establishment are located in Canada and the northern and eastern portions of the United States (Perala 1990, Namroud et al. 2005, Jelínková et al. 2009). In the western Rocky Mountain *P. tremuloides* ranges within the US, proper germination conditions are less frequent so clonal sizes are generally larger (Perala 1990, DeWoody et al. 2008, Mock et al. 2008). Such evidence suggests that genet size is inversely proportional to the rate of germination and seedling establishment. For these reasons, more stems should be genetically identified under the

following conditions: (1) when clonal stands have not reached sexual maturity, (2) in extremely dense stands where growth is limited by available space, or (3) in areas where the rates of successful germination and seedling establishment are higher. While I am unable to give an estimated number of stems that should be genetically identified, a general rule of thumb would be to genetically identify more stems in the presence of greater phenotypic variability. This is because phenotypic variability is largely associated with high genetic diversity (Barnes 1969, Jelínková et al. 2013). Genetic identification, though, is always recommended after any phenotypic delineation. Less genetic analysis would be required, however, when phenotypic delineation is applied in older and less dense stands where seedling recruitment is rare.

Second, a single phenotypic character is not adequate for an accurate assessment of stand clonality (Blake 1963). Phenotypic traits, as demonstrated by these findings, may overlap among clonal colonies when genets are unrelated and even when a trait is under strong genetic control. It is also expected that greater relatedness will lead to greater interclonal phenotypic agreement. However, I was unable to examine the potential relationship between relatedness and phenotypic expression because the selected clonal colonies were unrelated. While budbreak is often demonstrated as being the most robust phenotypic character for clonal delineation, differences in budbreak may not be apparent between certain genets (Jelínková et al. 2013). As such, it would be recommended to phenotypically group stems based on budbreak then separate those groups further using other phenotypic characters. Days until full budbreak was a particularly useful trait in this study, followed by leaf length ratio then bark brightness, leaf serrations, leaf senescence, and bark thickness. Blake (1963) and Barnes (1969) found that gender was another useful character for clonal delineation. I was unable to use gender, however, because flowering did not occur in this stand the year I recorded phenotypic characters.

Third, it is expected that single ramet genets will often be indistinguishable from multi-ramet genets. (Cheliak and Pitel (1984), Tuskan et al. (1996), Wyman et al. (2003)), and Jelínková et al. (2013) have all had difficulty phenotypically identifying single or small ramet genets. This is most likely because there are no genetic replicates to form distinct phenotypic clusters, making single ramet genets hard to discern from other clonal colonies. Single ramet genets, however, are not as frequent in older stands or in stands where successful germination is less common (DeWoody et al. 2009). This is yet another indication that more genetic analysis is needed in sexually immature stands where seedling establishment occurs periodically.

4.6 Future research

The results presented here validate, to some extent, the use of phenotypic characters to determine stand clonality. Nonetheless, only six of the 32 genetically identified genets were used for these analyses. Future statistical analyses will assess the other 26 genets to test how accurately phenotype characters can identify and delineate smaller ramet genets. Because several of the 26 genets were genetically related (results not shown), additional statistical tests will also investigate the relationship between relatedness and phenotypic expression. These genets, though, were excluded from these

results due to a lack of sufficient replicates (i.e. ramets) to test both individual phenotypic characters for significance and evaluate inter- and intracolonial phenotypic variation. Additional genetic conducted on more stems within this plot could potentially identify more ramets to these genets and add to the robustness of further phenotypic analyses.

Further research is necessary to confirm these findings in other *P. tremuloides* populations under different environmental and stand conditions. The clonal stand selected for observation was small and that had either recently attained, or was close to attaining sexual maturity. This study site was therefore not representative of an older clonal forest, but neither was it representative of a younger, sexually immature clonal stand. The stand was also in an area where soil conditions are conducive for seed germination and seedling establishment in *P. tremuloides*. Consequently, this plot is likely to contain more genets than older *P. tremuloides* stands found in the western US, such as those in the western Rocky Mountains or in other environmentally similar locations (Perala 1990). Additional research would also provide new insight regarding the number of stems that should be genetically identified for accurate clonal delineation under different environmental conditions.

Finally, current phenotypic delineation techniques have varied drastically among research studies (Kemperman and Barnes 1976, Cheliak and Pitel 1984, Tuskan et al. 1996, Wyman et al. 2003, Smith et al. 2011, Jelínková et al. 2013). Establishing phenotypic delineation procedures, which included standardized phenological and morphological charts for different habitats, would decrease inconsistencies present in current applications. Clonal delineation accuracy may subsequently increase simply due to the availability of standardized charts and guidelines for clonal assessment. Overall, the establishment of such procedures could assist both foresters and researchers who want to determine stand clonality, but do not have the resources for exhaustive genetic identification.

5 Conclusion

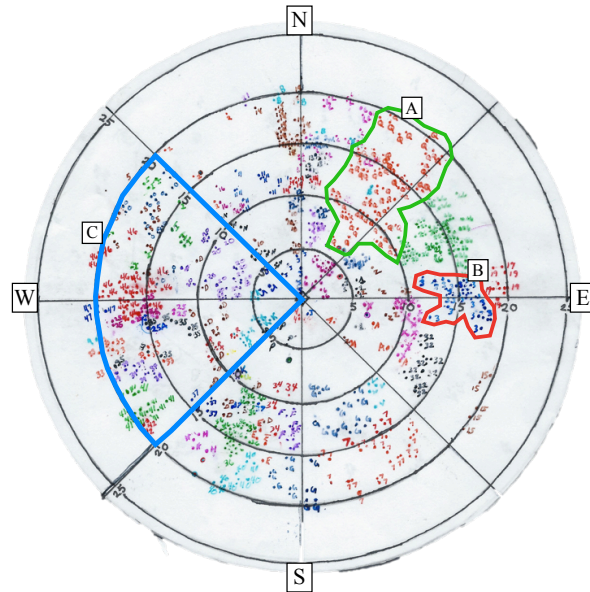
Determining population or stand clonality using phenotypic characters has had variable rates of success. Research conducted on *P. tremuloides* before the widespread establishment of molecular markers relied solely on phenotypic characters to identify and delineate clonal colonies (Blake 1963, Barnes 1966, 1969, Steneker 1973, Kemperman and Barnes 1976). Kemperman and Barnes (1976) accurately delineated the clone "Pando" using only phenotypic characters, which was genetically confirmed by DeWoody et al. (2008). Recent genetic research, though, has challenged the validity of these phenotypic delineation techniques (Cheliak and Pitel 1984, Tuskan et al. 1996, Wyman et al. 2003). However, the techniques used for phenotypic delineation are highly variable among studies using molecular identification. These genetic studies also provided very little information regarding the utilized phenotypic delineation procedures (Cheliak and Pitel 1984, Tuskan et al. 1996, Wyman et al. 2003).

To address these conflicting findings, I analyzed the phenotypic inter- and intracolonial variation of *P. tremuloides* in LTSP plot 28 located in the Ottawa NF of

Upper Michigan. Six of the 32 genetically identified genets were selected for phenotypic analyses, which included budbreak timing, DBH, bark thickness, bark color or brightness, leaf senescence, leaf serrations, and leaf length ratio. All phenotypic characters, except for DBH, were determined to be useful for the analysis of inter- and intraclonal variation and phenotypic delineation. Generally, phenotypic expression correlated with genotype. Interclonal phenotypic distances, though, were low between certain clones and intraclonal phenotypic distances were larger in some clones than others. When stems were assigned into phenotypic groups, six phenotypic groups were identified with each group containing a dominant genotype or clonal colony. However, all phenotypic groups contained stems from at least two clonal colonies and no clonal colony was entirely contained within one phenotypic group.

These findings indicate that clonal detection is possible using phenotypic characters and that phenotype varies with genotype. The precise delineation of clonal colonies, though, is more difficult since intraclonal variation may be large while interclonal distances may be small (Jelínková et al. 2013). I would always recommend that genetic identification occur to some extent after any phenotypic delineation. However, the exhaustive genetic identification of a stand or population is often impractical or impossible. Additionally, extensive genetic identification may not be needed in older clonal stands with lower genetic diversity. I therefore suggest that the amount of genetic identification applied to compliment phenotypic delineation should vary based on stem age and stand location. Finally, it is recommended that a standardized set of phenotypic delineation techniques be established for different stand and environmental conditions. This would remove many application inconsistencies and may ultimately increase the accuracy of phenotypic delineation.

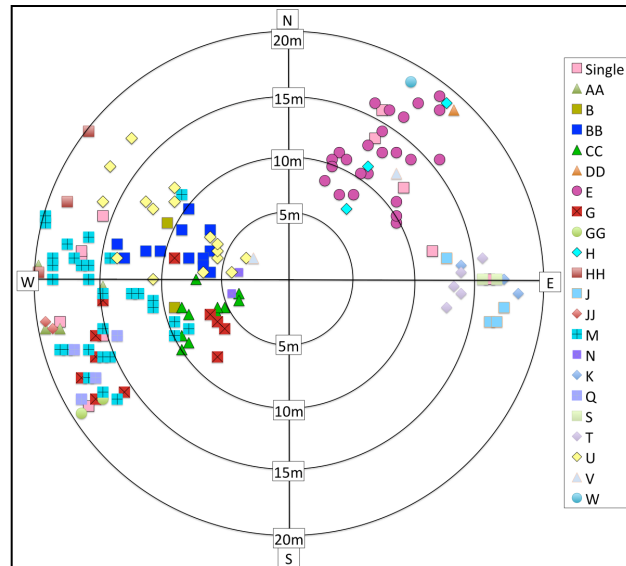
Supplementary figures



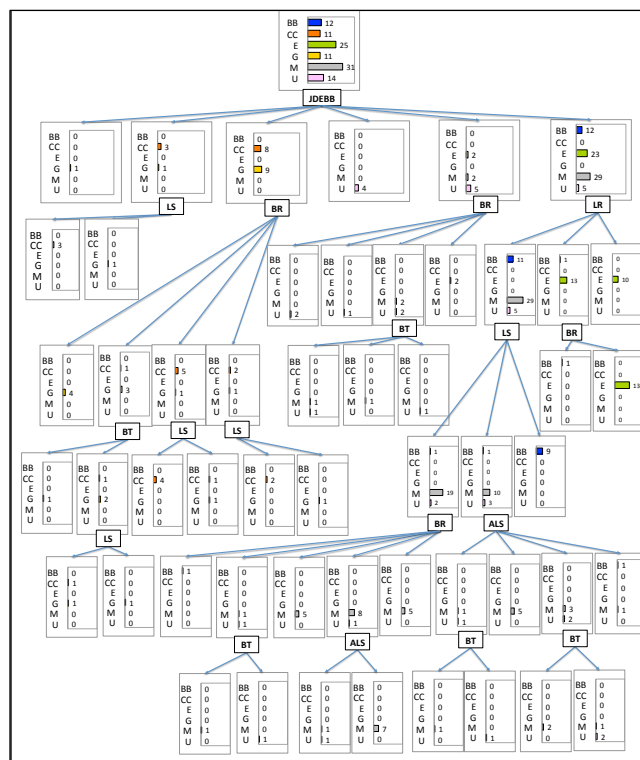
Suppl. Figure 1. Hand drawn *P. tremuloides* stem map of the Ottawa NF LTSP plot 28. The number of stems within 20 meters of plot center totals 585. Highlighted are the 181 stems selected for genetic and phenotypic analysis, which include putative clone "Q" (A), putative clone "3" (B), and the 120 exhaustive sampling (C).



Suppl. Figure 2. The 20-color chart used for initial bark color analysis. Colors were later converted into a "grey scale" format because most of the colors in sampled *P. tremuloides* were found within colors 1-5.



Suppl. Figure 3. LTSP Ottawa NF plot 28 map of all genetically identified *P. tremuloides* ramets and genets. Genets consisting of a single ramet are denoted with pink squares. Distances are given in 5-meter increments from plot center. N, E, S, and W represent cardinal directions.



Suppl. Figure 4. Classification and regression tree depicting *P. tremuloides* genetically identified clonal colonies (BB, CC, E, G, M, and U) separated phenotypically. Each branch separates stems based on one phenotypic variable with numbers representing the total number of ramets per genet per branch. BT=bark thickness; BR=bark brightness; LS=leaf senescence; ALS=average number of leaf serrations; LR=average leaf length ratio.

References

- Ally, D., K. Ritland, and S. P. Otto. 2008. Can clone size serve as a proxy for clone age? An exploration using microsatellite divergence in *Populus tremuloides*. *Mol Ecol* **17**:4897-4911.
- Aravanopoulos, F. 2010. Clonal Identification Based on Quantitative, Codominant, and Dominant Marker Data: A Comparative Analysis of Selected Willow (*Salix L.*) Clones. *Int J Forest Resour* **2010**.
- Bagley, M. J., S. L. Anderson, and B. May. 2001. Choice of methodology for assessing genetic impacts of environmental stressors: polymorphism and reproducibility of RAPD and AFLP fingerprints. *Ecotoxicology* **10**:239-244.
- Balloux, F., L. Lehmann, and T. de Meeûs. 2003. The population genetics of clonal and partially clonal diploids. *Genetics* **164**:1635-1644.
- Barnes, B. V. 1959. Natural variation and clonal development of *Populus tremuloides* and *Populus grandidentata* in northern lower michigan. DAI-PHD **20**.
- Barnes, B. V. 1966. The clonal growth habit of American aspens. *Ecology* **47**:439-447.
- Barnes, B. V. 1969. Natural variation and delineation of clones of *Populus tremuloides* and *P. grandidentata* in northern Lower Michigan. *Silvae Genetica* **18**:130-142.
- Barnes, B. V. 1975. Phenotypic variation of trembling aspen in western North America. *Forest Sci* **21**:319-328.
- Barnes, B. V. 1987. *Populus tremuloides* and *P. grandidentata* clonal isozyme identification in northern Lower Michigan. Unpublished raw data.
- Bergsma, W. 2012. A bias-correction for Cramér's V and Tschuprow's T. *J Korean Statist Soc* **42**:323-328.
- Blake, G. M. 1963. Clone identificaiton and delineation in the aspens. (PhD) Thesis. University of Minnesota.
- Borcard, D., and P. Legendre. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbor matrices. *Ecol Model* **153**:51-68.
- Boyko, A., J. Filkowski, D. Hudson, and I. Kovalchuk. 2006. Homologous recombination in plants is organ specific. *Mutat Res* **595**:145-155.
- Bradshaw, H., and R. F. Stettler. 1995. Molecular genetics of growth and development in *Populus*. IV. Mapping QTLs with large effects on growth, form, and phenology traits in a forest tree. *Genetics* **139**:963-973.
- Brown, A., and B. S. Weir. 1983. Measuring genetic variability in plant populations. Isozymes in plant plant genetics and breeding, part A. Elsevier Science Publishers BV, Amsterdam.
- Brown, A. H. D., and G. F. Moran. 1981. Isozymes and the genetic resources of forest trees. In Proceedings of the Symposium on Isozymes of North American Forest Trees and Forest Insects. PSW For Range Exp Stn Gen Tech Rep PSW-48:1-10.
- Brown, M. B., and A. B. Forsythe. 1974. Robust tests for the equality of variances. *Journal of the American Statistical Association* **69**:364-367.
- Buth, D. G., and R. W. Murphy. 1999. The use of isozyme characters in systematic studies. *Biochem Syst Ecol* **27**:117-129.

- Cheliak, W., and J. Pitel. 1984. Electrophoretic identification of clones in trembling aspen. *Can J Forest Res* **14**:740-743.
- Clinch, J. J., and H. J. Keselman. 1982. Parametric alternatives to the analysis of variance. *J Educ Stat* **7**:207-214.
- Conner, J. K., and D. L. Hartl. 2004. A primer of ecological genetics. Sinauer Associates Incorporated, Sunderland, MA.
- De Muth, J. E. 2006. Basic statistics and pharmaceutical statistical applications. CRC Press, New York, NY.
- Delacôte, F., and B. S. Lopez. 2008. Importance of the cell cycle phase for the choice of the appropriate DSB repair pathway, for genome stability maintenance. *Cell Cycle* **7**:33-38.
- DeWoody, J., T. H. Rickman, B. E. Jones, and V. D. Hipkins. 2009. Allozyme and microsatellite data reveal small clone size and high genetic diversity in aspen in the southern Cascade Mountains. *Forest Ecol Manag* **258**:687-696.
- DeWoody, J., C. A. Rowe, V. D. Hipkins, and K. E. Mock. 2008. "Pando" lives: molecular genetic evidence of a giant aspen clone in central Utah. *West N Am Naturalist* **68**:493-497.
- Douhovnikoff, V., and R. Dodd. 2003. Intraclonal variation and a similarity threshold for identification of clones: application to *Salix exigua* using AFLP molecular markers. *Theor Appl Genet* **106**:1307-1315.
- Dunnett, C. W. 1980. Pairwise multiple comparisons in the unequal variance case. *J Am Stat Assoc* **75**:796-800.
- Dunnett, C. W., and A. C. Tamhane. 1992. A step-up multiple test procedure. *J Am Stat Assoc* **87**:162-170.
- Ellegren, H. 2000. Microsatellite mutations in the germline. *Trends Genet* **16**:551-558.
- Famy, T., and P. Aubry. 2013. XLSTAT-Pro (version 2013.5.06). Society of Addinsoft **20**.
- Gilmore, A., W. Geyer, and W. Boggess. 1968. Microsite and height growth of yellow-poplar. *For Sci* **14**:420-426.
- Gom, L. A., and S. B. Rood. 1999. The discrimination of cottonwood clones in a mature grove along the Oldman River in southern Alberta. *Can J Bot* **77**:1084-1094.
- Grower, J. C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* **53**:325-338.
- Hadrys, H., M. Balick, and B. Schierwater. 1992. Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. *Mol Ecol* **1**:55-63.
- Hardy, O. J. 2003. Estimation of pairwise relatedness between individuals and characterization of isolation - by - distance processes using dominant genetic markers. *Mol Ecol* **12**:1577-1588.
- Hartl, D. L., and A. G. Clark. 1997. Principles of population genetics. Sinauer Associates, Sunderland, MA.
- Hill, T., and P. Lewicki. 2006. Statistics: methods and applications: a comprehensive reference for science, industry, and data mining. StatSoft, Inc., Tulsa, OK.
- Hilton, A., and R. A. Armstrong. 2006. Statnote 6: Post-hoc ANOVA tests. *Microbiologist* **2006**:34-36.

- IBM. 2013. IBM SPSS Statistics for Windows, Version 22.0. IBM Corp., Armonk, NY.
- Ingvarsson, P. K., M. V. Garcia, V. Luquez, D. Hall, and S. Jansson. 2008. Nucleotide polymorphism and phenotypic associations within and around the phytochrome B2 locus in European aspen (*Populus tremula*, Salicaceae). *Genetics* **178**:2217-2226.
- Iwata, H., and Y. Ukai. 2002. SHAPE: a computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. *J Hered* **93**:384-385.
- Jelínková, H., F. Tremblay, and A. DesRochers. 2009. Molecular and dendrochronological analysis of natural root grafting in *Populus tremuloides* (Salicaceae). *Am J Bot* **96**:1500-1505.
- Jelínková, H., F. Tremblay, and A. DesRochers. 2013. The use of digital morphometrics and spring phenology for clone recognition in trembling aspen (*Populus tremuloides* michx.) and its comparison to microsatellite markers. *Trees* **28**:389-398.
- Jelinski, D. E., and W. Cheliak. 1992. Genetic diversity and spatial subdivision of *Populus tremuloides* (Salicaceae) in a heterogeneous landscape. *Am J Bot* **79**:728-736.
- Jolliffe, I. 2002. Principal component analysis. 2nd edition. Springer, New York, NY.
- Jones, C. J., K. J. Edwards, S. Castaglione, M. O. Winfield, F. Sala, C. Van de Wiel, G. Bredemeijer, B. Vosman, M. Matthes, and A. Daly. 1997. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Mol breeding* **3**:381-390.
- Keith, H., K. Jacobsen, and R. Raison. 1997. Effects of soil phosphorus availability, temperature and moisture on soil respiration in *Eucalyptus pauciflora* forest. *Plant Soil* **190**:127-141.
- Kemperman, J. A., and B. V. Barnes. 1976. Clone size in American aspens. *Can J Bot* **54**:2603-2607.
- Kimura, M., and J. F. Crow. 1964. The number of alleles that can be maintained in a finite population. *Genetics* **49**:725-738.
- Klecka, W. R. 1980. Discriminant analysis. Sage, Beverly Hills, CA.
- Lachenbruch, P. A., and M. Goldstein. 1979. Discriminant analysis. *Biometrics* **35**:69-85.
- Lian, C., R. Oishi, N. Miyashita, and T. Hogetsu. 2004. High somatic instability of a microsatellite locus in a clonal tree, *Robinia pseudoacacia*. *Theor Appl Genet* **108**:836-841.
- Liu, Z., and G. Furnier. 1993. Comparison of allozyme, RFLP, and RAPD markers for revealing genetic variation within and between trembling aspen and bigtooth aspen. *Theor Appl Genet* **87**:97-105.
- Lynch, M., and K. Ritland. 1999. Estimation of pairwise relatedness with molecular markers. *Genetics* **152**:1753-1766.
- Mayes, S. G., M. A. McGinley, and C. R. Werth. 1998. Clonal population structure and genetic variation in sand-shinnery oak, *Quercus havardii* (Fagaceae). *Am J Bot* **85**:1609-1617.

- McCune, B., and J. B. Grace. 2002. Analysis of Ecological Communities. 2nd edition. MjM Software Design, Gleneden Beach, Oregon USA.
- McCune, B., and M. Mefford. 2006. PC-ORD 5.0. Multivariate analysis of ecological data. Gleneden Beach, Oregon, USA.
- Mielke Jr, P. W. 1984. 34 Meteorological applications of permutation techniques based on distance functions. *Handbook of statistics* **4**:813-830.
- Mitton, J. B., and M. C. Grant. 1980. Observations on the ecology and evolution of quaking aspen, *Populus tremuloides*, in the Colorado Front Range. *Am J Bot* **67**:202-209.
- Mock, K. E., C. A. Rowe, M. B. Hooten, J. Dewoody, and V. D. Hipkins. 2008. Clonal dynamics in western North American aspen (*Populus tremuloides*). *Mol Ecol* **17**:4827-4844.
- Mooi, E. A., and M. Sarstedt. 2011. A concise guide to market research: The process, data, and methods using IBM SPSS statistics. Springer.
- Namroud, M. C., A. Park, F. Tremblay, and Y. Bergeron. 2005. Clonal and spatial genetic structures of aspen (*Populus tremuloides* Michx.). *Mol Ecol* **14**:2969-2980.
- O'Brien, R. M. 2007. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity* **41**:673-690.
- Parks, J. C., and C. R. Werth. 1993. A study of spatial features of clones in a population of bracken fern (*Pteridium aquilium*). *Am J Bot* **80**:537-544.
- Pasteur, N., G. Pasteur, F. Bonhomme, J. Catalan, and J. Britton-Davidian. 1988. Practical isozyme genetics. Ellis Horwood Ltd.
- Peakall, R., D. Ebert, R. Cunningham, and D. Lindenmayer. 2006. Mark-recapture by genetic tagging reveals restricted movements by bush rats (*Rattus fuscipes*) in a fragmented landscape. *J Zool* **268**:207-216.
- Peakall, R., and P. E. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28**:2537-2539.
- Peakall, R. D., and P. E. Smouse. 2006. Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* **6**:288-295.
- Pellis, A., I. Laureysens, and R. Ceulemans. 2004. Genetic variation of the bud and leaf phenology of seventeen poplar clones in a short rotation coppice culture. *Plant biology* **6**:38-46.
- Perala, D. A. 1990. *Populus tremuloides* Michx. In: Silvics of North America. USDA Forest Service, Washington, D.C.
- Poly, W. J. 1997a. Nongenetic variation, genetic-environmental interactions and altered gene expression. I. Temperature, photoperiod, diet, pH and sex-related effects. *Comp Biochem Physiol Part B: Biochem Mol Biol* **117**:11-66.
- Poly, W. J. 1997b. Nongenetic variation, genetic-environmental interactions and altered gene expression. III. Posttranslational modifications. *Comp Biochem Physiol Part A: Physiology* **118**:551-572.

- Poly, W. J. 1997c. Nongenetic Variation, Genetic–Environmental Interactions and Altered Gene Expression. II. Disease, Parasite and Pollution Effects. *Comp Biochem Physiol Part B: Biochem Mol Biol* **117**:61-74.
- Queller, D. C., and K. F. Goodnight. 1989. Estimating relatedness using genetic markers. *Evolution* **43**:258-275.
- Rédei, G. P. 2008. Encyclopedia of genetics, genomics, proteomics, and informatics. 3rd edition. Springer, New York City, NY.
- Ritland, K. 1996. Estimators for pairwise relatedness and individual inbreeding coefficients. *Gen Res* **67**:175-186.
- Ritland, K. 2000. Marker-inferred relatedness as a tool for detecting heritability in nature. *Mol Ecol* **9**:1195-1204.
- Rohde, A., V. Storme, V. Jorge, M. Gaudet, N. Vitacolonna, F. Fabbri, T. Ruttink, G. Zaina, N. Marron, and S. Dillen. 2011. Bud set in poplar–genetic dissection of a complex trait in natural and hybrid populations. *New Phytol* **189**:106-121.
- Romme, W. H., M. G. Turner, and R. A. Reed. 2005. Establishment, Persistence, and Growth of Aspen (*Populus tremuloides*) Seedlings in Yellowstone National Park. *Ecology* **86**:404-418.
- Smith, E. A., S. B. Collette, T. A. Boynton, T. Lillrose, M. R. Stevens, M. F. Bekker, D. Eggett, and S. B. St Clair. 2011. Developmental contributions to phenotypic variation in functional leaf traits within quaking aspen clones. *Tree Physiol* **31**:68-77.
- Smulders, M. J. M., J. Van Der Schoot, P. Arens, and B. Vosman. 2001. Trinucleotide repeat microsatellite markers for black poplar (*Populus nigra* L.). *Mol Ecol Notes* **1**:188-190.
- St Clair, S. B., K. E. Mock, E. M. LaMalfa, R. B. Campbell, and R. J. Ryel. 2010. Genetic contributions to phenotypic variation in physiology, growth, and vigor of western aspen (*Populus tremuloides*) clones. *Forest Sci* **56**:222-230.
- Steinbach, M., G. Karypis, and V. Kumar. 2000. A comparison of document clustering techniques. Pages 525-526 in *KDD workshop on text mining*. Boston.
- Stenecker, G. 1973. The size of trembling aspen (*Populus tremuloides* Michx.) clones in Manitoba. *Can J Forest Res* **3**:472-478.
- Sunnucks, P. 2000. Efficient genetic markers for population biology. *Trends Ecol Evol* **15**:199-203.
- Suvanto, L. I., and T. B. Latva-Karjanmaa. 2005. Clone identification and clonal structure of the European aspen (*Populus tremula*). *Mol Ecol* **14**:2851-2860.
- Sydes, M. A., and R. Peakall. 1999. Extensive clonality in the endangered shrub *Haloragodendron lucasii* (Haloragaceae) revealed by allozymes and RAPDs. *Mol Ecol* **7**:87-93.
- Taberlet, P., and G. Luikart. 1999. Non-invasive genetic sampling and individual identification. *Biol J Linn Soc* **68**:41-55.
- Tamhane, A. C. 1977. Multiple comparisons in model-I one-way ANOVA with unequal variances. *Comm Statist Theory Methods* **6**:15-32.

- Tuskan, G., K. Francis, S. Russ, W. Romme, and M. Turner. 1996. RAPD markers reveal diversity within and among clonal and seedling stands of aspen in Yellowstone National Park, USA. *Can J Forest Res* **26**:2088-2098.
- Van Der Schoot, J., M. Pospíšková, and B. Vosman. 2000. Development and characterization of microsatellite markers in black poplar (*Populus nigra* L.). *Theor Appl Genet* **101**:317-322.
- Waits, L. P., G. Luikart, and P. Taberlet. 2001. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Mol Ecol* **10**:249-256.
- Weih, M. 2009. Genetic and environmental variation in spring and autumn phenology of biomass willows (*Salix spp.*): effects on shoot growth and nitrogen economy. *Tree Physiol* **29**:1479-1490.
- Weiner, J. 2004. Allocation, plasticity and allometry in plants. *Perspect Plant Ecol Evol Syst* **6**:207-215.
- Williams, J. G., A. R. Kubelik, K. J. Livak, J. A. Rafalski, and S. V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* **18**:6531-6535.
- Wyman, J., A. Bruneau, and M.-F. Tremblay. 2003. Microsatellite analysis of genetic diversity in four populations of *Populus tremuloides* in Quebec. *Can J Bot* **81**:360-367.
- Yao, Y., and I. Kovalchuk. 2011. Abiotic stress leads to somatic and heritable changes in homologous recombination frequency, point mutation frequency and microsatellite stability in *Arabidopsis* plants. *Mutat Res* **707**:61-66.
- Zane, L., L. Bargelloni, and T. Patarnello. 2002. Strategies for microsatellite isolation: a review. *Mol Ecol* **11**:1-16.