

REPRODUCTIVE DYNAMICS AND FINGERPRINTING EFFORT IN A  
DOUGLAS-FIR SEED ORCHARD

by

Joseph Anthony Kess  
B.Sc. University of Alberta 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate Studies  
(Forestry)

THE UNIVERSITY OF BRITISH COLUMBIA  
(Vancouver)

October 2012

© Joseph Anthony Kess, 2012

## ABSTRACT

Seed orchards are the tree improvement programs' production populations used to predictably package genetic gain and diversity achieved during the breeding cycle. Genetic gain and diversity delivered by seed orchards is calculated under the assumption of reproductive randomness, equality, and synchrony. These ideal expectations are not fulfilled by any existent seed orchards and deviations in gametic contribution by orchard parents' makes genetic gain and diversity unpredictable. In this study, five Douglas-fir (*Pseudotsuga menziesii*) microsatellite markers (Slavov et al 2004) were used to genotype 66 orchard parents, 14 of which were also supplemental mass pollination (SMP) pollen donors, and 396 bulk seeds from the 2009 seed crop of a wind-pollinated Douglas-fir seed orchard. Genotype data were analyzed using the likelihood based CERVUS parentage analysis program (Kalinowski et al 2007) for full pedigree reconstruction. In this orchard, 14% of paternal gametic contributions came from outside males. Parental balance curves showed that 80% of paternal, maternal, and gametic contributions were made by 38 (58%), 34 (52%) and 37 (56%) orchard parents, indicating that the greatest gametic contribution inequality was attributable to maternal gametic contribution. Differences in gametic contribution and common ancestry between orchard parents decreased the effective number of males, females, and population size to 42, 37, and 41, lower than the census number of 66 parents. Selfing was 24.24%, higher than that reported for many Douglas-fir seed orchards. High selfing may be attributed to reproductive asynchrony or differences in parental reproductive output. Supplemental mass pollination did not result in significantly higher paternal gametic contribution. Failure of SMP may be attributed to either incorrect timing of application or competition with ambient pollen. The minimum number of genotyped seeds required for accurate contamination estimate was 150, identified by jackknife sampling of the total genotyped seed sample.

## TABLE OF CONTENTS

<b>ABSTRACT .....</b>	<b>ii</b>
<b>TABLE OF CONTENTS .....</b>	<b>iii</b>
<b>LIST OF TABLES .....</b>	<b>v</b>
<b>LIST OF FIGURES .....</b>	<b>vi</b>
<b>ACKNOWLEDGEMENTS.....</b>	<b>vii</b>
<b>INTRODUCTION.....</b>	<b>1</b>
Tree Improvement.....	1
Quantitative Traits .....	1
Heritability.....	2
Genetic Quality .....	3
Genetic Gain.....	3
Genetic Diversity .....	4
Breeding Cycle .....	6
Seed Orchards .....	7
Seed Orchard Assumptions.....	7
Seed Orchard Management .....	7
Pollen Contamination .....	9
Detection of Pollen Contamination .....	10
Parentage Assignment Methods .....	13
<b>OBJECTIVES .....</b>	<b>13</b>
<b>MATERIALS AND METHODS .....</b>	<b>15</b>
Study Population .....	15
DNA Isolation.....	15
PCR Protocol.....	16
Product Visualization and Genotyping.....	17
Parentage Analysis.....	17
Gametic Contribution and Effective Population Size .....	18
Estimation of Supplemental Mass Pollination Effectiveness .....	20
Determination of Minimum Fingerprinting Effort .....	20

<b>RESULTS AND DISCUSSION .....</b>	<b>21</b>
Pollen Contamination .....	21
Parental Gametic Contribution and Effective Population Size.....	22
Self Fertilization .....	25
Estimation of Supplemental Mass Pollination Effectiveness .....	27
Determination of Minimum Fingerprinting Effort.....	29
<b>CONCLUSION .....</b>	<b>31</b>
<b>WORKS CITED.....</b>	<b>32</b>

## LIST OF TABLES

Table 1. Primer set, annealing temperature and genetic information of 5 Douglas-fir microsatellite loci (Slavov et al 2004).....	19
Table 2. Mean contamination estimate and confidence intervals of randomly sampled sets of genotyped seeds under low, medium and high contamination levels .....	29

## LIST OF FIGURES

Figure 1. Parental balance curves depicting cumulative gametic contribution vs number of parents contributing genetic material .....	23
Figure 2. Distribution and size of male and female half sib families of 66 orchard parents .....	24
Figure 3. Comparison of ranking for maternal and paternal gametic contribution between orchard parents .....	25
Figure 4. Correlation between selfing and maternal, parental and paternal gametic contribution .....	26
Figure 5. Contamination rate and standard deviation of contamination estimates detected from 30 randomly sampled seed genotype sets of different population sizes .....	30

## **ACKNOWLEDGEMENTS**

This thesis has been a difficult but rewarding process, and it has been finished due to the combined effort of many people, all of whom have my gratitude. First, I would like to thank my supervisor, Dr. Yousry El-Kassaby, whose advice, support, humor, considerable knowledge of forest genetics, and editing prowess made completion of this project possible. I would especially like to thank him for providing me with the opportunity to conduct this research project in April 2010 and allowing me to continue my studies at UBC. I would also like to thank my committee members Dr. Richard Hamelin and Dr. Michael Stoehr for their supervision. I would like to thank the following people for their help with my lab work, without whose efforts I would still be genotyping seedlings right now: Irena Fundova, for her help in teaching me the techniques used in this project, Dr. Tomas Funda, for his instruction in laboratory techniques, knowledge of forest genetics, support through difficult periods in this project, and patience with answering many questions, sometimes at unreasonable hours, and Ben Lai, whose experience, humor, management of the lab and sharing of the secrets of mass genotyping are greatly appreciated. This project would have been impossible without the help I received from summer students Xiao Song, Jiayin Song and Blaise Ratcliffe. I am also grateful to Dr. Jaroslav Klapste for his help with the statistical analysis in this project, as well as his design of the images in this thesis, and his easy-going nature. I thank Dr. Cindy Prescott and Gayle Kosh for their guidance through my program. I am also grateful to Simren Brar for her advice and friendship during the journey through graduate school, as well as the many other friends I've made in classes and in labs during my time at UBC. I would like to thank my friends outside the lab for their constant support and confidence. Finally, and most of all, I would like to thank my family and my wonderful partner Lynn Kane for their love, support, reassurance, and patience; I owe any success here primarily to their efforts.

## **INTRODUCTION**

### **Tree Improvement**

The purpose of tree improvement programs is to develop genetically and phenotypically superior seed stock for reforestation programs and future harvest. These programs are designed to identify desirable forms of quantitative traits from the existing genetic variation within natural populations and to reliably recombine these superior alleles in production populations to enhance phenotype in seed used for reforestation. This enhancement is achieved by matings among select parent trees that are genetically superior to the natural populations from which they were drawn (Zobel and Talbert 1984). Initially, improvement programs focused solely on individual trait enhancement without consideration for genetic diversity or the danger of exhausting the genetic potential of selected populations, leading to difficulty in sustaining future improvement (Zobel 1976). Current improvement programs have been designed with consideration for future breeding and environmental scenarios. The objectives of current programs are to enhance the economic value of specific traits of interest while conserving genes identified as potentially advantageous under future environmental regimes, and maintaining a threshold level of genetic diversity in the offspring for future selection and environmental adaptation (Burdon and Shelbourne 1971; Woods et al 1996; Forest Genetics Council of British Columbia 2009).

### **Quantitative Traits**

Traits targeted for tree improvement programs are primarily determined by the additive effect of many genes with small individual contribution to a trait. Traits under quantitative control show a continuous range of possible phenotypes due to the cumulative contribution of many genes, each of which comprise a small portion of phenotypic variability. While some alleles of quantitative genes at any given locus for a trait are of greater phenotypic effect than another allele at the same locus, when observing the genotype across all loci for a trait, most individuals within a population will likely possess only some of these alleles of



greater effect and some of lesser effect in multiple combinations across all loci, leading to a continuous range of phenotypes (Hartl and Clark 1997). The low probabilistic likelihood of many individuals possessing several alleles of greater effect or lesser effect creates a normal distribution in which most individuals in a population have a phenotypic value that is within one standard deviation of the population mean. Those few individuals that do possess many superior alleles of great effect or many inferior alleles of lesser effect for a given trait across all loci contributing to that trait are expected to be rare within the population and exhibit a phenotypic value that is far greater or less than the population mean (Charlesworth and Charlesworth 2010). Individuals exhibiting these rare phenotypes for economically desirable traits are selected for improvement programs.

### **Heritability**

Differences in phenotype of continuous traits are also affected by environmental variability; these effects need to be accounted for in order to identify how much variability in a trait is under genetic control. Determining the slope of the line of regression from plots of phenotypic values between parents and offspring provides an estimate of heritability of a continuous trait. Heritability ( $h^2$ ) refers to the genetic contribution made to a trait across all loci after accounting for environmental variance; it is expressed as the ratio of genetic variance ( $\sigma^2_g$ ) to phenotypic variance ( $\sigma^2_{ph}$ ) within the population studied. Additive genetic variance ( $\sigma^2_a$ ) is the amount of cumulative genetic variance in a trait that can be reliably transmitted to the next generation, and is the component of genetic variance relevant for tree improvement (Eriksson et al. 2006). Dominance and epistatic interactions between alleles also account for genetic variability of traits within individuals, but the phenotypes that result from these interactions cannot be reliably transmitted to the next generation because they are dependent on the interaction of unique alleles at specific loci. The effect of dominant and epistatic alleles cannot be maintained across multiple allelic combinations at a locus and

are not considered to be suitable for selection in improvement programs, because they require inheritance of a specific haplotype (Hartl and Clark 1997).

### **Genetic Quality**

The overall genetic quality of a tree improvement program is assessed on the criteria of genetic gain and genetic diversity. Of primary importance for the economic improvement in a crop is genetic gain, which determines how greatly improved a seed crop is compared to a natural founding population as a result of breeding. Much like directional selection in natural populations, selective breeding enhances the frequency of superior quantitative alleles in a population. However, emphasizing genetic gain comes at the expense of genetic diversity, because only a subset of the natural population is represented in the breeding population of an improvement program. Some genetic variation in natural populations is excluded in improvement programs, limiting the adaptability of the resulting population and increasing the risk of inbreeding (El-Kassaby and Ritland 1996). Because of these two opposing objectives, tree improvement programs are managed to both enhance genetic gain and maintain a baseline level of genetic diversity (Funda and El-Kassaby 2012).

### **Genetic Gain**

To achieve genetic gain, tree improvement programs are designed to increase the frequency of superior additive alleles for a trait in a population of selected parents and in the subsequent offspring generation, relative to the original base population. By increasing the frequency of superior alleles across all loci for a trait, the phenotypic mean of the population will shift to a greater value than that of the natural population from which parents were selected (Eriksson 2006). This shift in the mean phenotype due to changing the frequency of alleles contributing to a quantitative trait is referred to as genetic gain. The change in phenotypic mean as a result of artificial selection is determined by a population's response to selection ( $R$ ), which is calculated by multiplying the heritability ( $h^2$ ) of a trait by the selection differential ( $S$ ), which is the difference in mean phenotype between

selected parents and the base population (Hartl and Clark 1997). For a tree improvement program, the genetic gain expected from a production population (seed orchard) is determined by the heritability of a trait, and the breeding value of each individual in the seed orchard, weighted by the gametic contribution of each of these parents. These values are also adjusted to account for loss of genetic gain caused by contamination (gene flow) from outside unselected genotypes (Stoeher et al. 2004; Woods et al. 1996). Breeding value can be thought of as the individual deviation in phenotype for each parent tree compared to the population mean, which is determined by the quality of each allele it possesses that contributes to a quantitative trait. Since it is difficult to directly identify all quantitative trait loci and the alleles each parent tree possesses at each locus, parent breeding values are assessed by measuring average phenotypic improvement obtained from progeny tests. Because not all parents included in improvement programs have equal breeding values, the greater the contribution of the best performing parents, the greater the genetic gain. Maximum gain could technically be achieved by clonal propagation of one individual tree with the “best” alleles at all loci that contribute to a trait. However, this overemphasis on genetic gain will reduce the genetic diversity and future adaptability of an improved population (Funda and El-Kassaby 2012).

### **Genetic Diversity**

Genetic diversity is also important for improvement programs, because it acts as insurance for adaptability to future environmental conditions and breeding scenarios (Zobel and Talbert 1984). Genetic diversity is determined by the richness and variability of a population or species' genetic base. Diversity is a prerequisite for evolutionary response to selection because selection acts on superior variants of alleles; more alleles available for selection allows for greater response to selection. However, selection or random genetic drift will reduce diversity as some alleles favored by selection or by chance will be preferentially passed on to the next generation. Over multiple generations, this process decreases the number of different alleles at a locus. The overall genetic diversity

of a population can be estimated by observing the average number of alleles per locus, the number of loci with multiple alleles, or heterozygosity. Genetic diversity can also be inferred from calculating the probability that alleles are identical by descent; if any two alleles of the same gene are identical, there is a chance they are descended from the same ancestral allele. Because selection or drift cause some alleles to be passed on more successfully, identity by descent increases each generation as these alleles are more likely to be identical copies of the same ancestral allele. Genetic diversity decreases because alleles identical by descent must also be identical in state; loss of genetic diversity by increasing probability of identity by descent will also decrease measures of polymorphism and heterozygosity (Hartl and Clark 1997). An ideal population is a hypothetical population of infinite size experiencing no drift, selection, mutation, or reproductive imbalance. In an ideal population, genetic diversity will not decrease over time. However, as no population is infinite, some random fluctuation in allele frequencies due to drift is expected to occur (Andrews 2010). In practically every population, other evolutionary factors such as mutation, migration and reproductive imbalance are also occurring. Under these circumstances, the census size of a population does not reflect the population's actual reproductive dynamics; changes in allele frequency in the offspring generation do not reflect the allele frequencies in the census population due to random drift alone. Effective population size is the size of an ideal population that would explain the allele frequencies in the offspring generation due to drift (Wright 1931). Effective population size differs depending on how it is calculated; there are multiple formulae to estimate population size depending on the ideal population assumption being violated.

In tree improvement programs, there is often reproductive imbalance, migration from outside the breeding population, and a level of coancestry (alleles identical by descent) among individuals in the selected population. Measures of effective population size relevant to breeding programs have been devised to account for these factors, and production populations are managed in an effort to keep effective population size as high as possible. Genetic diversity is managed at the

level of effective population size, as reduction in effective population size will also decrease values for all other measures of diversity (Lindgren et al. 1996).

### **Breeding Cycle**

Tree improvement programs require controlled, recurrent cycles of artificial selection on traits of interest, in which the genetic contribution from candidate parents in each subsequent cycle is modified with information gained in previous cycles (White 1987). Before individuals enter a production population, they are evaluated in a breeding cycle in three steps: selection, breeding and testing (Funda and El-Kassaby 2012). The purpose of a separate breeding cycle is to evaluate the genetic quality of individual genotypes as candidates for entry into production and subsequent breeding populations (Zobel and Talbert 1984). In the selection step, trees exhibiting desirable phenotypes are selected from natural populations. In subsequent breeding cycles, these trees can also be selected from previously tested parents. These trees are expected to possess above average phenotypes due to corresponding genetic quality. However, genetic quality cannot be assumed from phenotype, especially when selecting trees from natural stands. Parent trees can be tested for genetic worth by identifying whether superior phenotype can be genetically transmitted to the offspring generation. These trees undergo breeding to recombine genes in the offspring generation; crosses between parent trees in this stage are controlled and recorded to ensure that the offspring of each cross can be matched to both parent trees. Offspring are then tested to determine whether the phenotypes of the parent trees can be attributed to superior genetic worth – parent trees that produce the best offspring consistent over multiple crosses are deemed genetically superior and can then be incorporated into a production population (i.e., backwards selection) (Zobel 1971). Additionally, superior individuals from the progeny testing phase can also be selected for incorporation in either the production or the breeding population (i.e., forward selection); however, their superiority requires further testing to parse out the degree of genetic and environmental effects.

## **Seed Orchards**

Seed orchards are production populations designed to capture the genetic value present in selected superior parent trees, and to package that genetic material in superior seed. These populations serve as the link between planned breeding programs and reforestation efforts because they act as a mass production system of genetic material selected and verified in breeding and testing cycles. Ideally, each seed produced by a seed orchard will possess superior genetic material in a quantity that reflects the genetic value of the parental generation. However, factors affecting reproductive output equality, synchrony, and isolation from undesirable gene flow (pollen) often compromise these estimates (El-Kassaby 1989).

## **Seed Orchard Assumptions**

The genetic gain and diversity of a seed orchard are calculated under the assumption of reproductive isolation, equality and synchrony. Seed orchards should perform as ideal populations meeting Hardy Weinberg expectations; allele frequencies between generations are not changed due to the effects of random genetic drift, differences in reproductive timing or success, mutation, or gene flow. Any deviation from these ideal conditions will alter the allele frequency and genetic quality of seed crops, because superior alleles cannot be assured transmission into the offspring crop, despite inclusion in the parental breeding population (Eriksson et al. 2006). Substantial deviation in all of these assumptions have been identified (El-Kassaby 1989, El-Kassaby et al. 2010); estimates of genetic quality are more accurately calculated by weighting breeding value with estimates of reproductive success and outside gametic contribution (Stoehr et al. 2004).

## **Seed Orchard Management**

Reproductive asynchrony and parental gametic contribution imbalances are common in seed orchards, and detract from achieved genetic gain. Reproductive asynchrony can be a result of differences in reproductive

phenology; asynchrony in pollen production and receptivity will create small subpopulations of breeding individuals, preventing total panmixia of all orchard parents (El-Kassaby 1995; Funda and El-Kassaby 2012). Variability in reproductive phenology is under genetic control, and so individual parents selected from differing climates will still require a genetically determined level of heat-sum accumulation before pollen production and receptivity, creating phenological differences within the seed orchard (Worral 1983). Seed orchards can be “cooled” by spraying mist from overhead irrigation systems in a process called bloom delay, used to delay heat sum accumulation to a period in which heat sum can be accumulated rapidly by all parents in the orchard (Silen and Keane 1969). This process will synchronize pollen production and receptivity for orchard parents to promote outcrossing within the seed orchard (Fashler and El-Kassaby 1987). Bloom delay will also create asynchrony in reproductive phenology between the seed orchard and surrounding stands contributing contaminant pollen, decreasing undesirable gene flow from these sources (El-Kassaby and Ritland 1986).

Gametic contribution imbalances will also decrease the level of panmixia and genetic gain achieved by a seed orchard. To increase the representation of desirable genotypes in the seed orchard, pollen from selected parents may also be applied to all receptive clones in the orchard (Wakeley et al. 1966). This process, called supplemental mass pollination (SMP), has been found to increase outcrossing rate in the seed orchard, especially for parents that may not have access to ambient pollen due to phenology differences (Askew 1986). Supplemental mass pollination has also been found to decrease levels of contamination within the seed orchard by outcompeting ambient contaminant pollen (Bridgwater et al. 1993). The application of pollen from select genotypes can also increase the genetic gain in the seed orchard, and can facilitate panmixia by increasing the gametic contribution of parents with otherwise low reproductive output (Stoehr et al. 2006). However, the success of SMP treatment has varied between studies, ranging between 4 and 80% effectiveness (Yazdani et al. 1986; Bridgwater et al. 1987). Treatment effectiveness is

influenced by the timing of pollen application (Owens et al. 1981), the number of applications (El-Kassaby et al. 1993), volume of pollen applied (Webber and Painter 1996), the level of ambient pollen in the orchard (Webber and Yeh 1987), the genetic quality of applied pollen (Nakamura and Wheeler 1992), pollen receptivity of parents within the orchard, the effect of other orchard management practices (El-Kassaby and Ritland 1986), and the amount of receptive female cones (El-Kassaby et al. 2010).

### **Pollen Contamination**

Under 100% gene flow, seed orchards' crops are expected to deliver 50% of their genetic gain potential; thus gene flow is considered to be the most serious threat to genetic gain capture (Fast et al. 1986; Wheeler and Jech 1986). Contribution of genetic material from outside the seed orchard impedes tree improvement efforts by diminishing the realized genetic gain achieved from breeding and selection. Wind pollination facilitates random mating of parents within a seed orchard, but also allows male genetic contributions from outside, unselected stands to contribute genes to the seed crop. Species in tree improvement programs have developed specialized adaptations to pollen morphology that enable transmission of pollen over considerable distances, promoting unwanted gene flow from outside donors (Adams and Birkes 1989). The inclusion of genes from unselected parents is expected to diminish genetic gain, because trees from outside seed orchards frequently have lower breeding values (Eriksson et al. 2006). Although genetic gain is depressed in contaminated orchards, genetic diversity may be enhanced. This change in genetic diversity depends on the extent of pollen contamination, the amount of genetic variability within stands contributing contaminant pollen, and the reproductive success of every individual contributing contaminant pollen (Adams and Burczyk 2000). Though pollen contamination can increase genetic diversity, many alleles introduced may be detrimental to the survival of the seed crop. Orchard parents have been selected partially for their adaptation to specific seed planting zones; the introduction of alleles from outside these zones can lead to the loss of fitness and survival



effects expected from the inheritance of adaptive groups of additive alleles. This decrease in fitness due to gene flow is referred to as outbreeding depression (Nikkanen 2002). The severity of outbreeding depression depends on the differences in local adaptation between seed orchard parents and surrounding stands, and the extent to which these adaptive benefits are phenotypically plastic. In some instances, pollen contamination can negate all gains achieved from breeding by generating seed that is completely maladapted to its planting zone. An example of this maladaptation has been observed in a Scots pine (*Pinus sylvestris*) seed orchard; stands surrounding this orchard are adapted to milder climatic conditions than required for planting enhanced seed. The outbreeding depression caused by pollen contamination decreased survival rate of seedlings produced from these seeds compared to those from natural stands within the planting zone, making them unsuitable for reforestation in their planting zone (Pulkkinen et al. 1995).

### **Detection of Pollen Contamination**

Estimates of pollen contamination are variable, ranging from nearly zero (El-Kassaby and Ritland 1986) to 90% (Fast et al. 1986). Pollen traps are used to compare the level of pollen produced within an orchard to the background level of pollen originating in surrounding stands. These contamination estimates tend to be high, and reflect only the difference in pollen production between seed orchards and surrounding stands (Slavov et al 2004). While reproductive output and success are correlated, high pollen production cannot guarantee fertilization, so measures of pollen contamination obtained via pollen traps can only provide an estimate of actual gene flow. Comparisons of pollen levels outside and within the seed orchard cannot provide insight into the variable levels of reproductive success between these pollen sources; a high estimate of pollen contamination may still lead to low fertilization success, depending on the reproductive success of that pollen relative to the reproductive success of within-orchard pollen. While current pollen trap methods show concordance with molecular estimates that directly measure reproductive success, the correlation between these estimates

may change as molecular marker based methods are refined to allow greater resolution of reproductive success.

Molecular markers can be used to estimate contamination and reconstruct pedigrees because they are reliably inherited between generations. Parentage can be assigned to individuals based on observed inheritance of molecular markers. In this method, individuals in a seed sample are assigned to each parent based on inheritance of specific alleles of a given marker. Assigning parentage to a small sample of seed from a seed orchard can estimate the overall level of gene flow from outside the seed orchard. The haploid megagametophyte tissue surrounding a seed germinant can be genotyped to determine the maternal gametic contribution; all maternal haplotypes will originate from within the orchard, barring any labeling errors of seed samples or ramets. Seed germinant tissue is then analyzed for paternally contributed haplotypes; germinants with genotypes that cannot originate from a cross between parents within the orchard are assumed to be sired by outside males (Smith and Adams 1983).

Early molecular marker based estimations identified structural polymorphism in proteins. In this technique, called isozyme analysis, a multilocus genotype used for assigning parentage can be determined by electrophoretically analyzing which form of multiple isozyme proteins an individual possesses (Smith and Adams 1983). While this method allows direct observation of within and outside orchard reproductive success, the resolution provided by isozyme markers tends to be poor (Adams 1992). These markers are limited by their low levels of polymorphism; contaminant pollen may possess the same multilocus haplotype as gametes produced by orchard parents and will go undetected in contamination assessments. The failure to identify outside pollen sources due to poor resolution of molecular markers is referred to as cryptic gene flow, and will lead to reduced estimates of pollen contamination. The level of cryptic gene flow increases with the size of the population studied, and is inversely proportional to the polymorphism of the markers; this limitation has prevented application of isozyme analysis of large populations (Slavov et al 2004), and creates large

measures of standard error in pollen contamination estimates obtained using these markers (Adams and Bursczyk 2000).

DNA-based microsatellite markers provide higher resolution for monitoring pollen contamination and reproductive dynamics in seed orchards. These markers are variable regions of repeated nucleotides that can be amplified in polymerase chain reactions (PCR). While the size of the repeat region is variable between individuals, the flanking region is highly conserved between individuals and species, allowing design of region-specific primers. Regions of repeated nucleotides have a high rate of mutation caused by slipped-strand mis-pairing during DNA replication, making them highly polymorphic and well-suited for parentage analyses (Ashley 2010). Microsatellites follow a Mendelian inheritance pattern, allowing for easy interpretation of transmission between generations. An individual-specific DNA “fingerprint” can be generated using only a few highly polymorphic microsatellite markers (Dow and Ashley 1998), and studies of orchard contamination have successfully revealed gene flow and mating dynamics using only three highly variable microsatellite markers (Fernandes et al. 2008). These markers have been developed for most commercial species in improvement programs, and currently provide estimations of gene flow with the highest resolution per loci typed. However, there are still several accuracy issues with microsatellites that can bias parentage assignment and gene flow detection. Due to their high variability, microsatellites can experience between-generation mutations either in the hyper variable repeat region, creating a different allele, or in the conserved primer binding region, preventing amplification. Both of these mutations will create genotypic mismatches between generations, due either to the appearance of an allelic mismatch between parents and offspring, or the disappearance of an allele that should be present (null allele). Genotyping errors due to incorrect allele scoring are also frequent with microsatellites, leading to inflated contamination estimates, depending on how stringently genotypes are assigned to orchard and natural populations (Slavov et al. 2004).

## **Parentage Assignment Methods**

In addition to the type of molecular markers used, pedigree reconstruction and detection of gene flow between populations is dependent on the method used to assign parentage. Determining matches between parent haplotypes was initially conducted by an exclusion method that considers any multilocus haplotype that mismatches orchard males at one locus as a contaminant. There are few statistical assumptions made with this method, but it is highly sensitive to genotyping error, and can lead to unrealistically high contaminant levels, especially when considering the frequency of mutation, null alleles, and genotyping error inherent in using microsatellites (Jones et al. 2009). Likelihood methods, such as the one utilized in the program Cervus 3.0 (Kalinowski et al 2007), will instead calculate probabilities of parentage based on allele frequencies, and then assign offspring to the most likely parents. Cervus 3.0 employs a pairwise likelihood assignment method, in which every individual in the offspring population is compared to every parent individually. This method will never actually exclude non-parents, but these individuals will be assigned probability of parentage values that are significantly lower than parents assigned as the most likely (Marshall et al. 1998). Likelihood methods can also be adjusted to accommodate for genotyping error, mutation, and null alleles, and are helpful in compensating for some of the errors in genotyping caused by microsatellites (Kalinowski et al. 2007).

## **OBJECTIVES**

The goal of this study was to use pedigree reconstruction to determine the extent of reproductive imbalance and gene flow from outside sources into an advanced generation Douglas-fir (*Pseudotsuga menziesii*) seed orchard. Several studies have been conducted on pollen contamination in Douglas-fir seed orchards using molecular markers, but few have considered the minimum genotyping effort required given current advances in parentage assignment methods and the recent development of highly variable microsatellite DNA markers. The specific objectives of this study are to:

1. Measure the extent of pollen contamination of a Douglas-fir seed orchard from a bulk seed sample,
2. Reconstruct the seed orchard pedigree to identify deviations from panmixia, and
3. Determine minimum fingerprinting effort required to identify pollen contamination in a bulk seed sample.

## **MATERIALS AND METHODS**

### **Study Population**

Parents' bud tissue and random bulk seed were obtained from the 2009 seed crop of a second-generation clonal Douglas-fir seed orchard, operated by Western Forest Products Inc. and established in 1990. In 2009, this orchard consisted of 66 parents identified from either forward or backward selection for breeding in the British Columbia Ministry of Forests, Lands and Natural Resource Operations' low-elevation coastal Douglas-fir tree improvement program. This orchard is located on Vancouver Island in Saanichton, British Columbia (latitude 48°35'N, longitude 123°24'W, elevation 50 m). Parental ramets are arranged according to the permuted neighbourhood design described by Bell and Fletcher (1978). In addition to wind-pollination, supplemental (SMP) pollen of 14 within-orchard parents has been used to enhance the probability of reproductive success for these clones.

### **DNA Isolation**

Douglas-fir seeds were germinated following stratification to synchronize germination time following the protocol described by El-Kassaby and Edwards (1995). Germinated seeds were dissected into diploid embryo and haploid megagametophyte tissue. All tissue used for DNA extraction was stored at -80 °C. DNA was extracted from diploid parental buds (2n), diploid offspring embryo tissue (2n), and haploid megagametophyte (n) tissue following a modified version of the CTAB extraction method described by Doyle and Doyle (1987). Samples were placed in 1.5mL tubes and submerged in liquid nitrogen, then ground using bleach-sterilized plastic grinding pestles. Each sample was further ground in 800 µL warmed (65 °C) extraction buffer (0.1M CTAB, 0.1M Tris base (pH 8.0), 0.04M EDTA (pH 8.0), 1.4M NaCl, and 0.2% beta-mercaptoethanol), and incubated at 65 °C for one hour. Samples were gently vortexed every 10 minutes to break up clumped tissue. Following incubation, samples were centrifuged at 13,000 RPM for 10 minutes. The supernatant layer was

transferred to new 2.0 mL tubes and incubated at 37 °C for 45 minutes with 4 µL of 10mg/mL RNase A. After RNase incubation, 750 µL of 24:1 chloroform:isoamyl alcohol was added to each tube to remove proteins. Samples were mixed using a rotomixer for 15 minutes and then centrifuged at 13,000 RPM for 10 minutes to separate layers. The supernatant layer was transferred to new 1.5mL tubes and precipitated overnight at -20 °C with 400 µL 100% isopropanol. Precipitated DNA was concentrated by centrifugation at 11,000 RPM at 4 °C for 30 minutes. Isopropanol was poured off, samples were washed with 400 µL of 70% ethanol and left to dry on the bench. Following drying, samples were suspended in 50 µL of deionized DNase/RNase free water. DNA quality and quantity were assessed using agarose gel electrophoresis and spectrophotometry.

### **PCR Protocol**

In this study, 5 Douglas-fir microsatellite loci designed by Slavov et al (2004) were used to successfully genotype 396 offspring and 66 parents from a bulk sample of 544 seeds. DNA samples were amplified by polymerase chain reaction using a PE Applied Biosystems Gene Amp PCR System 9700 thermal cycler, and an Eppendorf Mastercycler gradient thermal cycler. Samples were initially denatured for 5 minutes at 95 °C, followed by another 33 cycles of 30 seconds of denaturing at 95 °C, 30 seconds at a primer-specific annealing temperature ( $T_a$ , Table 1), and 45 seconds of extension at 72 °C, and a 10 minutes extension at 72 °C following completion of the 33 amplification cycles. All PCR reactions were carried out in a final volume of 10 µL. Reactions for primer sets Pm\_OSU1F9, Pm\_OSU3F1, Pm\_OSU3G9 and Pm\_OSU2D4 contained 0.25µM of forward primer and 0.25µM of reverse primer, and all reactions for the Pm\_OSU2G12 primer set contained 0.125µM forward and 0.125µM reverse primer. All primer sets contained one tailed primer with an additional 19-20 nucleotide tail complementary to infrared-labeled M13 primers, to allow incorporation of M13 labeled primers into PCR products for visualization on polyacrylamide gel. All reactions contained 75ng template DNA, 0.25mM of all four dNTPs, 1X Amplitaq

Gold PCR buffer (Roche, Laval Que), 3.0mM MgCl<sub>2</sub>, 0.3 pmol of M13 labeled primers (LiCor Inc., Lincoln, NE), and 1 unit of Amplitaq Gold (Roche, Laval, Que).

### **Product Visualization and Genotyping**

After PCR, all samples were mixed with 2 µL stop dye (LiCor Inc., Lincoln, NE) and denatured at 95 °C for 3 minutes, and held at 4 °C. Denatured samples were separated by LiCor 4300 automated sequencer based on molecular weight corresponding to the number of base pair repeats, on 25 cm long, 0.4 mm thick, 6% Long Ranger polyacrylamide gels (LiCor Inc., Lincoln, NE). Megagametophyte DNA samples were run next to corresponding embryo DNA samples to ensure that the maternal contribution to the offspring genotype could be accurately inferred. A total of 544 offspring and 66 parents were genotyped by visual scoring using SAGA<sup>TM</sup> software (LiCor Inc., Lincoln, NE). Of the 544 offspring genotyped, 396 produced genotypes at a minimum of 3 loci and were sufficient for parentage analysis. Allele size was determined by 50 - 350 base pair sizing standards (LiCor Inc., Lincoln, NE) loaded on the gel. Parents were amplified in replicate reactions and run on multiple gels to ensure consistency in genotyping. Offspring genotypes were checked against parental reference genotypes, and re-genotyped twice to increase confidence in accuracy.

### **Parentage Analysis**

Parentage was assigned to offspring using the program CERVUS 3.0.3 (Kalinowski et al. 2007). Simulation of parentage analysis in CERVUS generating 10,000 offspring genotypes found that 100% of parents could be assigned with 5 loci, with 3 loci set as the minimum necessary for parentage assignment. Parentage was assigned at 95% confidence, accounting for genotyping errors (0.01), selfing and mutation. Maternity was determined by submitting diploid, homozygote doubles of haploid megagametophyte genotypes to parentage analysis without selfing in CERVUS to assign the two individuals most likely to have contributed the maternal haplotype. In all cases where



megagametophyte DNA was of sufficient quality for PCR with 3 or more loci, one of the two most likely assigned mothers was matched to one of the two most likely assigned parents in CERVUS parentage analysis with selfing, allowing for either identification of the most likely father, or the maternal contribution to a genotyped seed produced by outside-orchard pollen. Individuals that could be assigned a maternal parent, but were not assigned two parents in parentage analysis were considered to be the result of pollen contamination.

### **Gametic Contribution and Effective Population Size**

The pedigree of the seed orchard was reconstructed from molecular data. Proportional paternal, maternal and total gametic contributions were determined by direct count. These values were used to assess the level of reproductive imbalance within the seed orchard. Effective number of mothers, fathers, and population size was also calculated by weighting summed proportional gametic contribution by the average level of relatedness ( $r$ ) between parents within the seed orchard, using the equation:

$$N_E = \frac{1}{\sum_{i=1}^N \sum_{j=1}^N p_j p_i r_{ij}} \quad (1)$$

where  $p_i$  and  $p_j$  are proportional gametic contributions of parents  $i$  and  $j$ , and  $r_{ij}$  is the coefficient of relatedness between parents  $i$  and  $j$ , obtained from the available pedigree of the orchard parents. Relatedness is calculated as the average probability of possessing an autosomally inherited allele identical by descent (Charlesworth and Charlesworth 2010). Relatedness between an orchard parent and itself is 1.0, a parent or full sib is 0.5, a half sib is 0.25, and an unrelated individual is 0. Effective numbers of fathers and mothers in the seed orchard were calculated using the same equation, with proportional paternal or maternal gametic contribution replacing proportional total gametic contribution.

Table 1. Primer set, annealing temperature and genetic information of 5 Douglas-fir microsatellite loci (Slavov et al 2004).

Locus	Forward Primer (5'-3')	Reverse Primer (5'-3')	Repeat Motif	Allele Range (bp)	A	PIC	NE- PP	T <sub>a</sub> °C
Pm_OSU1 F9	CCTCATGC ATTGGACA CTC	GGATTCTTG AGCAGGTAG G	(AG)34	208 - 270	26	0.941	0.022	57
Pm_OSU3 F1	GACTAGAT CATCGCAA CTT	GGTATTCTT ATGGTTTTTA T	(AT)4...(TG)1 8 (AG)26	188 - 254	19	0.909	0.047	52
Pm_OSU3 G9	ATTCCTTT TGAGACCT ACTT	CTTCAAAAA TTCCTACAA CA	(TG) 12(AG)28	157 - 239	23	0.897	0.055	52
Pm_OSU2 D4	TTATTGCA CATGAGTA TTATGA	CAGATGTTG TTTTTTATAC CAC	(TG)6...(TG) 7 (AG)27...(AC) 4	141 - 207	27	0.928	0.031	50
Pm_OSU2 G12	CAAGGACT CATATGGG AAA	AACATCAGT AATAACCTTT T	(AC)11...(AC) 19...(GCAC) 5(GCAC)4(A C)7...(AC)6	251 - 303	19	0.873	0.078	50
M13 FPrimer	CACGACGTTGTAAAACGAC							
M13 RPrimer	GGATAACAATTTACACAGG							

**A**, number of alleles per locus identified in 66 orchard parents, **PIC**, polymorphic information content (Botstein et al. 1980), **NE-PP** Non-exclusion probability of a parent pair, **T<sub>a</sub> °C**, optimum annealing temperature for primers designed for a given locus.

### **Estimation of Supplemental Mass Pollination Effectiveness**

All parents donating pollen for supplemental mass pollination (SMP) were from the seed orchard, and thus the success of SMP compared to the reproductive success of ambient pollen from these parents could not be determined. However, a statistically significant difference in reproductive success of SMP parents compared to all parents donating only ambient pollen was used as an indirect measure of the success of SMP. The mean, standard deviation, error and confidence intervals of paternal reproductive success of all orchard parents, SMP parents, and parents donating only ambient pollen were compared at the 95% confidence level. SMP was considered successful if there was no overlap in mean and confidence interval between SMP parents and parents donating only ambient pollen.

### **Determination of Minimum Fingerprinting Effort**

Minimum fingerprinting effort was identified for low (6%), moderate (14%) and high contamination (54%) scenarios using simulated contaminant and within-orchard offspring genotypes. Simulated genotypes were generated by assigning one of either a set “contaminant” genotype that would not be matched to parents or a “within-orchard genotype” that would always match to parents to 400 individual offspring, so that while there were 400 offspring, there were only two genotypes within the actual sample. Samples of 100, 150, 200, and 250 offspring were randomly drawn without replacement using an R-based sampling program. A total of 30 random sets of offspring genotypes of each sample size were assigned parentage to determine mean pollen contamination level and standard error for each sample size. CERVUS 3.0 was used for estimation of pollen contamination level under the same parameters used for assigning parentage to the full sample of seeds. Confidence intervals were calculated for each sample size at 95% confidence level. Accuracy of simulated genotyping effort was confirmed by repeating random sampling and contamination estimating using actual seed genotype samples.

## RESULTS AND DISCUSSION

### Pollen Contamination

Of the 396 seeds genotyped, 55 could be assigned a maternal parent but did not assign to any paternal parent, indicating a 13.8% rate of pollen contamination. This estimate is close to the contamination rate (10.36%) of the 2005 seed crop from the same seed orchard (Lai et al. 2010; El-Kassaby et al. 2010) managed with supplemental mass pollination and bloom delay. This small increase in pollen contamination level could be attributed to differences in management practices between years – the current seed crop studied was not cooled for bloom delay, and had proportionally fewer supplemental pollen donors, potentially causing a slight increase in the reproductive success of ambient contaminant pollen. The absence of bloom delay in the orchard may have led to differences in reproductive phenology between parents. Because pollen production and receptivity has not been synchronized among clones in this orchard, parents may be separated into breeding subpopulations, which could include contaminant pollen. Due to the decreased number of pollen donors in each of these subpopulations, the gametic competition between orchard parents and outside parents is reduced, increasing the chance of gene flow. A multi-year (1999, 2000, 2003) study of a Douglas-fir seed orchard in Oregon using the same molecular markers identified an average contamination rate of 35.7% (Slavov et al. 2005). While the Oregon orchard studied by Slavov et al. is in close proximity to other Douglas-fir orchard blocks and natural stands and was not managed for SMP or bloom delay, the Saanichton BC orchard used in the current study is relatively isolated from contaminant pollen, and has been managed to increase reproductive success of orchard parents in both 2005 and 2009.

The assignment methods used in these studies may have also lead to different contamination estimates. Because contamination in Slavov et al.'s study was calculated using paternity exclusion as opposed to likelihood assignment, mutation, scoring errors and incomplete genotypes can compromise true

estimates of paternity in these methods. In exclusion based methods, efforts to account for genotyping and scoring errors require multiple mismatches to exclude a parent, but will require genotyping at several additional loci, which may actually compound genotyping and mutational mismatches, introducing statistical noise and leading to erroneous exclusion (Wang and Santure 2009). Conversely, likelihood methods never exclude individual parents, and may incorrectly assign parentage based on only limited genotypic information, artificially decreasing the contamination estimate. However, given the high polymorphic information content ( $PIC = 0.9097$ ) and low probability of erroneous non-exclusion of all loci used in this study, even a small amount of genotypic information (2-3 loci) has the potential to correctly assign parentage (Slavov et al. 2004).

While different parentage assignment methods might be responsible for some differences in the contamination estimates between orchards, differences in orchard isolation and management practices probably account for the high contamination level of the 2005 orchard. Previous studies in Douglas-fir orchards have illustrated the importance of bloom delay, isolation, and supplemental mass pollination in preventing contamination (Adams et al. 1997; El-Kassaby and Ritland 1986). Knowing that differences in management practices in the same orchard lead to the different estimates of contamination when assessed using the same loci and assignment method, it is more likely that differences in management are responsible for the differing contamination estimates between the orchard in the 2005 study and the orchard in the present study, though temporal variation in outside and within-orchard paternal fecundity cannot be discounted.

### **Parental Gametic Contribution and Effective Population Size**

Through pedigree reconstruction, the gametic contributions of each parental clone to the bulk seed sample were assessed in terms of maternal, paternal, and total gametic contribution (Figure 1). Of the 66 orchard parents, 64 (97%) made gametic contributions to the orchard crop, with 62 (94%) parents contributing

seed and 64 (97%) contributing pollen. Calculation of cumulative gametic contribution showed that 80% of gametes were contributed by 37 (56%) orchard parents. Cumulative maternal and paternal contribution assessment showed that 80% of pollen and seed were contributed by 38 (58%) and 34 (52%) parents, respectively (Figure 2). Of the 20 top parents ranked for maternal and paternal gametic contribution, 12 parents were included as the top 20 maternal and paternal gametic contributors (Figure 3). Maternal and paternal contribution showed positive correlation ( $r = 0.666$ ), but stronger correlation was observed between paternal and gametic contribution ( $r = 0.874$ ), and maternal and gametic contribution ( $r = 0.944$ ).

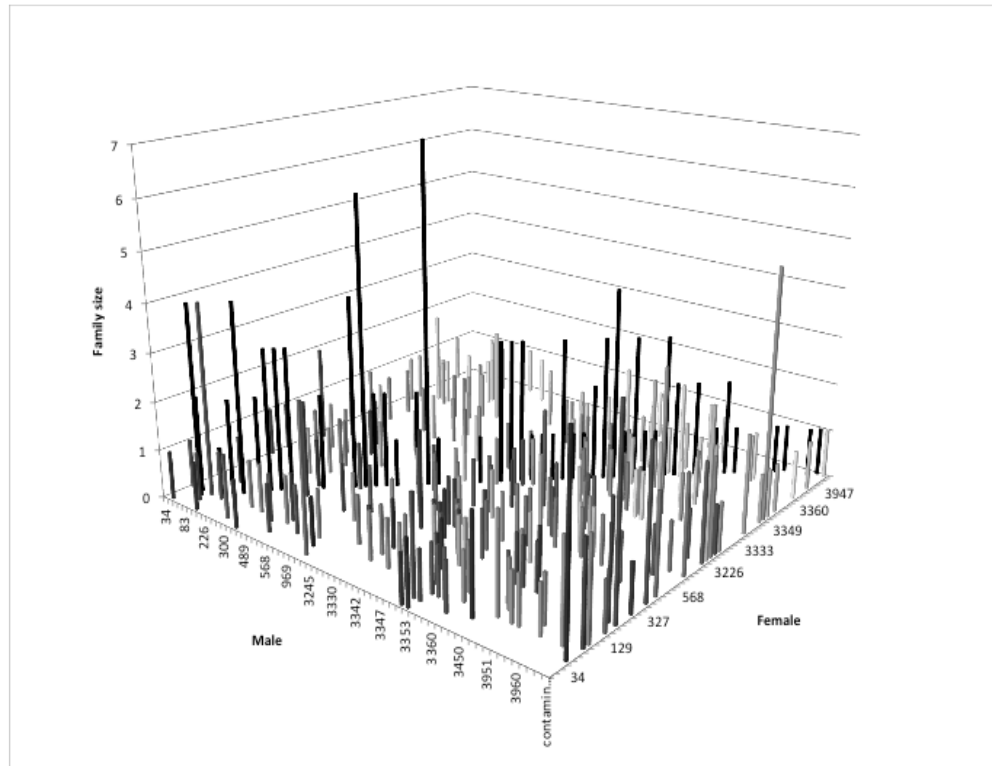


Figure 1. Distribution and size of male and female half sib families of 66 orchard parents.

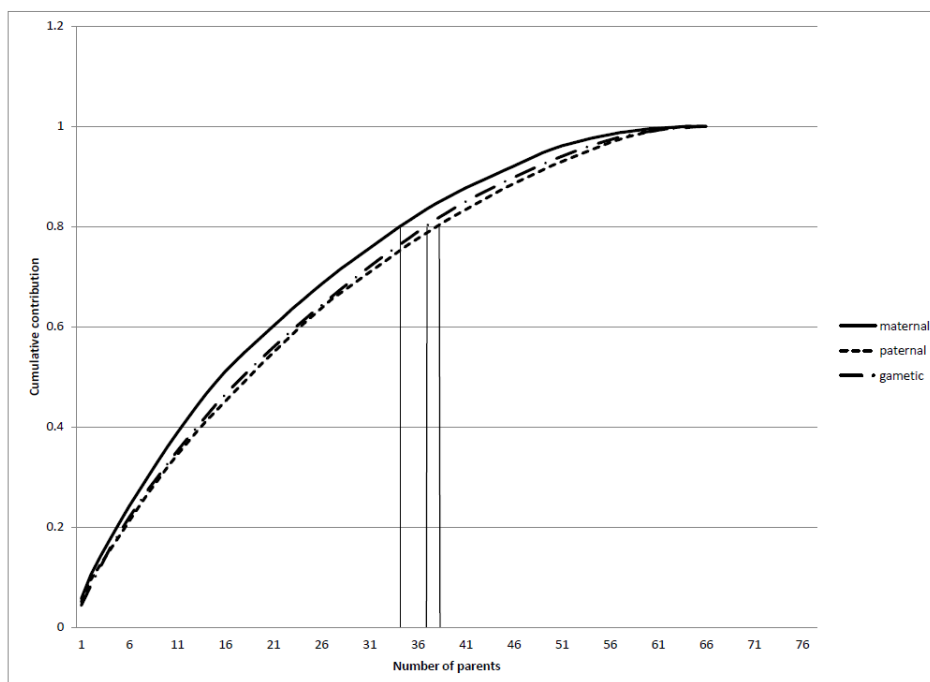


Figure 2. Parental balance curves depicting cumulative maternal, paternal and gametic contribution vs. number of contributing parents.

These differences in contribution indicate deviation in maternal reproductive success had greater effect on overall reproductive success compared to paternal reproductive success. Variations in female reproductive success are expected, as individual reproductive success differences have been identified previously in Douglas-fir seed orchards (El-Kassaby and Cook 1994), and have been shown to have greater variability than male reproductive success (El-Kassaby et al. 2010). Effective population size and number of effective males and females were smaller than the census number of orchard parents. The overall effective population size was 41, indicating that 62% of parents made the majority of gametic contributions. The effective male population size was 42, while the effective number of female parents was 37. All of these values are considerably higher than the threshold effective population size ( $N_e = 10$ ) required for capturing 95% of base population genetic diversity (Yanchuk 2001).

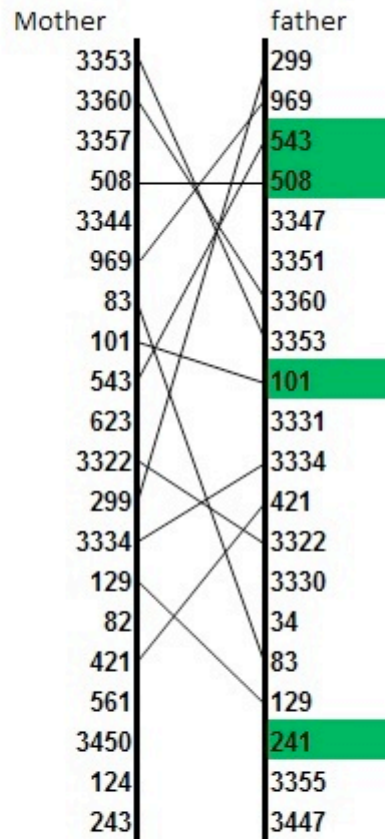


Figure 3. Comparison of ranking for maternal and paternal gametic contribution between orchard parents. (Parents contributing both ambient pollen and pollen for supplemental mass pollination (SMP) are highlighted in green).

### Self Fertilization

Of the 396 seeds assigned parents, 96 were assigned the same parent as both mother and father, giving a 24.24% selfing rate. Self-fertilization rates were relatively high in this study compared to some previous estimates of selfing in Douglas-fir seed orchards and wild populations (Slavov et al 2005, El-Kassaby and Ritland 1985), but relatively close to the selfing rates identified in other studies, including previous monitoring of this seed orchard (Lai et al. 2010; El-Kassaby et al. 2010, Fast et al. 1986). This variation in selfing between studies indicates that selfing rate is affected by variation in orchard management and design, as well as environmental and temporal factors affecting male



reproductive output, and individual genetic variation in male output and self-compatibility capabilities. Selfing was correlated with total gametic contribution ( $r = 0.70$ ), male gametic contribution ( $r = 0.69$ ), and female gametic contribution ( $r = 0.64$ ), indicating that imbalances in gametic contribution between parents are the likely cause of high selfing rate (Figure 4). Parents that produce excess pollen and receptive seed are more likely to produce self-fertilized seed, due to decreased competition between pollen from the same clone and pollen from other orchard parents (Erickson and Adams 1990).

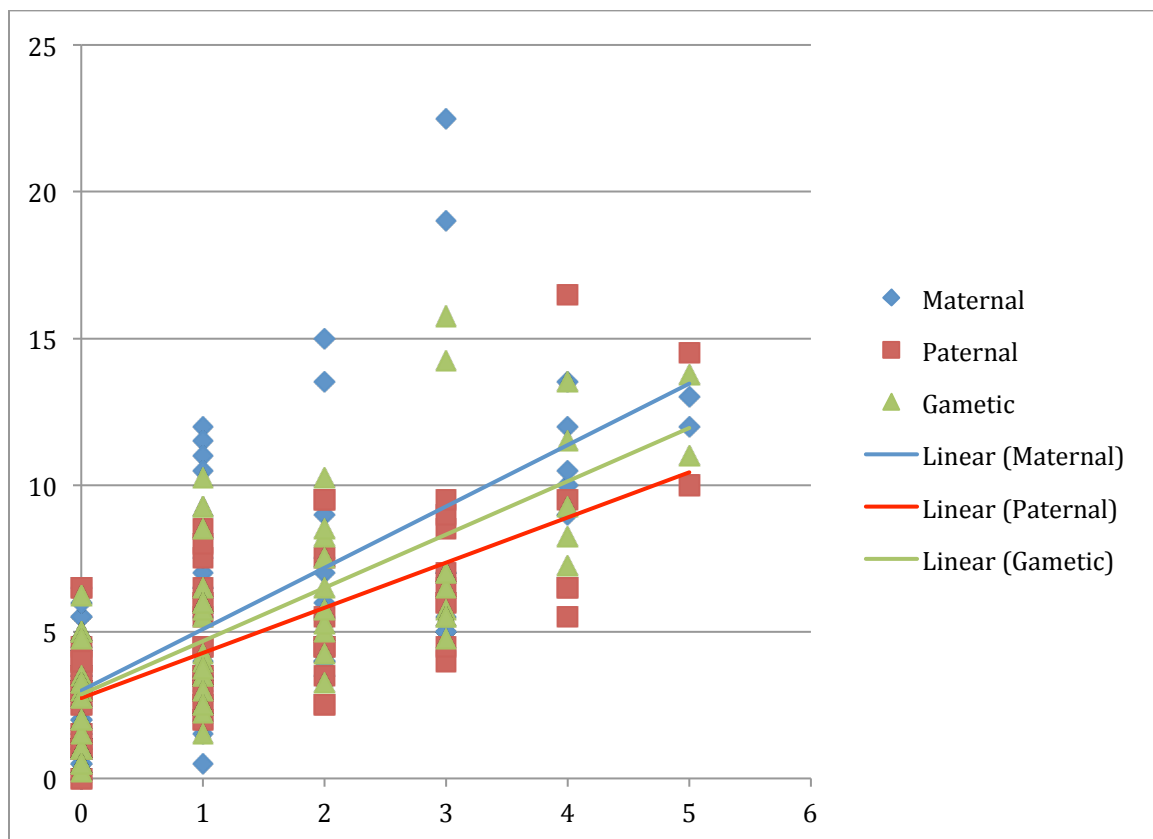


Figure 4. Correlation between amount of self seed produced and number of maternal, parental and paternal gametic contributions.

Another factor affecting the level of selfing is the reproductive synchrony of clones within the orchard. Previous studies have shown that bloom delay synchronizes pollen production and receptivity, and promotes outcrossing by ensuring all orchard parents are receptive to pollen. In orchards that have not

been cooled for bloom delay, the period of pollen production and receptivity is longer (Fashler and El-Kassaby 1987), and the orchard is reproductively separated into multiple mating subpopulations, each characterized by different periods of pollen production and receptivity. Outcrossing rates in orchards not cooled for bloom delay are lower, as there is less pollen from other parents competing with pollen from the same individual (Erickson and Adams 1990, El-Kassaby and Davidson 1991). In the orchard currently studied, selfing was high (15.3%) in a previous year when orchard parents were cooled for bloom delay, due to high gametic contribution of within-orchard parents (Lai et al. 2010; El-Kassaby et al 2010). In the present study, it is likely that the level of selfing has been compounded by the absence of bloom delay treatment and high contribution in some clones (Figure 1), creating subpopulations containing clones with high pollen production and fewer receptive parents available for outcrossing. Selfing also showed a moderate correlation with contamination ( $r = 0.462$ ). While the relationship between pollen contamination and selfing is unexpected, this correlation may be a result of the division of orchard parents into smaller breeding populations due to temporal reproductive asynchrony; small breeding groups will have a greater risk of both selfing and contamination due to the low number of within-orchard pollen donors, especially if maternal gametic contribution of parents in these groups is high. Reproductive asynchrony may create scenarios in which parents are more receptive to self and contaminant pollen than pollen produced later by other orchard parents. Because both selfing and contamination decrease the genetic gain achieved by breeding, bloom delay is recommended whenever possible.

### **Estimation of Supplemental Pollination Effectiveness**

Supplemental pollen donors also contributed ambient pollen in the orchard studied, making it impossible to directly differentiate between the success of supplemental pollination and the reproductive success of these orchard parents due to ambient pollen contribution. Significant differences at the 95% confidence level in reproductive success between the supplemental pollen donor population

and the population of parents donating only ambient pollen could be interpreted as successful supplemental pollination. However, there was only a small increase in the reproductive success of parents that contributed supplemental pollen. The average paternal gametic contributions for SMP parents was  $5.132 \pm 0.776$ , compared to an average paternal gametic contribution of  $4.913 \pm 0.922$  from parents donating only ambient pollen. This increase in mean paternal contribution was not statistically significant at the 95% confidence level when compared to the reproductive success of ambient pollen. The top 20 parents ranked for paternal gametic contribution contained only 4 SMP donors (Figure 3). There are several factors that could have caused SMP failure. Due to the high number of effective male pollen donors in this population (Figure 1), the pollen added via SMP may have been outcompeted by the ambient pollen present. In the study of the 2005 seed crop from this orchard, parents donating supplemental pollen had greater male reproductive equality, with 63% of parents donating 80% of pollen, compared to the differences of paternal gametic contribution in the orchard overall, with 45% of parents donating 80% of pollen (El-Kassaby et al. 2010). In contrast, the base level of male reproductive equality of the orchard in the present study was already high, with 63% of parents contributing 80% of pollen. Because one of the main objectives of SMP treatment is to equalize the reproductive output of orchard parents, supplemental pollen served as an extraneous treatment to an orchard with relatively high male contribution equality. Douglas-fir reproduction is characterized by “first on - first in” pollination, in which receptivity to the first arriving pollen granules is highest, and most likely to result in fertilization (Webber and Yeh 1987). In the orchard studied, ambient pollen may have successfully pollinated the majority of receptive females, and rendered supplemental pollen ineffective.

### **Determination of Minimum Fingerprinting Effort**

Previous studies estimating pollen contamination in seed orchards have not included a threshold level of genotyping past which the cost of genotyping provides diminishing returns in accurately estimating gene flow. The majority of

contamination studies have involved genotyping a seemingly arbitrary number of offspring, often far greater than would be required to predict gene flow accurately. While increased sample size will always increase statistical certainty of an estimate, there is considerable cost in laboratory reagents and time that can be minimized by determining minimum fingerprinting effort. Contamination estimates from simulated random samples identified a minimum genotyping effort of 150 seeds for low, moderate and high contamination scenarios (Table 2).

Table 2. Mean contamination estimate and confidence intervals of randomly sampled sets of genotyped seeds under low, medium and high contamination levels.

Contamination Level	# of genotyped seeds			
	100	150	200	250
6% (low)	8.70 ± 2.94%	6.60 ± 0.74%	6.20 ± 0.44%	6.20 ± 0.32%
14% (medium)	15.3 ± 1.76%	14.9 ± 0.84%	14.5 ± 0.81%	14.5 ± 0.79%
54% (high)	53.1 ± 1.96%	53.9 ± 1.29%	54.2 ± 0.92%	53.4 ± 0.63%

The size of the confidence interval decreased more quickly in low and medium contamination scenarios than in the high contamination scenario. Using the genotypes generated from the bulk seed sample in this study, it was found that while genotyping 100 seeds led to larger confidence interval than other sample sizes (Figure 5), genotyping of 150 seeds was sufficient to estimate contamination with a small confidence interval ( $14.9 \pm 0.84\%$ ), consistent with the minimum sample size identified using simulated genotypes. Genotyping of samples beyond this number served only to further increase confidence in the contamination estimate (350 seeds,  $14.0 \pm 0.27\%$ ) at the 95% confidence level. This low genotyping effort is comparable to the low genotyping effort required to estimate the threshold effective population required for genetic diversity capture (Yanchuk 2001; Lstiburek et al. 2011). The identified genotyping effort will allow for cost and time efficient contamination monitoring, and can be used to quickly

identify realized gene flow when pollen trap measures are inconsistent or unexpectedly high.

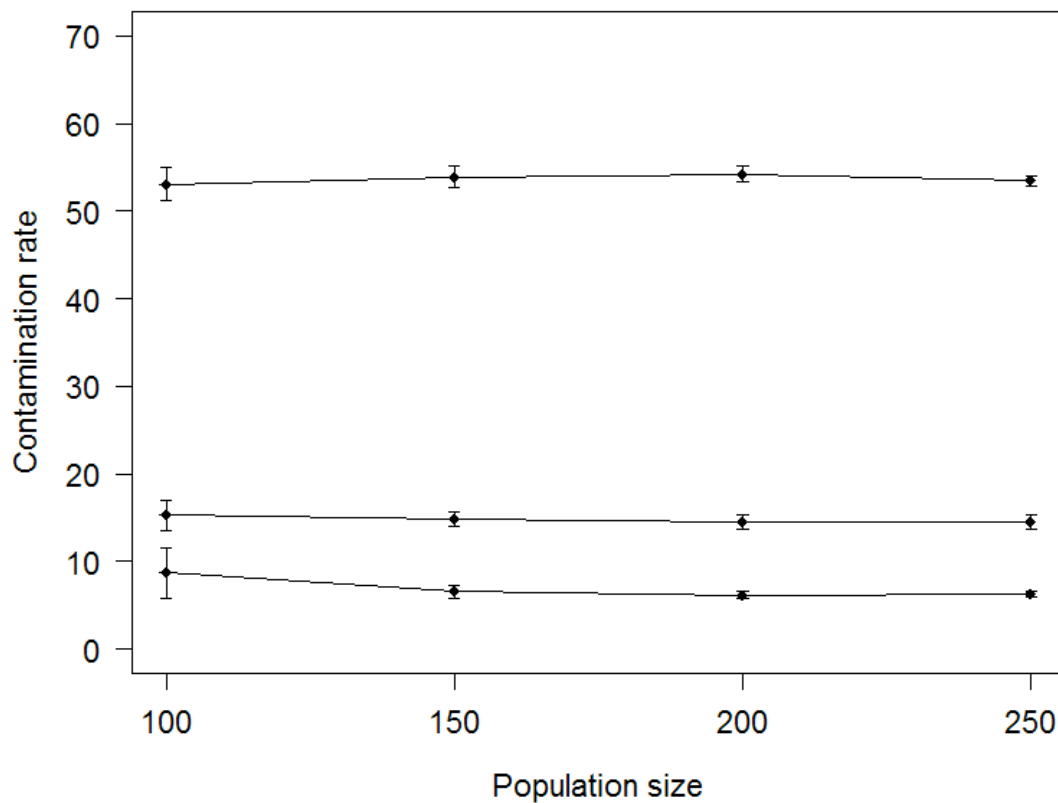


Figure 5. Contamination rate and confidence intervals of contamination estimates at different contamination rates detected from 30 randomly sampled seed genotype sets of different population sizes.

## CONCLUSION

Production of improved seed for reforestation relies on seed orchards functioning properly to package the genetic gain and diversity selected during breeding cycles. Deviations in reproductive synchrony and equality within the seed orchard will alter predicted genetic gain and diversity. In this study, I used highly variable microsatellite markers designed by Slavov et al (2004) to reconstruct the pedigree of the 2009 seed crop of a Douglas-fir seed orchard. Pedigree reconstruction of bulk seed sample allowed for identification of the reproductive dynamics within the seed orchard. Male, female, and parental gametic contribution and effective population size were all lower than expected under panmixia in this orchard. Pedigree reconstruction with molecular markers also allowed for estimation of selfing and contamination rate. The use of a large bulk seed sample and application of the jackknife random sampling method provided high confidence and low standard error in contamination estimates at the 95% confidence level. Jackknife random sampling also allowed for identification of a minimum seed sample size for contamination estimation in future studies with 95% confidence and low standard error. This study was conducted in a seed orchard previously studied by Lai et al. (2010) and El-Kassaby et al (2010), and so comparisons could be made between the reproductive dynamics between the 2005 crop and the 2009 crop used in this study. Differences in genetic diversity, selfing and contamination between years in this orchard were attributed to the absence of bloom delay and lower levels of reproductive imbalance in the seed crop of the current study. Pedigree reconstruction of future seed crops from the same seed orchard with or without bloom delay will allow for more accurate comparison of the effects of bloom delay on selfing and pollen contamination. Future monitoring of contamination and reproductive dynamics in this and other Douglas-fir seed orchards can use the decreased genotyping effort identified in this study to minimize the time, cost and effort of pedigree reconstruction using molecular markers.

## WORKS CITED

- Adams, WT. 1992. Gene dispersal within forest tree populations. *New Forests* 6:217-240.
- Adams, WT, Birkes, DS. 1989. Mating patterns in seed orchards. In: *Proceedings from the 20th Southern Forest Tree Improvement Conference*, Charleston, S.C.
- Adams, WT, Burczyk, J. 2000. Magnitude and implications of gene flow in gene conservation reserves. In: *Forest Conservation Genetics: Principles and practice*. Young, A, Boshier, D, Boyle, T. (eds.). CSIRO Publishing, Collingwood, Victoria, Australia. pp.215-224.
- Andrews, CA. 2010. Natural selection, genetic drift, and gene flow do not act in isolation in natural populations. *Nature Education Knowledge* 1.10:5
- Ashley, MV. 2010. Plant parentage, pollination and dispersal: How DNA microsatellites have altered the landscape. *Critical Reviews in Plant Sciences* 29.3
- Askew, GR. 1986. Implications of non-synchronous flowering in clonal conifer seed orchards. In: *Proc. IUFRO Conf. on Breeding Theory, Progeny Testing and Seed Orchards*. Oct. 13–17, 1986. Williamsburg, VA. pp.182-191.
- Bell, GD, and Fletcher, AM. 1978. Computer organized orchard layouts (COOL) based on the permuted neighbourhood design concept. *Silvae Genetica* 27:223-225.
- Botstein, D, White, RL, Skolnick, M., Davis, RW. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32(3):314-331.
- Bridgwater, FE, Blush, TD, Wheeler, NC. 1993. Supplemental mass pollination. In: *Advances in pollen management USDA For. Serv. Agric. Handbook* 698, pp. 69-77.
- Bridgwater, FE, Bramlett, DL, Matthews, FR. 1987. Supplemental mass pollination is feasible on an operational scale. In: *Proc. 19th South. For. Tree Improv. Conf.* June 16–18, 1987. College Sta, TX, pp.216-222.

- Burdon, RD, Shelbourne, CJA. 1971. Breeding populations for recurrent selection conflicts and possible solutions. *New Zealand Journal of Forestry Science* 1:174–93.
- Charlesworth, B, Charlesworth, D. 2010. *Elements of evolutionary genetics*. Roberts and Company Publishers. Greenwood Village, CO.
- Dow, BD, Ashley, MV. 1998. High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. *Heredity* 89:62-70.
- Doyle, JJ, Doyle, JL. 1987. A rapid DNA isolation procedure for small quantities of fresh tissue. *Phytochemical Bulletin* 19:11-15.
- Edwards, DGW, El-Kassaby, YA. 1995. Douglas-fir genotypic response to seed stratification. *Canadian Journal of Forest Research* 23:771-778.
- El-Kassaby, YA. 1989. Genetics of Douglas-fir seed orchards: expectations and realities. In: *Proceedings from the 20th Southern Forest Tree Improvement Conference*, Charleston, S.C.
- El-Kassaby, YA. 1995. Evaluation of the tree-improvement delivery system - factors affecting genetic potential. *Tree Physiology* 15:545–50
- El-Kassaby, YA, Barnes, S, Cook, C, MacLeod, DA. 1993. Supplemental mass pollination success rate in a mature Douglas-fir seed orchard. *Canadian Journal of Forest Research* 23:1069-1099
- El-Kassaby, YA, Cook, C. 1994. Female reproductive energy and reproductive success in a Douglas-fir seed orchard and its impact on genetic diversity. *Silvae Genetica* 43:243-246.
- El-Kassaby, YA, Davidson, R. 1991. Impact of pollination environment manipulation on the apparent outcrossing rate in a Douglas-fir seed orchard. *Heredity* 66:55-59.
- El-Kassaby, YA, Funda, T, Lai, BSK. 2010. Female Reproductive Success Variation in a *Pseudotsuga menziesii* Seed Orchard as Revealed by Pedigree Reconstruction from a Bulk Seed Collection. *Journal of Heredity* 101:164-168.



- El-Kassaby YA, Ritland, K. 1986a. Low levels of pollen contamination in a Douglas-fir seed orchard as detected by allozyme markers. *Silvae Genetica* 35:224-229.
- El-Kassaby, YA, Ritland, K. 1986b. The relation of outcrossing and contamination to reproductive phenology and supplemental mass pollination in a Douglas-fir seed orchard. *Silvae Genetica* 35:240-244.
- El-Kassaby, YA, Ritland, K. 1996. Impact of selection and breeding on the genetic diversity in Douglas-fir. *Biodiversity and Conservation*. 5:795-813.
- Erickson, VJ, Adams, WT. 1990. Mating system variation among individual ramets in a Douglas-fir seed orchard. *Canadian Journal of Forestry Research* 20:1248-1255.
- Eriksson, G, Clapham, D, Ekberg, I. 2006. An introduction to forest genetics. 2nd Edition. Swedish University of Agricultural Sciences, Uppsala, Sweden.
- Forest Genetics Council of B.C. Business Plan. 2009. [www.fgcouncil.bc.ca/FGC-BusinessPlan-2009-10.pdf](http://www.fgcouncil.bc.ca/FGC-BusinessPlan-2009-10.pdf) (Woods, JH., compiler and ed.).
- Fashler, AMK, El-Kassaby, YA. 1987. The effect of water spray cooling treatment on reproductive phenology in a Douglas-fir seed orchard. *Silvae Genetica* 36:245-249.
- Fast, W, Dancik, BP, Bower, RC. 1986. Mating system and pollen contamination in a Douglas-fir clone bank. *Canadian Journal of Forest Research* 16:1314-19.
- Fernandes, L, Rocheta, M, Cordeiro, J, Pereira, S, Gerber, S, Oliveira, M, Ribeiro, MM. 2008. Genetic variation, mating patterns and gene flow in a *Pinus pinaster* Aiton clonal seed orchard. *Annals of Forest Science* 65:706.
- Funda, T, El-Kassaby, YA. 2012. Seed orchard genetics. *CAB Reviews*. 7
- Hartl, DL, Clark, AG. 1997. Principles of population genetics. Sinauer Associates, Sunderland, MA.
- Jones, A, Clayton, M, Small, K, Paczolt, A, Ratterman, N 2009. A practical guide to methods of parentage analysis. *Molecular Ecology Resources* 10:6-30.

- Kalinowski, ST, Taper, ML, Marshall, TC. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* 16:1099-1006.
- Lai, BK, Funda, T, Liewlaksaneeyanwin, C, Klapstie, J, van Niejenhuis, A, Cook, C, Stoehr, MU, Woods, J, El-Kassaby, YA. 2010. Pollination dynamics in a Douglas-fir seed orchard as revealed by pedigree reconstruction. *Ann. For. Sci.* 67:808 DOI: 10.1051/forest/2010044
- Lindgren, D, Gea L, Jefferson, P. 1996. Loss of genetic diversity monitored by status number. *Silvae Genetica* 45:52-59.
- Lstiburek, M, Ivanková, K, Kadlec, J, Kobliha, J, Klápšte, J, El-Kassaby, YA. 2011. Breeding without breeding: minimum fingerprinting effort with respect to the effective population size. *Tree Genetics and Genomes* 7:1069-1078.
- Marshall, TC, Slate, J, Kruuk, LEB, Pemberton, JM. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* 7.5:639-655.
- Nakamura, RR, Wheeler, NC. 1992. Pollen competition and paternal success in Douglas-fir. *Evolution* 46:846-851.
- Namkoong G, Kang, HC, Brouard, JS. 1988. Tree breeding: Principles and strategies. Monographs on Theoretical and Applied Genetics 11. Springer-Verlag, New York.
- Nikkanen, T. 2002. Functioning of a Norway spruce (*Picea abies* (L.) Karst.) seed orchard. Finnish Forest Research Institute, Research papers 850. 58 pp. + 6 appendices.
- Owens, JN, Simpson, SJ, Molder, M. 1981. The pollination mechanism and the optimal time of pollination in Douglas-fir (*Pseudotsuga menziesii*). *Canadian Journal of Forest Research* 11:36-50.
- Pulkinen, P, Haapanen, M, Mikola, J. 1995. Effect of southern pollination on the survival and growth of seed orchard progenies of northern Scots pine (*Pinus sylvestris*) clones. *Forest Ecology and Management* 73:75-84.

- Ritland, K, El-Kassaby, YA. 1985. The nature of inbreeding in a seed orchard of Douglas-fir as shown by an efficient multilocus model. *Theoretical and Applied Genetics* 71:101:375-384.
- Silen RR, Keane, G. 1969. Cooling a Douglas-fir seed orchard to avoid pollen contamination, USDA For. Serv. Res. Note PNW-101.
- Slavov, GT, Howe, GT, Yakovlev, I, Edwards, KJ, Krutovskii, KV, Tuskan, GA, Carlson, JE, Strauss, SH, Adams, WT. 2004. Highly variable SSR markers in Douglas-fir: Mendelian inheritance and map locations. *Theoretical and Applied Genetics* 108:873-880.
- Slavov, GT, Howe, GT, Adams, WT. 2005. Pollen contamination and mating patterns in a Douglas-fir seed orchards as measured by simple sequence repeat markers. *Canadian Journal of Forest Research* 35:1592-1603.
- Smith, DB, Adams, WT. 1983. Measuring pollen contamination in clonal seed orchards with the aid of genetic markers. In: *Proceedings of the 17th Southern Forest Tree Improvement Conference*, p.64-73. Athens, GA, US.
- Stoehr, M, Mehl, H, Nicholson, G, Pieper, G, Newton, C. 2006. Evaluating supplemental mass pollination efficacy in a lodgepole pine orchard in British Columbia using chloroplast DNA markers. *New Forests* 31:83-90.
- Stoehr, M, Webber, J, Woods, J. 2004. Protocol for rating seed orchard seedlots in British Columbia: quantifying genetic gain and diversity. *Forestry* 77:297-303.
- Wakeley, PC, Wells, OO, Campbell, TE. 1966. Mass production of shortleaf-slash pine hybrids -by pollinating unbagged female flowers. In: *Joint Proceedings of the Second Genetics Workshop of the Society of American Foresters and the Seventh Lake States Forest Tree Improvement Conference*. U.S.D.A. Forest Service, St. Paul, USA. p. 78-89.
- Wang, J, Santure, A 2009. Parentage and sibship inference from multilocus genotype data under polygamy. *Genetics* 181:1-16.
- Webber, JE, Painter, RA. 1996. Douglas-fir pollen management manual. Second edition, Res. Program WP 02/96. B.C. Ministry of Forests, Victoria, Canada.

- Webber, JE, Yeh, FCH. 1987: Test of the *first-on, first-in* pollination hypothesis in coastal Douglas-fir. Canadian Journal of Forest Research 17:63-68.
- Wheeler, N, Jech, K. 1986. Pollen contamination in a mature Douglas-fir seed orchard. In: Proceedings from IUFRO Conference on Breeding Theory, Progeny Testing and Seed Orchards, Williamsburg, VA; 1986, pp.160-171.
- Woods, JH, Stoehr, MU, Webber, JE. 1996. Protocols for rating seed orchard seedlots in British Columbia. B.C. Ministry of Forests, Forestry Division Services Branch, Research Report 06. p. 26.
- Worrall, J. 1983. Temperature-bud-burst relationships in amabilis and subalpine fir provenance tests replicated at different elevations. Silvae Genetica 32:203-209.
- Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97-159.
- Yanchuk, AD. 2001. A quantitative framework for breeding and conservation of forest tree genetic resources in British Columbia. Canadian Journal of Forest Research 31:566-576.
- Yazdani, R, Hadders, G, Szmidt, A. 1986. Supplemental mass pollination in a seed orchard of *Pinus sylvestris* L. investigated by isozyme analyses. Scandinavian Journal of Forest Research 1:309-315.
- Zobel, BJ. 1971. The Genetic Improvement of Southern Pines. Scientific American 225.55:94-103.
- Zobel, BJ. 1976. Gene Conservation as viewed by a forest tree breeder. Forest Ecology and Management 1:339-344.
- Zobel BJ, Talbert, JT. 1984. Applied Forest Tree Improvement. John Wiley and Sons, New York.