

POPULATION GENETICS OF CONIFER SEED ORCHARDS

by

TOMAS FUNDA

Ing., Czech University of Life Sciences Prague, 2003

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

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Forestry)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

March 2012

© Tomas Funda, 2012

Abstract

Seed orchards represent the link between breeding and silvicultural activities. They were expected to act as closed, panmictic populations in Hardy-Weinberg equilibrium, meaning that desired genes drafted during previous selection stage would be effectively transmitted from parental to offspring populations; however, extensive research has indicated that this expectation is not met. Scrutinizing seed orchards' efficiency is of vital importance as it determines the genetic quality of future forest stands.

Population genetics of four tree species' seed orchards (western larch, Douglas-fir, lodgepole pine, and western redcedar) was studied using microsatellite DNA markers. Partial (family array) and full (bulk seed) pedigree reconstruction of offspring population (seed crops) were conducted using the likelihood-based parentage inference program CERVUS to estimate parental reproductive success, selfing rate, pollen contamination, effective number of parents (N_e), and seedlot genetic worth. Several simplified methods for predicting seed crops' genetic quality and quantity were evaluated by comparing parental reproductive success with parental fecundities.

In all species, the top 20% of males contributed approximately one half of successful within-orchard pollen, substantially reducing male N_e (45 to 62% of the orchards' census numbers). Even larger distortion was observed among females (the top 20% of females produced 77% of seed crop in Douglas-fir), reducing female N_e to as little as 13% of the census. Selfing and pollen contamination rates were in the range of previously reported studies, with the exception of high (15.2%) and low (7.3%) selfing rates in Douglas-fir and western redcedar, respectively. Pollen bud production and seed-cone volume were found to be the most reliable proxies to parental reproductive success, genetic worth, and N_e estimates.

An optimization protocol was developed for creating custom seedlots with maximized genetic gain at any N_e while collectively considering parental male and female fecundities, co-ancestry among parents, inbreeding, and variation in seed germination capacity. This protocol can be utilized in any generation's seed orchard e.g. when seed supply exceeds demand, for mixing surpluses from multiple years, or if a given seed lot fails to meet minimum N_e requirements.

Preface

Chapter 1 is a version of an invited review paper: Funda T., El-Kassaby Y.A. Seed Orchard Genetics. CAB Reviews (In press). T. Funda wrote the chapter and Y.A. El-Kassaby revised it.

Chapter 2 is a revised version of a published paper: Funda T., Chen C.C., Liewlaksaneeyanawin C., Kenawy A.M.A., El-Kassaby Y.A. 2008. Pedigree and mating system analyses in a western larch (*Larix occidentalis* Nutt.) experimental population. *Annals of Forest Science* 65(7): 705. DOI: 10.1051/forest:2008055. T. Funda and A.K.A. Kenawy did the lab work (DNA extraction, PCR, and PAA gels); T. Funda, C. Liewlaksaneeyanawin, and C.C. Chen analyzed the data; T. Funda wrote the chapter with contribution of C.C. Chen and C. Liewlaksaneeyanawin; and Y.A. El-Kassaby revised it. Seed collection was provided by C. Walsh (Kalamalka Research Station of the B.C. Ministry of Forests, Mines and Lands). The chapter is reproduced with kind permission of EDP Sciences (<http://www.afs-journal.org/>).

For **Chapter 3**, T. Funda collected parental foliar tissue, ran the labwork (DNA extraction, PCR, and PAA gels), analyzed the data, and wrote the chapter; Y.A. El-Kassaby revised it. Seed collection was provided by A. Van Niejenhuis (Western Forest Products Ltd.).

Chapter 4 is a revised version of a published paper: El-Kassaby, Y.A., Funda, T., and Lai, B.S.K. 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(2): 164-168. DOI: 10.1093/jhered/esp126. T. Funda analyzed data provided by B.S.K. Lai and Y.A. El-Kassaby and T. Funda wrote the manuscript. Seed collection and parental foliage was provided by A. Van Niejenhuis (Western Forest Products Ltd.). The chapter is reproduced with kind permission of Oxford University Press (<http://jhered.oxfordjournals.org/>).

For **Chapter 5**, T. Funda, I. Fundova, C. Na Takuathung, and C. Liewlaksaneeyanawin did the lab work (DNA extraction, PCR, and PAA gels); T. Funda and C. Liewlaksaneeyanawin analyzed the data; T. Funda wrote the chapter with contribution of C. Liewlaksaneeyanawin;

and Y.A. El-Kassaby revised it. Seed collection and parental foliage was provided by H. Graham (Pacific Regeneration Technologies Inc.).

Chapter 6 is a revised version of a published paper: Funda T., Liewlaksaneeyanawin C., Fundova I., Lai B.S.K., Walsh C., Van Niejenhuis A., Cook C., Graham H., Woods J., El-Kassaby, Y.A. 2011. Congruence between parental reproductive investment and success determined by DNA-based pedigree reconstruction in conifer seed orchards. *Canadian Journal of Forest Research* 41(2): 380-389. DOI: 10.1139/X10-190. This paper is based on genetic data obtained from studies included in Chapters 2–5 (see the previous four paragraphs for the relative contribution of each co-author regarding the genetic analyses) and from El-Kassaby et al. (2011). T. Funda analyzed the data with contribution of I. Fundova and wrote the chapter; Y.A. El-Kassaby revised it. C. Walsh, A. Van Niejenhuis, H. Graham, and C. Cook provided parental fecundity data from western larch, Douglas-fir, lodgepole pine, and western redcedar seed orchards, respectively. J. Woods had developed the evaluated simplified methods for predicting parental reproductive success. The chapter is reproduced with kind permission of Canadian Science Publishing/NRC Research Press.

Chapter 7 is a revised version of a published paper: Funda T., Lstibůrek M., Lachout P., Klápště J., El-Kassaby Y.A. 2009. Optimization of combined genetic gain and diversity for collection and deployment of seed orchard crops. *Tree Genetics & Genomes* 5(4): 583-593. DOI: 10.1007/s11295-009-0211-3. T. Funda, M. Lstibůrek, and J. Klápště developed the mathematical model following previous work by Lindgren and Mullin (1998); T. Funda analyzed the data and wrote the chapter and M. Lstibůrek and Y.A. El-Kassaby revised it. M. Lstibůrek and J. Klápště wrote the code in C++ and computed optimum solutions and P. Lachout provided a mathematical proof of the positive semidefiniteness of the co-ancestry matrix (not included in this thesis). Parental fecundity data was provided by C. Walsh (Kalamalka Research Station of the B.C. Ministry of Forests, Mines and Lands). The chapter is reproduced with kind permission of Springer Science and Business Media (<http://www.springerlink.com>).

Chapter 8 is a version of a submitted manuscript: Funda T. Lstibůrek M., Klápště J., El-Kassaby Y.A. Optimization of genetic gain and diversity in seed orchard crops considering

germination variation. Scandinavian Journal of Forest Research (In press). T. Funda modified the original optimization protocol (Chapter 7) to include variation in seed germination capacity, analyzed the data and wrote the chapter; M. Lstibůrek and J. Klápště wrote the code in C++ and computed optimum solutions; Y.A. El-Kassaby revised the chapter. L.M. El-Kassaby and T. Funda collected seed germination data.

Appendix A is a version of a table extracted from a published paper: Chen C.C., Liewlaksaneeyanawin C., Funda T., Kenawy A.M.A., Newton C.H., El-Kassaby Y.A. 2009. Development and characterization of microsatellite loci in western larch (*Larix occidentalis* Nutt.). Molecular Ecology Resources 9(3): 843-845. DOI: 10.1111/j.1755-0998.2008.02289.x. C.C. Chen supervised lab work conducted by T. Funda and A.K.A. Kenawy (DNA extraction, PCR, and PAA gels), analyzed the data with contribution of C. Liewlaksaneeyanawin and T. Funda and wrote the paper. C.H. Newton developed primer sequences for the tested loci. Y.A. El-Kassaby revised the paper. The table is reproduced with kind permission of John Wiley & Sons Ltd. (<http://www.wiley.com>).

Table of Contents

Abstract.....	ii
Preface.....	iii
Table of Contents	vi
List of Tables	xi
List of Figures.....	xii
List of Tree Species	xv
List of Abbreviations	xvi
Acknowledgements	xvii
Dedication	xix
1. Literature Review	1
1.1 Forest Tree Improvement.....	1
1.2 Population Description.....	2
1.3 Seed Orchards	4
1.4 Genetic Quality	6
1.4.1 Genetic gain	7
1.4.2 Genetic diversity	8
1.5 Factors Affecting Genetic Quality in Seed Orchard Crops	12
1.5.1 Reproductive success	13
1.5.2 Reproductive phenology	16
1.5.3 Inbreeding	17
1.5.4 Gene flow	19
1.5.5 Seed germination	21
1.6 Methods to Manipulate Genetic Quality.....	22
1.6.1 Population size	22
1.6.2 Parental representation.....	23
1.6.3 Orchard design	24
1.6.4 Crown management	27
1.6.5 Supplemental mass pollination	28

1.6.6	Bloom delay	30
1.6.7	Selective harvesting and production of custom seedlots	30
1.7	Genetic Markers	31
1.8	Parentage Analysis	33
1.9	Objectives	35
2.	Pedigree and Mating System Analyses in a Western Larch (<i>Larix occidentalis</i> Nutt.)	
	Seed Orchard.....	36
2.1	Introduction.....	36
2.2	Materials and Methods.....	37
2.2.1	Seed orchard, plant material, DNA extraction and fingerprinting	37
2.2.2	Mode of inheritance of microsatellite loci	39
2.2.3	Paternity and mating system analyses.....	39
2.2.4	Assessment of genetic gain and diversity	40
2.3	Results and Discussion	40
2.3.1	Microsatellite loci mode of inheritance	40
2.3.2	Mating system analysis	42
2.3.3	Paternity analysis	42
2.4	Conclusion	46
3.	Mating System Analyses in a Western Redcedar (<i>Thuja plicata</i> Donn. ex D. Don)	
	Seed Orchard.....	47
3.1	Introduction.....	47
3.2	Materials and Methods.....	49
3.2.1	Study seed orchard	49
3.2.2	DNA extraction, PCR and genotyping.....	50
3.2.3	Paternity analysis	50
3.3	Results and Discussion	51
3.3.1	Mislabeling and alien genotypes.....	51
3.3.2	Paternity analysis	52
3.3.3	Male reproductive success variation.....	52
3.3.4	Selfing	53
3.4	Conclusion	54

4. Female Reproductive Success Variation in a Douglas-fir (<i>Pseudotsuga menziesii</i> (Mirb.) Franco) Seed Orchard Revealed by Pedigree Reconstruction of a Bulk Seed Collection	55
4.1 Introduction.....	55
4.2 Materials and Methods.....	56
4.2.1 Seed orchard population and seed sampling.....	56
4.2.2 Microsatellite genotyping	56
4.2.3 Data analysis	57
4.3 Results and Discussion	57
4.4 Conclusion	61
5. Partial Versus Full Pedigree Reconstruction in a Lodgepole Pine (<i>Pinus contorta</i> Dougl. ex Loud.) Seed Orchard	62
5.1 Introduction.....	62
5.2 Materials and Methods.....	64
5.2.1 Seed orchard population and plant material sampling.....	64
5.2.2 DNA extraction and SSR genotyping	64
5.2.3 Mendelian inheritance tests.....	65
5.2.4 Parentage analyses	65
5.2.5 Reproductive success and genetic parameters estimates	66
5.3 Results and Discussion	67
5.3.1 Mendelian inheritance.....	67
5.3.2 Genetic diversity analysis	68
5.3.3 Verification of parental affiliation	69
5.3.4 Parentage analysis and gene flow	69
5.3.5 Pedigree reconstruction and parental reproductive success.....	71
5.3.6 Selfing.....	74
5.3.7 Effective number of parents.....	75
5.4 Conclusion	75
6. Congruence Between Parental Reproductive Investment and Success Determined by Microsatellite DNA-based Pedigree Reconstruction in Conifer Seed Orchards	77
6.1 Introduction.....	77

6.2	Materials and Methods.....	78
6.2.1	Seed orchard populations.....	78
6.2.2	Reproductive investment	79
6.2.2.1	Female reproductive investment	79
6.2.2.2	Male reproductive investment.....	80
6.2.2.3	Parental representation	80
6.2.3	Reproductive success	80
6.2.4	Survey methods evaluation.....	81
6.2.5	Genetic worth and diversity	82
6.3	Results and Discussion	83
6.3.1	Survey methods evaluation.....	83
6.3.1.1	Female component	83
6.3.1.2	Male component.....	85
6.3.2	Genetic diversity	86
6.3.3	Genetic worth.....	88
6.4	Conclusion	89
7.	Optimization of Combined Genetic Gain and Diversity for Collection and	
	Deployment of Seed Orchard Crops	93
7.1	Introduction.....	93
7.2	Materials and Methods.....	95
7.2.1	Seed orchard and fecundity assessment.....	95
7.2.2	Mathematical model.....	96
7.2.3	Model constraints.....	97
7.2.4	Limits to gametic contributions	98
7.2.5	Optimization	98
7.3	Results.....	100
7.3.1	Correlation	100
7.3.2	Genetic gain and diversity estimates.....	101
7.3.3	Optimization	101
7.4	Discussion	104
7.5	Conclusion	108

8. Optimization of Genetic Gain and Diversity in Seed Orchard Crops Considering Variation in Seed Germination.....	109
8.1 Introduction.....	109
8.2 Materials and Methods.....	110
8.2.1 Seed orchard and parental fecundity assessment	110
8.2.2 Seed germination	110
8.2.3 Optimization	110
8.3 Results and Discussion	111
8.3.1 Seed germination	111
8.3.2 Genetic gain and diversity	112
8.3.3 Optimization	113
8.4 Conclusion	115
9. Conclusion	116
9.1 Thesis Summary.....	116
9.2 Research Limitations	121
9.3 Future Directions	122
Bibliography	124
Appendix A Characterization of microsatellite loci for western larch (<i>Larix occidentalis</i> Nutt.) (after Chen et al. 2009).	165

List of Tables

Table 2.1 Mode of inheritance for western larch microsatellite loci using wind-pollinated seed from 14 parents. <i>G</i> -test for the goodness-of-fit to the 1:1 segregation ratio for pooled data (G_P) and test of heterogeneity among multiple parents sharing the same heterozygous genotype (G_H).	38
Table 3.1 Characterization of parental and offspring populations in a western redcedar seed sample using eight nuclear microsatellite loci.	53
Table 5.1 Log-likelihood <i>G</i> -test on segregation ratio of nine microsatellite loci for the same heterozygous genotype of two parents from a lodgepole pine seed orchard.	68
Table 5.2 Number of alleles (A), observed (H_o) and expected (H_e) heterozygosities, frequency of null alleles (f_n), and genotyping error rate (e) for nine microsatellite loci based on genetic analysis of 74 parents in a lodgepole pine experimental population.	69
Table 5.3 Summary of parentage analysis of 11 maternal half-sib families (known female parentage) and bulk seed (unknown male and female parentage) at 95% confidence levels.	70
Table 6.1 Fit statistics (R^2 : coefficient of determination and <i>RMSE</i> : root mean square error) for linear regressions between female (F) and male (M) reproductive investment assessment methods and reproductive success as determined by microsatellite-DNA-based pedigree reconstruction and crops' genetic parameters (N_e : status effective number; <i>GW</i> : genetic worth; <i>m</i> : male; <i>f</i> : female; <i>p</i> : joint male and female (parental)) for Douglas-fir, lodgepole pine, western larch, and western redcedar seed orchards.	90
Table 7.1 List of variables used in the optimization protocol.	100
Table 7.2 Pearson's correlation coefficients between female (below diagonal), male (above diagonal), and between male and female reproductive investment (diagonal).	100
Table 8.1 List of variables used in the optimization protocol.	111
Table 8.2 Germination capacity variation among 37 western larch parents.	112

List of Figures

Figure 1.1 Depiction of a tree improvement cycle after White (1987) and Ruotsalainen (2002). Symbols S, T and X represent artificial selection, progeny testing, and crossing among individuals, respectively.....	2
Figure 2.1 Western larch male half-sib family size listed in descending order.	44
Figure 2.2 The distribution of 221 full-sib families from 14 wind-pollinated western larch seed donors.....	45
Figure 3.1 Verification of parental affiliation of ramets in the western redcedar seed orchard using microsatellite DNA markers. Genotypic profiles are shown for 5 ramets of 12 parents (total 60 ramets) on loci TP3, TP4, TP11, TP 7, TP9, and TP5 (top to bottom) and allelic standards surround each group of 3 parents (15 ramets). One incident of clone contamination (C) and mislabeling error (M) are indicated above.	51
Figure 3.2 Cumulative reproductive success of 27 western redcedar male parents determined from microsatellite-DNA-based partial pedigree reconstruction. The diagonal line represents a hypothetical scenario with equal success among parents.	53
Figure 4.1 Female and male cumulative reproductive success in a Douglas-fir seed orchard showing a greater female than male gamete contribution distortion.	58
Figure 4.2 Pearson's correlation between female and male reproductive success for 49 parents in a Douglas-fir seed orchard.	58
Figure 4.3 Pearson's correlations between female (a), male (b), and combined male and female (parental; c) reproductive success and selfing rate within a Douglas-fir clonal seed orchard.	60
Figure 5.1 Distribution of naturally occurred crosses (mating design) in a lodgepole pine seed orchard revealed by partial pedigree reconstruction of 11 wind-pollinated maternal half-sib families using nine nuclear microsatellite loci.	72

Figure 5.2 Distribution of naturally occurred crosses (mating design) in a lodgepole pine seed orchard revealed by full pedigree reconstruction of offspring with unknown male and female parentage using a combination of nine nuclear and six chloroplast microsatellite loci.....	72
Figure 5.3 Comparison of paternal half-sib family sizes obtained from partial (family array) and full (bulk seed) pedigree reconstruction of offspring from a lodgepole pine seed orchard.	73
Figure 5.4 Comparison between cumulative male reproductive success from partial (family array) and full (bulk seed) pedigree reconstruction (left) and between male, female, and combined (parental) reproductive success from full pedigree reconstruction (right) of offspring from a lodgepole pine seed orchard. The diagonals represent equal contribution among parents.....	74
Figure 6.1 Linear regressions between female reproductive investment (survey methods F0–F3; see Section 6.2.2.1 for details) and transformed female reproductive success (microsatellite-DNA-based pedigree reconstruction) for lodgepole pine (a) and Douglas-fir (b). The three curves represent the linear regression line (solid) and 95% confidence limits for individual predictions (dashed).	91
Figure 6.2 Linear regressions between male reproductive investment (survey methods M0–M3; see Section 6.2.2.2 for details) and transformed male reproductive success (microsatellite-DNA-based pedigree reconstruction) for lodgepole pine (a), Douglas-fir (b), western larch (c), and western red cedar (d). The three curves represent the linear regression line (solid) and 95% confidence limits for individual predictions (dashed).	92
Figure 7.1 Estimated genetic gain (left axes, solid markers) and corresponding effective status number (right axes, open markers) in the western larch seed orchard over three years (2004–2006) after a stepwise exclusion of lower-breeding-value seed constrained by $N_{e\ min} = 10$. The three pairs of curves in each year represent results from the three different scenarios applied (A- male and female reproductive equality, B- male as a function of parental representation, and C- actual data).	102

Figure 7.2 Estimated genetic gain (left axes, solid markers) and corresponding effective status number (right axes, open markers) in a western larch seed orchard over three years (2004–2006) after a stepwise exclusion of lower-breeding-value seed considering breeding values and relatedness following Scenario C and constrained by $N_{e\ min} = 10$ (top) and $N_{e\ min} = 15$ (bottom). The three pairs of curves represent no, moderate, and strong relatedness among parents as specified in Section 7.2.5 (hypothetical scenario). 104

Figure 8.1 Germination curves for 37 western larch parents fitted using the four-parametric Hill function (El-Kassaby et al. 2008). 112

Figure 8.2 Genetic gain (ΔG ; left axes, solid markers) and effective status number (N_e ; right axes, open markers) estimates from a western larch seed orchard crop after a stepwise (10%) exclusion of seed. The two marker shapes, square and triangle, represent $N_{e\ min}$ constraints of 15 and 20, respectively, and the point and line dashing represent the posterior adjustment of the genetic parameters due to variation in germination capacity following Scenario 1 and the inclusion of the germination capacity in the optimization model following Scenario 2. 114

Figure 8.3 Genetic gain (ΔG ; left axes, solid markers) and effective status number (N_e ; right axes, open markers) estimates from a western larch seed orchard crop after a stepwise (10%) exclusion of seed. The three marker shapes, square, triangle, and circle, stand for simulated mean germination capacities of 10, 20, and 30%, respectively, and N_e is constrained by the value of 20 for illustration. The left and right parts of the figure are based on Scenarios 1 and 2, respectively. 115

List of Tree Species

black pine (*Pinus nigra* Arnold)

Chinese arborvitae (*Platycladus orientalis* Franco), formerly oriental-cedar (*Thuja orientalis* L.)

Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco)

Engelmann spruce (*Picea engelmannii* Engel.)

Japanese black pine (*Pinus thunbergii* Parl.)

Japanese red pine (*Pinus densiflora* Sieb. et Zucc.)

knobcone pine (*Pinus attenuata* Lemmon)

loblolly pine (*Pinus taeda* L.)

lodgepole pine (*Pinus contorta* Dougl. ex Loud.)

Nordmann fir (*Abies nordmanniana* (Steven) Spach)

Norway spruce (*Picea abies* (L.) Karst.)

radiata pine (*Pinus radiata* D. Don)

Sitka spruce (*Picea sitchensis* (Bong.) Carrière)

slash pine (*Pinus elliottii* Engelm.)

subalpine fir (*Abies lasiocarpa* (Hook) Nutt.)

sugi (*Cryptomeria japonica* (L.f.) D. Don))

western larch (*Larix occidentalis* Nutt.)

western redcedar (*Thuja plicata* Donn. ex D. Don)

white spruce (*Picea glauca* (Moench) Voss)

wild cherry (*Prunus avium* L. (Batsch))

List of Abbreviations

CTAB	Cetyltrimethylammonium bromide
DNA	Deoxyribonucleic acid
IBD	Identity/identical by descent
PAA	Polyacrylamide
PCR	Polymerase chain reaction
SMP	Supplemental mass pollination
SNP	Single nucleotide polymorphism
SSR	Simple sequence repeats (a.k.a. microsatellites)
VNTR	Variable number of tandem repeats

Acknowledgements

I wish to thank everyone who contributed to this dissertation and helped me succeed. Foremost I wish to thank my research supervisor Yousry El-Kassaby for the tremendous support and encouragement that he has provided throughout my study. I am grateful for the opportunity to be part of his research group and to work on a number of interesting projects. Despite his busy schedule filled with both academic and research responsibilities, Yousry has always had time to discuss my ideas with me and has been a great guide through the field of forest genetics. I wish to thank to my committee member, Milan Lstibůrek who was especially supportive during my early years at UBC and who helped me develop the optimization protocol used in this thesis. I am grateful for the time he has dedicated to our discussions as well as for putting the ideas into a C++ code. Without his priceless contribution the optimization protocol would never have been what it is now. I also thank my committee members Michael Stoehr and Kim Cheng for helping me see things “from the other side” and for providing feedback on my progress.

It was my great pleasure to attend courses taught by Sally Otto, Mike Whitlock, Valerie LeMay, Kermit Ritland, Andrew Riseman and Yousry who have all provided me with a strong background for my research.

I am grateful to Charles Chen and Cherdsak Liewlaksaneeyanawin for their guidance through all steps of microsatellite DNA genotyping and for unraveling the mysteries and tricks of parentage analyses to me. I have had the pleasure of sharing the lab with many students and research assistants whose presence during my everlasting DNA extractions, PCRs, and gel processing made the lab a warm and nice place: Ben Lai, Irena Fundová, Simren Brar, Stéphanie Beauseigle, Lina Farfan, Manal Fayed, Chakrit Na Takuathung, Tony Kess, Mohamed Ismail, and Ahmed Kenawy. My special thanks go to Ben who has been a great friend and who always made sure I had all lab supplies needed for my work. I thank Jaroslav Klápště for testing countless versions of the optimization code and for “executing my orders” to generate data for my optimization analyses. I would also like to thank Ilga Porth and Ian MacLachlan for their friendship and help with this dissertation.

I wish to thank Annette Van Niejenhuis and Cathy Cook from Western Forest Products Ltd., Hilary Graham from Pacific Regeneration Technologies Inc., and Chris Walsh from the B.C.

Ministry of Forests, Mines and Lands Kalamalka Research Station for kindly providing plant material as well as parental fecundity data from their seed orchards.

I also wish to thank Seppo Ruotsalainen, Kyu-Suk Kang and Finnvid Prescher for the inspiration I have drawn from their wonderful dissertations, and to thank their supervisor Dag Lindgren, an inexhaustible well of knowledge and ideas, whom I studied under in Umeå, Sweden during my master's exchange program.

I would also like to acknowledge Cindy Prescott and Gayle Kosh from the Graduate Student Office for their moral support and formal guidance during my studies as well as Rosemarie Cheng and Andrea Chan from the Forest Sciences Department office for their assistance with lots of formal paperwork.

Most of all, I would like to thank my wife Irena for her love, support, and help during my PhD program and my parents for always believing in my ability to succeed.

Studies included in this thesis were financially supported by the Johnson's Family Forest Biotechnology Endowment, Forest Genetic Council of British Columbia, Natural Sciences and Engineering Research Council of Canada, and University of British Columbia. Their contribution is much appreciated.

To my family

1. Literature Review

1.1 Forest Tree Improvement

Forest tree improvement is a combination of breeding and silvicultural activities that are commonly and recurrently applied for exploiting the existing natural genetic variation in forest tree populations to increase the economic value of artificially regenerated forests and their products (Zobel and Talbert 1984). Unlike breeding and domestication of agricultural crops and animals (at least in their unstructured forms) which have been occurring on Earth for several millennia, the modern deliberate and organized tree improvement as we know it today did not start until as late as 1950s (van Buijtenen 1984). This significant delay had two major consequences: firstly, it left most forest tree populations largely untouched by genetic manipulation and, secondly, it allowed tree breeders to utilize the wealth of knowledge and draw on the experience accumulated over centuries in other fields (Namkoong et al. 1988). At its inception, tree improvement activities were mostly launched with the aim to immediately satisfy demand for high quality timber products, without any planning for future generations. Achieving reasonable gains quickly and effectively was possible when a very intensive selection was applied (i.e., selecting only a small fraction of the best individuals); however, this short-term thinking was eventually abandoned because it would leave too shallow a genetic base for future improvement, especially when the short-term programs were being converted into long-term endeavors (Burdon and Shelbourne 1971).

Most long-term tree-improvement programs can be described using the framework of a recurrent selection breeding cycle (White 1987) that comprises of two independent, but not mutually exclusive, breeding and production phases (Zobel and Talbert 1984). These phases are tightly linked and their main objective is to obtain and retain a broad genetic base while combining the desired characteristics into suitable high-quality, well-adapted trees that would be valuable for future generations (Zobel and Talbert 1984). The former phase consists of **selection** of the most desirable individuals either in natural forest stands, artificial plantations, or advanced generations of improvement according to specific criteria that would secure meeting the defined breeding goals, **breeding** which involves the production of controlled crosses among these selected individuals to permit the extraction of the genetic parameters needed to execute the second round of breeding, and **testing** to evaluate the

genetic quality of the selected trees via testing their progenies or the progenies themselves (i.e., backward and forward selection). In the latter phase, breeding efforts are “packaged” into seed crops and delivered to produce new, genetically improved forest stands or plantations.

1.2 Population Description

Tree improvement programs contain a variety of populations. According to the conceptual framework of White (1987) and later modifications by Ruosalainen (2002), these populations are defined by their function rather than their physical appearance, genetic constitution, or location (in fact, some functionally different populations may be physically overlapping (Ruotsalainen 2002)). One of the concept’s most distinct features is that populations used in the breeding cycle are separated from populations used to produce genetically improved material for operational plantations (e.g. Kang 1982; van Buijtenen 1984) as illustrated in Figure 1.1.

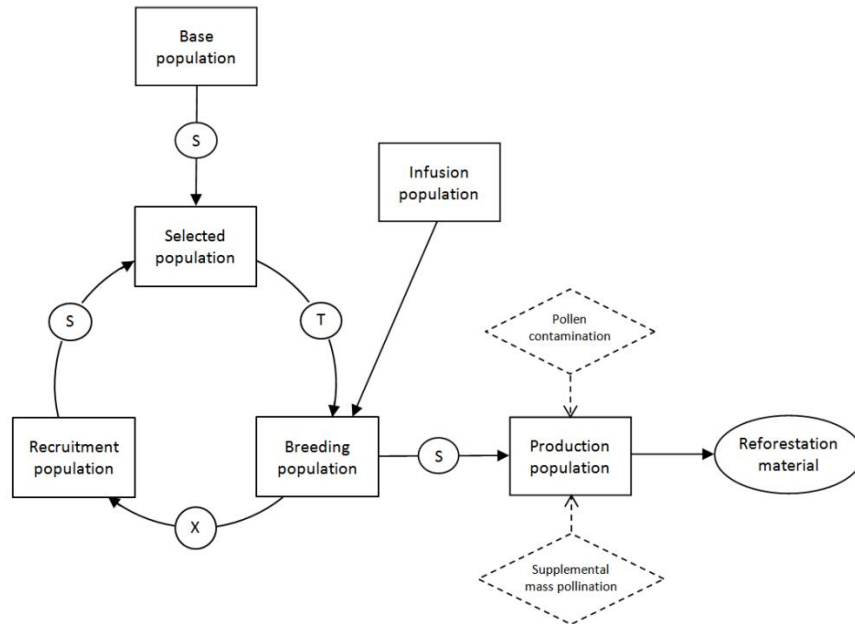


Figure 1.1 Depiction of a tree improvement cycle after White (1987) and Ruotsalainen (2002). Symbols S, T and X represent artificial selection, progeny testing, and crossing among individuals, respectively.

A population from which individuals are selected to enter a breeding program is called a base (Zobel and Talbert 1984) or synonymously a source (Kang and Nienstaedt 1987) population. Since both of these terms have been used interchangeably (for instance, the former was used for any starting material of a breeding program irrespective of its genetic background (Stonecypher 1969), Ruotsalainen (2002) suggested distinguishing between different generations of breeding by using the terms founder and recruitment population. The founder population consists of individuals selected in wild forests with no previous breeding history, usually using individual-tree selection methods (Gullberg and Kang 1985), whereas the recruitment population is functionally identical, but originates from advanced generations of breeding. Therefore, founder populations can be viewed as special, first-generation recruitment populations containing individuals that are assumed to be unrelated and non-inbred. Similarly, a gene resource population is defined either in the form of natural stands or clonal archives, but its main purpose is preserving the existing genetic legacy, i.e., gene conservation, rather than being considered for a given improvement program (Eriksson et al. 2006).

In each generation, a candidate (Ruotsalainen and Lindgren 2000) or selected population (White 1987) is established from a selected subset of phenotypically superior individuals drawn from the base or recruitment populations. This population usually has a census number several orders of magnitude lower than that of the base population, but its genetic composition is expected to carry sufficient potential to generate gains for the desirable trait(s) in the next generation. The actual genetic superiority of these selected trees is evaluated based on the performance of their offspring in progeny tests (Allard 1960), which represent an essential component of all programs and serve, among other purposes, to exclude low-ranking parents from both breeding and production populations.

A breeding population, therefore, is a subset of the selected individuals that have proven their desirable genetic qualities through progeny testing and will serve as parents to subsequent generations. At establishment, a compromise is made between the attainment of high gains through a strong selection differential, i.e., the difference between the mean of the base population and the breeding population, and the maintenance of a sufficient level of genetic variation by retaining enough individuals in the breeding population so that future progress is secured (Zobel and Talbert 1984). At any stage of any breeding generation, the breeding

population can be supplemented by material from external sources, such as a gene resource population or another ongoing breeding program. The supplementary material is called an infusion population (White 1987), and this step is usually preceded by selection so that the material fits well into a given improvement program's objectives.

Independent from all populations, and the breeding cycle in particular, lies the production population, sometimes also called deployment population. It is a highly selected subset of the breeding population, irrespective of its level of advancement (i.e., first or advanced breeding cycles), that usually includes as few as several dozen of the very best individuals, and is solely dedicated to the production of genetically improved seeds (seed orchards) or vegetative propagules (stool-beds) for operational reforestation programs (El-Kassaby 1992). Although the establishment of production populations is sometimes considered as the final step in each tree improvement cycle before operational plantations are established (e.g. White 1987; Zobel and Talbert 1984), El-Kassaby (1992) also stresses the importance of considering other factors such as seed harvest, handling, and storage practices, variance in seed germination, and thinning and culling in forestry nurseries. Such factors were proven by El-Kassaby (1992) to significantly impact both the genetic gain and diversity of the regeneration material. Some management practices relevant to aspects of this dissertation are outlined in later sections.

1.3 Seed Orchards

Seed orchards are by far the most frequently used sexually reproducing forestry production populations in the world. They consist of selected, genetically superior parents or their progeny with the main objective to produce consistent, abundant, and easy-to harvest yields of genetically improved seed (e.g. Feilberg and Søgaard 1975; Zobel et al. 1958). Seed orchards represent the link between tree breeding and subsequent silvicultural activities (El-Kassaby 1992, 2000), because breeding efforts are packaged into improved seed crops and subsequently conveyed into operational forestry in the form of genetically improved seedlings (Smith and Adam 2003). According to Schreiner (1961), the first seed orchards were formed as early as 1880 by Dutch colonists in Java to propagate the indigenous plant, *Cinchona ledgeriana*; however, use of this concept did not become widespread until about seven decades later when evidence for substantial improvement through selection and breeding became apparent (Schreiner 1950; Zobel et al. 1958). Aside from the genetic

improvement *per se*, seed orchards are also used for the production of seed that is well adapted to specific environments (Nanson 1972).

Seed orchards can be categorized according to the generation of breeding or the type of plant material used – seedling or clonal (Eriksson et al. 2006). First-generation orchards are planted with parents selected based on their phenotypes in forest stands or unimproved plantations (i.e., founder populations) and their genetic worth is largely unknown. These orchards can be improved by roguing of inferior parents (in earlier-published literature also called 1.5th-generation orchards), following results obtained from progeny testing (i.e., backward selections). As breeding programs progress, advanced generation orchards are established using individuals from controlled crosses between the elite genotypes (or from their open pollinated progenies with a subsequent reconstruction of the paternal parentage (El-Kassaby et al. 2011; El-Kassaby and Lstibůrek 2009)) to obtain increased genetic gain (i.e., forward selection). The effectiveness of these forward and backward selection options, including combinations of both, have previously been evaluated (Burdon and Kumar 2004; Hodge and White 1993; Ruotsalainen and Lindgren 1998).

Regarding the material used, seedling seed orchards are established using the progeny of plus trees (either maternal half-siblings from open-pollination or full-siblings from controlled crosses), whereas clonally established seed orchards utilize vegetative propagules such as grafts, cuttings, or tissue culture plantlets collected or cultivated from plus trees (El-Kassaby and Askew 1998; Zobel and Talbert 1984). Seedling seed orchards usually encompass a broader genetic base due to the larger number of individuals involved in mating (the male component of the open-pollinated progenies is often unknown); on the other hand, they are expected to provide lower genetic gain because their pedigrees are not fully known (Toda 1964). Use of seedling seed orchards is limited to early-flowering species such as *Eucalyptus* spp., black spruce, some pines and hardwoods (Kang 2001) or those which may suffer from scion-rootstock incompatibility such as Douglas-fir (Copes 1974). An alternative to seedling seed orchards, called polycross seed orchards, was proposed by Baradat (1987). In a polycross seed orchard, the genetically tested plus trees are crossed following a polycross mating design instead of being vegetatively propagated, and their progenies then serve as the material for the seed orchard's establishment. However, the vast majority of

seed orchards are clonally established and on average consist of 30–50 different parents, each of which is represented by one or more vegetatively propagated units/copies, called ramets. Seed orchards have been established with the assumption that they would function as closed, perfect, panmictic populations (Eriksson et al. 1973); in other words, they would possess neither genetic nor behavioral restrictions to random mating, and thus all individual ramets would have an equal chance to mate with all other ramets, including themselves. If this assumption held true, then both allele and genotype frequencies in the offspring population would reflect those of the parental population, in the absence of disturbing forces such as migration, selection, mutations, and random genetic drift (Hardy 1908; Weinberg 1908). Therefore, the genetic gain attained during previous cycle(s) of breeding would be effectively transmitted to the next generation, i.e., a seed crop (El-Kassaby 1989). It is not surprising that this assumption was found to be far too optimistic and that expectations differ greatly from reality (Askew 1988; Burczyk et al. 1997; Edwards and El-Kassaby 1996; El-Kassaby 2000; Gömöry et al. 2003; Kjær and Wellendorf 1998; Matziris 1994), making the actual genetic quality of seed crops unpredictable.

1.4 Genetic Quality

Genetic quality is a rather general expression with ambiguous usage that can be viewed from a number of perspectives to serve a given purpose. For instance, Hunt et al. (2004) concluded that the only meaningful definition of genetic quality is a breeding value of an individual for total fitness. In this case, total fitness is the number of descendants produced by an individual, relative to the average number produced by other individuals in the population (Dawkins 1982). Hunt et al. (2004) also pointed out that in a number of population genetics studies, only one or a few fitness components are used as indices of genetic quality because they are believed to be positively correlated with the total fitness. Although the genetic quality of an individual is mostly related to the survival rate of its offspring (in other words, an individual is of higher genetic quality when it possesses a genotype that increases its fitness relative to that of an individual with a different genotype (Neff and Pitcher 2005)), in forest tree breeding this concept would not suffice as it only represents one of many aspects that need to be considered. Seed yield in production populations, seed germination in forest nurseries, and seedling survival in operational plantations are important, but genetic quality in forest tree improvement programs, and in

breeding in general, is rather understood in a different way: to reflect the values from the breeders' point of view that can be readily quantified and compared with that of natural, founder populations. In this context, the term quality actually embodies two components – genetic gain and genetic diversity, which act in opposite directions, meaning that one can only be increased at the expense of the other. Balancing these key variables has been the subject of extensive research in both plant and animal breeding (Brisbane and Gibson 1995; Funda et al. 2009 (Chapter 7); Lindgren and Mullin 1997; Meuwissen 1997; Toro and Pérez-Enciso 1990; Villanueva and Woolliams 1997; Wray and Goddard 1994), although Rosvall (1999) points out that it is not a straightforward task due to the complexity of breeding programs and the variety of methods implemented (Caballero et al. 1996; Quinton and Smith 1995).

1.4.1 Genetic gain

Most of the phenotypic traits under consideration by tree breeders are polygenic in nature and follow the quantitative mode of inheritance where the environment could be a major factor. These traits are referred to as quantitative or continuous traits due to their control by many genes of minor effect, and because along with the genes' interactions with the environment they produce continuous phenotypic variability (Falconer and Mackay 1996). The essence of genetic gain for quantitative traits lies in the response to selection (R), which is the shift between the mean phenotypic value for the offspring of the selected parents and that of the parental generation due to selection. It can be expressed as

$$R = b_{op} S \quad [1.1]$$

where b_{op} is the regression of offspring on parents and S is the selection differential, i.e., the average superiority of the selected parents. Provided there is no non-genetic cause of parent-offspring resemblance and no natural selection, the regression can be replaced by narrow-sense heritability (h^2), which is the ratio of additive genetic variance to total phenotypic variance, and thus Eq. 1.1 becomes

$$R = S h^2 \quad [1.2].$$

In order for breeders to quantify the success of their improvement programs, the superiority of selected parents must be evaluated relative to the original population using the concept of

breeding values. A breeding value is the sum of the average effects of all quantitative trait loci that an individual carries (summed over the pair of alleles at each locus and over all loci affecting the trait of interest) or, more practically, it is a value of an individual based on the field performance of its progeny. Breeding values can be expressed in the same absolute units as the trait in question or, after standardization, as a deviation from the population's mean. Genetic gain then equals the average breeding value of all individuals in a production population, which under the assumption of random mating is transmitted to the next generation. In seed orchards the potential or expected genetic gain is calculated in a similar fashion, but the breeding values are weighted by parental representation (i.e., the number of ramets representing a particular parent) (Stoehr et al. 2004). However, the actual genetic gain realized by means of seed crops is unknown and may substantially differ from the expected value (St. Clair 1993) depending on the variation in parental reproductive success and the magnitude and quality of pollen contamination (Lindgren et al. 2004; Stoehr et al. 1998). Genetic gain (ΔG) is calculated as

$$\Delta G = \sum_{i=1}^N x_i p_i \quad [1.3]$$

where x_i and p_i denote breeding value and reproductive success of parent i , respectively, and N is the number of individuals in the seed orchard population.

1.4.2 Genetic diversity

Genetic diversity is one of the most important parameters to be considered in breeding programs. It is defined as the total amount of genetic variation carried by individuals within or among genetic units as a consequence of their evolutionary pathways and forms the basis for their response towards biotic and abiotic influences; thus, genetic diversity is regarded as a prerequisite for adaptability and further evolution (e.g. Avise and Hamrick 1996; Hedrick 2001; Reed and Frankham 2003). According to Fisher's fundamental theorem of natural selection (Fisher 1999), the rate of increase in the mean fitness of any population at any time is directly related to its genetic variance in fitness at that time (Edwards 1994; Wallace 1981). Genetic diversity can be understood in several different ways. It is commonly expressed as heterozygosity, average number of alleles per locus, or average proportion of polymorphic loci (Hedrick 2005), and can be estimated from different types of genetic

markers (see Vendramin and Hansen 2005; Vignal et al. 2002 for reviews). Another option is to relate genetic diversity to the concept of common ancestry (co-ancestry), which is defined as the probability that any two randomly sampled alleles (one from each parent) at a given locus are identical by descent (IBD) (Malécot 1948). Genetic diversity is then calculated as

$$GD = 1 - \Theta \quad [1.4]$$

(Weir 1990) where Θ is the average co-ancestry of all pairs of population members including themselves (also termed group co-ancestry) (Cockerham 1967) or, in other words, the accumulated loss of genetic diversity (Lindgren and Kang 1997).

Most often, genetic diversity is approximated by effective population size, N_e (Wright 1931), which is a concept that has been widely used by quantitative and population geneticists and included in most animal and plant breeding programs. Effective population size is an abstract concept representing the size of an ideal population that produces the same level of genetic drift as the population under consideration (Crow and Kimura 1970). It was developed to account for all possible factors affecting changes in allele frequencies among generations due to non-random sampling of the gametes that are transmitted to the next generation (Wright 1969). In other words, the concept of effective population size accounts for factors which reduce a population's census number to a size that is relevant for evolution (Begon et al. 1996; Krebs 2002). These include situations when: 1- breeding individuals only represent a portion of the entire population because some individuals have not reached sexual maturity or participated in mating, 2- there are unequal numbers of male and female breeding individuals, 3- parental gametic contributions vary (i.e., there is unequal reproductive success among parents), 4- there is spatial or temporal variation in population size, and 5- individuals that mate are related or inbred. Effective population size can be calculated as inbreeding, variance, or eigenvalue effective population size (Crow and Denniston 1988), which relate to the sizes of ideal populations that would produce the same amount of inbreeding, variance in allele frequency, or loss of heterozygosity as the actual population, respectively. A ratio between the effective and census population sizes is often calculated so that the impact of the above-mentioned factors can be directly quantified.

Monitoring and managing genetic diversity is of a major concern to breeders because its reduction is associated with severe consequences: genetic drift leads to unpredictable random changes in allele frequencies, whilst co-ancestry and inbreeding gradually accumulate over generations, causing a fitness decline in affected individuals. This phenomenon is called inbreeding depression, and although it appears to be nearly universal, its impact varies across species (e.g. Keller and Waller 2002; Lynch and Walsh 1998) and even between populations within species (Kärkkäinen et al. 1996). To provide a tool for breeders that quantifies the reduction in diversity due to co-ancestry or inbreeding (for instance, in a population of parents in a seed orchard), Lindgren et al. (1995) and Lindgren et al. (1997) developed a measure called effective status number (N_e) which can be interpreted as the size of an idealized population consisting solely of unrelated and non-inbred individuals. It is estimated as

$$N_e = \frac{1}{2\Theta} = \frac{1}{2 \sum_{i=1}^N \sum_{j=1}^N p_i p_j c_{ij}} \quad [1.5]$$

where N is the population size, p_i and p_j are parental representations of parent i and j , respectively (i.e., their expected or actual reproductive successes) and c_{ij} is co-ancestry between them (Lindgren and Mullin 1998). Consequently, status number can be derived for any population linked by a pedigree to an initial population with known co-ancestry and inbreeding (Lindgren et al. 1996). In addition to the parental population, this concept can be applied to estimate the genetic diversity of seed crops (Lindgren and Mullin 1998), whereby parental representations p_i from Equation 1.5 (which are equal to the expected gametic contributions of the respective parents) can be replaced by actual parental reproductive success determined from parentage analyses using DNA markers (Chapters 2–5) or approximated by visual assessment of their fecundities (Funda et al. 2011 (Chapter 6); Stoehr et al. 2004; Woods 2005). If all parents are unrelated and non-inbred, then the co-ancestry matrix with elements c_{ij} takes the form $0.5\mathbf{I}$ (where \mathbf{I} is an identity matrix; details are provided in Chapter 7) and Equation 1.5 simplifies to

$$N_e = \frac{1}{\sum_{i=1}^N p_i^2} \quad [1.6]$$

a form known under the term “effective number of parents”, but used and interpreted by many authors as effective population size (e.g. Kang and Namkoong 1988; Kimura and Crow 1963; Robertson 1961; Woods et al. 1996) because it reduces the census number of parents to that of an ideal population in which all parents have an equal chance of mating with all other parents. Furthermore, if all parents have equal reproductive success, then the status number reaches its maximum and is equal to the actual census number of the parental population (Lindgren et al. 1996). This feature is particularly useful because it facilitates simple and straightforward interpretations of both the expected impact of relatedness and inbreeding present in the population, as well as the variance in parental reproductive success on the reduction in genetic diversity experienced by the next generation (i.e., a seed crop). Lindgren and Mullin (1998) also point out that unlike the effective population size this concept makes no comparison with the Wright-Fisher’s ideal population, but rather it simply makes comparison against the reference population. Burrows (1984) proposed a similar method, but due to the exclusion of self co-ancestries from his calculations this concept was unsuitable for self-fertilizing species, and thus for forest tree breeding in general. Assuming both sexes contribute unequally to the gamete pool, the contribution of parent i can be split into female (f) and male (m) component such that

$$p_i = \left(\frac{f_i + m_i}{2} \right) \quad [1.7]$$

where $0 \leq f_i \leq 1$, $0 \leq m_i \leq 1$, $\sum_{i=1}^N f_i = 1$, and $\sum_{i=1}^N m_i = 1$.

Reviewing the factors that reduce census to effective population size, all the factors applicable to seed orchards are well covered by this concept. Therefore, in the relevant chapters of this thesis I will be using status number for estimating genetic diversity and will consider “effective population size”, “effective number of (male/female) parents” and “(effective) status number” to be interchangeable terms.

Godt et al. (2001) investigated whether seed orchards are capable of capturing the majority of genetic diversity existing within species’ natural populations. Comparing allozyme diversity of white spruce and jack pine seed orchards with the species’ natural populations occurring within the source area, they found no significant differences in expected heterozygosities and mean genetic identities, indicating that the seed orchards captured the major alleles present in

natural populations in similar frequencies and thus that they are not associated with gene diversity reduction. However, they showed that the plus tree selection led to missing several low-frequency alleles that existed in the natural populations. These findings were consistent with other studies reporting a similarly high genetic diversity (Sitka spruce (Chaisurisri and El-Kassaby 1994), Douglas-fir (El-Kassaby and Ritland 1996), Engelmann spruce (Stoehr and El-Kassaby 1997), longleaf pine (Schmidtling and Hipkins 1998)) and fewer rare alleles (El-Kassaby and Ritland 1996; Stoehr and El-Kassaby 1997) in seed orchards compared with the species' natural populations; however, seed orchards may possess an even greater allelic richness than is on average found in natural populations (Chaisurisri and El-Kassaby 1994). This phenomenon is probably due to the plus tree selection being commonly conducted over a wider geographic area, thus covering individuals representing a large number of subpopulations, and in fact stresses the positive role of seed orchards as they associate individuals that would normally not be able to interbreed due to the presence of physical barriers. This feature is particularly important in insect-pollinated species with a scattered distribution such as the highly valuable wild cherry or wild service tree (*Sorbus aucuparia*) in Europe which otherwise have a very limited access to pollen from more distant and thus (probably) less related sources.

Hosius (Hosius et al. 2000) proposed a multi-step selection approach where the original selection of a larger number of plus trees for desired quantitative traits in the base population is followed by the trees' genetic screening for marker-based diversity, which subsequently serves as a means for further selection (i.e., for 1st generation seed orchard establishment). The final pool of genotypes that meet both breeders' objectives and optimum genetic diversity requirements (homozygosity for different common alleles and preservation of rare alleles) is then used for an orchard's establishment. While originally tested using isozymes, this approach can be further employed for seed orchard establishment or thinning using genome-wide scans where knowledge on significant associations between SNPs and the traits of interest could be exploited during the selection phase.

1.5 Factors Affecting Genetic Quality in Seed Orchard Crops

Both genetic gain and diversity are controlled as complex measures of the overall genetic quality of orchard seedlots (Stoehr et al. 2004; Woods et al. 1996). The former provides an estimate of a seedlot's average level of improvement for a desired commercial trait whereas

the latter indirectly quantifies its potential for coping with unpredictable events such as extreme weather conditions or pest problems. As mentioned above, there are several factors that obstruct random mating among seed orchard parents, and thus reduce effective transmission of desired genes from parental to offspring populations, causing deviation from Hardy-Weinberg expectations (i.e., allelic disequilibrium). The following section addresses each factor individually and, where possible, their interactions will also be discussed.

1.5.1 Reproductive success

Reproductive success is the relative gametic contribution of a given parent to the next generation's gene pool. It is often derived from the number of offspring that a parent has produced, but a more strict definition also considers the fertility of the offspring (Clutton-Brock 1990); thus, all sterile offspring are excluded from reproductive success calculations regardless of their actual performance. Although considering offspring fertility is valid when looking at reproductive success from an evolutionary perspective, in this thesis I assume that all individual trees that would develop from the tested embryos/seedlings are fertile, because testing for this characteristic would require completion of the reproductive cycle. Additionally, under the strict definition of plantation forestry, the planted individuals' main function is fibre production rather than their reproductive capability.

There are a variety of methods for predicting or estimating parental reproductive success in forest tree seed orchards. The simplest and least expensive method assumes that parental reproductive success is a function of parental representation; therefore, the production of male and female gametes reflects the number of ramets representing a given parent (Bondesson and Lindgren 1993; Lindgren et al. 2009; Lindgren and Matheson 1986) or their total crown volume (Burczyk et al. 1996). Other commonly used methods approximate reproductive success from reproductive investment and assume that a high positive correlation exists between the production of male and female gametes (fecundity) and the actual reproductive success (Gregorius 1989). These methods include ocular scoring of generative buds (Philipson 1997) or male and female strobili production (Kang 2000; Muona and Harju 1989; Philipson 1997; Schoen and Stewart 1987) on standing trees, measuring seed-cone volume (Savolainen et al. 1993) or mass (Woods 2005), counting the number of seeds per cone and estimating the proportion of filled seeds per parent (El-Kassaby and Cook 1994; Reynolds and El-Kassaby 1990), or classifying individual ramets into groups of similar

fecundities (Burczyk and Prat 1997). The utilization of knowledge on the variation in parental reproductive phenology has also been proposed (Chaix et al. 2007). According to a number of published studies, substantial variation in reproductive investment has been observed among parents of practically all forest tree species for both male (e.g. Kang 2000; Savolainen et al. 1993) and female components (Askew 1988; El-Kassaby and Askew 1991; El-Kassaby and Reynolds 1990; Kang and Lindgren 1998; Savolainen et al. 1993). In addition, the “20:80 rule” has been coined, asserting that 20% of an orchard’s parents produce as much as 80% of the entire seed-cone crop (Anonymous 1976), and only small deviations from this ratio were later observed for black spruce (O'Reilly et al. 1982), radiata pine (Griffin 1982), loblolly pine (Schmidtling 1983), and Douglas-fir (El-Kassaby et al. 1989). Sometimes for economic reasons, the male component of parental reproductive success is not considered separately and is assumed to equal to that of females (e.g. Kang et al. 2001b; Lindgren et al. 2004; Prescher et al. 2006). However, this approach is not appropriate due to weak correlations between male and female fecundities (e.g. Funda et al. 2011 (Chapter 6); Kang 2000), that probably results from a limited supply of resources and their unequal allocation between the two genders (Savolainen et al. 1993). Information on among-family variation in germination capacity, i.e., the total number of germinated seeds relative to the total number of seeds tested after a given period of time (Int. Seed Test. Assoc. 1985), may also be utilized because the number of seeds produced by a female parent (even if assessed without error) may not truly correspond with the development of viable seedlings.

Following the rapid development and spread of a variety of molecular genetic markers in recent decades, previously reported conclusions based on scoring parental fecundities in seed orchards could be verified. A substantial male imbalance has been reported by many authors (Funda et al. 2008 (Chapter 2); Goto et al. 2002; Moriguchi et al. 2004; Schoen and Stewart 1986, 1987; Slavov et al. 2005b; Stoehr and Newton 2002; Stoehr et al. 1998; Xie and Knowles 1992), confirming the apprehension of Kjær (1996) regarding over-representation of a few highly fecund male parents in tree seed crops. However, only a few of these studies employed molecular markers to investigate the relationship between reproductive investment and reproductive success in seed orchards. The first molecular results were produced using allozymes by Schoen and Stewart (1986), who found a significant positive relationship

between parental male-cone production and proportion of seed sired in white spruce, with the former explaining 61% of the variance in the latter. Similar results have been obtained by Burczyk and Prat (1997) in a Douglas-fir seed orchard, with the exception of early-flowering phenology class (the proportion of variance explained was not provided), and several studies using microsatellite DNA markers in Nordmann fir (Hansen and Nielsen 2010), Japanese black pine (Goto et al. 2005), and sugi (Moriguchi et al. 2007) seed orchards, in which male fecundity explained 76, 43, and 15% of the variance in reproductive success, respectively. It should be pointed out, however, that in all of these studies an equal or at least similar quantity of seed with known maternal parentage was analyzed and thus only male reproductive success could be inferred. In order to obtain unbiased estimates of the female component of the reproductive success, a random sample of bulk seed (i.e., seed mixed from all ramets of all parents into one seedlot) with unknown male and female parentage would be required. Using this bulk seed sample, independent genetic analysis of embryos ($2n$) and their corresponding megagametophytes ($1n$) (El-Kassaby et al. 2010) (Chapter 4) or of uniparentally inherited extranuclear chloroplast or mitochondrial DNA markers would be needed (Chapter 5). Plomion et al. (2001) also point out that a conifer seed (at least that of pine species) possesses two different chloroplast haplotypes. Although both of them are inherited exclusively through the father, the one in the embryo is a result of the current fertilization while the one in the megagametophyte had originated from the previous generation's pollination event and thus is identical to that of the maternal parent. This unique seed biology enables an independent inference of the two parentage components as well.

First attempts to study the female component of reproductive success were presented by El-Kassaby and Cook (1994), who compared parental seed-cone and filled-seed production (the latter was a proxy for reproductive success in their study). However, since they did not take into account the among-family variation in germination capacity, nor did they conduct parentage analyses on bulk seed as described above, their conclusions could not advance beyond stating that gametic contributions based on the two approaches differed, whilst leaving the actual female reproductive success unknown. Funda et al. (2011) (Chapter 6) presented a comprehensive study where four different male and six different female reproductive investment methods (Woods 2005) were evaluated using microsatellite DNA markers. They found that parental representation was the weakest predictor of parental

reproductive success for both male and female component as the proportion of variance explained by the respective fecundities was 42 and 19% for Douglas-fir and 20 and 31% for lodgepole pine. By contrast, they confirmed the previously published work (Burczyk and Prat 1997; Goto et al. 2005; Hansen and Nielsen 2010; Moriguchi et al. 2007; Schoen and Stewart 1986) concluding that accounting for parental fecundities does substantially improve the prediction of the reproductive success. For instance, the explained variation increased to 55 and 89% for Douglas-fir and 58 and 66% in lodgepole pine when data on pollen production (visually assessed number of male strobili) and volume of seed cones after harvest for each parent in the seed orchards were utilized, respectively. Although Denti and Schoen (1988) analyzed seed samples weighted by the total parental seed production, whereby a bulk seed collection was simulated, they did it with the aim of comparing selfing rate with that experienced by an equal sample size, rather than to infer the female component of the reproductive success.

1.5.2 Reproductive phenology

Depending on the species or the breeding program, parents in seed orchards often originate from larger geographic areas and thus may possess adaptation to different climatic conditions. Synchrony in reproductive phenology is one of the main prerequisites for random mating and thus balanced reproductive success among parents in seed orchards (Eriksson et al. 1973). Reproductive phenology pertains to life cycle events involving sexual reproduction and their influence by seasonal and inter-annual environmental variation such as bud burst (Boyer 1973) or duration of pollen shedding and seed-cone receptivity during the pollination period. Because some components of reproductive phenology are known to have strong genetic control (Blush et al. 1993; El-Kassaby et al. 1984; Matziris 1994; Worrall 1983), parents selected from founder populations are also expected to retain their relative phenological characteristics regardless of the *ex situ* conditions to which they have been transferred. Consequently, although physically occurring next to one another, the parents are split into a number of non-overlapping phenological classes that create temporally isolated subpopulations with no or very limited gene exchange (El-Kassaby 1995). Asynchronous phenology has been observed in seed orchards of Douglas-fir (Copes and Snieszko 1991; El-Kassaby et al. 1984; Fashler and El-Kassaby 1987), Sitka spruce (El-Kassaby and Reynolds 1990), radiata pine (Griffin 1984), loblolly pine (Askew 1986; Askew

and Blush 1990) and black pine (Matziris 1994), while it was found to have stronger synchronization in seed orchards of Scots pine (Jonsson et al. 1976), black spruce (O'Reilly et al. 1982), and Norway spruce (Nikkanen 2001). Female reproductive phenology and the initiation time of female receptivity in particular appeared to have a stronger genetic control than its male counterpart (Griffin 1984; Matziris 1994; Nikkanen 2001) and the duration of pollen shedding has been found to be completely determined by environmental factors (Matziris 1994). Although in a western redcedar seed orchard Worrall and El-Kassaby (in El-Kassaby 1999) observed a nearly uniform heat-sum requirement for reproductive bud burst within the entire population (100 parents), they attributed their findings to the species' high phenotypic plasticity as its survival strategy rather than a lack of genetic variability for this trait. The presence of variation in reproductive phenology among a seed orchard's parents leads to a reduction of the possible mating combinations, causing positive assortative mating among them (i.e., the preferential mating between individuals with similar phenotypes) and imbalanced reproductive success. As such, even the most fecund parents (both male and female) may be partially or completely excluded from sexual reproduction with other mates if males shed pollen while females are not yet or no longer receptive, or if females have receptive seed-cones while no ambient pollen from within-orchard individuals is available for their fertilization.

1.5.3 Inbreeding

Inbreeding refers to mating between genetically related individuals including self-fertilization and is usually quantified by the coefficient of inbreeding F , which equals the coefficient of co-ancestry or kinship among individuals in the previous generation (Malécot 1948). This has similarities with assortative mating; however, whilst assortative mating affects only a limited number of loci that control or are linked with a trait of interest, inbreeding affects all loci indiscriminately and causes genome-wide excesses of homozygosity (Hedrick 2005).

In most species, inbreeding results in a fitness decline due to the increase in homozygous genotypes with deleterious alleles; this is called inbreeding depression. Studies on the genetic basis of inbreeding depression indicate that a small number of genes confer large deleterious effects, whereas most other deleterious effects are minor (Crow and Simmons 1983). Although the genetic consequences of inbreeding depression may vary among species

(Lynch and Walsh 1998) and even among populations within species (Kärkkäinen et al. 1996), inbreeding is a critical factor that must be considered in all breeding programs. This is especially true for hermaphroditic species such as forest trees because of their self-fertilization capability; in each cycle of self-fertilization individuals' homozygosity increases by one half, thus magnifying the chance of producing genotypes with only rare deleterious alleles. Rare deleterious alleles arise mainly as unfavorable mutations or are evolutionary relics, and although some of them may become favorable later (Lindgren and Gregorius 1976; Williams et al. 1995), their homozygosity via inbreeding is supposed to be the primary cause of inbreeding depression. Examples of unfavorable rare alleles can be found in the resistance of Douglas-fir to damaging biotic agents (Chen et al. 2001) or the survival of loblolly pine (reviewed by Mitton 1995), in which homozygotes for the rarer allele had lower values than homozygotes for common allele; resistance to pathogens in agricultural crops is often connected to such alleles (Burdon 2001; Lindgren and Gregorius 1976). Some plant species have developed a variety of pre- and post-pollination mechanisms to minimize or fully avoid selfing, such as self-incompatibility in angiosperms (reviewed by de Nettancourt 1997), dioeciousness, physical separation of male and female strobili within trees of monoecious species (Park and Fowler 1984), asynchrony of male and female reproductive phenology (El-Kassaby et al. 1984; Erickson and Adams 1990; Eriksson et al. 1973), or polyembryony (Sorensen 1982). Aside from self-fertilization which can occur in any generation's seed orchards, half-sib, full-sib and parent-offspring matings are likely in advanced-generations orchards where parents have been selected from material obtained during previous cycle(s) of breeding.

Negative effects of inbreeding such as reduced seed set, survival, growth, and productivity have been observed in many forest tree species (Denti and Schoen 1988; Eriksson et al. 1973; Geburek 1986; Griffin and Lindgren 1985; Johnsen et al. 2003; Koelewijn et al. 1999; Morgante et al. 1993; Orr-Ewing 1974; Park and Fowler 1984; Samuel et al. 1972; Snyder 1972; Sorensen 1999; Sorensen et al. 1976; Squillace and Kraus 1963; Wang et al. 1999; Woods and Heaman 1989; Wu et al. 1998). Progeny of self-pollinated or otherwise inbred trees are typically substantially smaller than those of outcrossed trees and this difference can be further accentuated during stand development (Williams and Savolainen 1996) as the slow-growing inbred trees are eventually outcompeted for light and other resources by more

vigorous outcrossed trees. Western redcedar appears to be an exception from this norm: its selfing rate was found to be relatively high in both natural and artificial populations, but nearly no inbreeding depression was determined (Russell et al. 2003). In a study on redwood, Libby et al. (1981) found no consistent effect of selfing on number of seeds per cone, but they reported that under stress conditions, survival from self-fertilized seed was lower than from outcross seed because height was just 65–80% after one year's growth, but it was greater and more balanced under good nursery conditions. After 14 years in the field, the difference in growth was further accentuated and the self-fertilized individuals averaged only 42% the height of their outcrossed counterparts. It is also possible that low to moderate levels of inbreeding such as crossing half-sibs may increase production of filled seeds (Andersson et al. 1974; Sorensen and Cress 1994), leading to larger proportions of inbred seedlings that still meet nursery culling standards. Consequently, these inbred seeds may have an even more severe impact on growth and yield of forestry plantation than self-fertilization (Wang et al. 2004), although Wang et al. (2004) found that inbreeding in seed orchards is likely to have little effect on stand yield.

1.5.4 Gene flow

Gene flow is genetic exchange between populations. Although in many situations, especially when dealing with animal species, gene flow was difficult to estimate because migrants are hard to identify, leave no or fewer offspring than individuals that have not moved, or stay briefly but still mate, the introduction of molecular genetic markers enabled unambiguous identification of parentage and determination of gene flow between generations without any information about their direct movement (Hedrick 2005). In forest tree seed orchards, gene flow is easier to determine because trees are sessile and the gene flow of interest is only limited to one direction: from background, external pollen sources into the seed orchards. Genotyping all candidate male parents within an orchard and assigning the offspring (seed crop) to one of the candidate fathers makes it possible to partition male gametes originating within orchard trees from gametes arising from external sources. Gene flow into a seed orchard from unselected, background pollen sources is known as pollen contamination and it seriously affects seed crops' genetic quality (Adams and Birkes 1989).

The level of pollen contamination depends on the species, their mating system, reproductive phenology of both within- and outside-orchard parents, seed orchard size and the degree of

isolation, and actual environmental conditions. A great number of pollen contamination studies have been reported for many species to date, with the rate ranging from nearly zero (El-Kassaby and Ritland 1986a) to up to 90% (Fast et al. 1986; Kaya et al. 2006). A study on a Norway spruce seed orchard covering three different years showed a high but consistent level of contamination (69-71%) (Pakkanen et al. 2000) whereas a more than two-fold difference (44 and 89%) was observed in a Douglas-fir seed orchard over two consecutive years (Fast et al. 1986).

As the two components of seed crop genetic quality, genetic gain and genetic diversity are affected by contamination in the opposite direction. Genetic gain decreases as contamination increases (Wheeler and Jech 1986) unless the average breeding value of contaminant sources is equal to or higher than an orchard's parents. Under equal reproductive success among parents the decrease in genetic gain equals half the contamination rate. On the other hand, additional parental contributions from pollen contamination usually increases genetic diversity of the seed crop (Lindgren and Mullin 1998). Also, if contamination is high, then diversity may either increase or decrease, depending on genetic characteristics of the source population (Adams and Kunze 1996). In conifers, which are all anemophilous, pollen is adapted for aerial transport by the presence of special bladdery wings in pine, cedar, fir, and spruce, or large spherical grains as in Douglas-fir and larch (Wodehouse 1935). Consequently, pollen contamination may reach high levels, especially if conspecific stands are present in the vicinity of a seed orchard.

Since pollen contamination affects only the male component of reproductive success, half of all genes will originate from within-orchard parents even under 100% contamination rates. However, the genetic gain reduction may be further exacerbated if the contaminating pollen comes from trees with poor adaptation to an intended planting site (Kang 2001) such as the original location of the parental population, because adaptation depends on both male and female gametic contributions. Another issue to be considered is the possible preferential nursery-stage selection of faster-growing seedlings sired by contaminant pollen, which may be outcompeted by other plant species on sites to which they are maladapted.

While quantification of pollen contamination in clonal seed orchards is relatively straightforward using paternity analysis from molecular genetic markers, in seedling seed orchards it is relatively more difficult due to the higher number of genotypes present.

Therefore, estimation of pollen contamination based on differences between allele frequencies derived from single-locus outcrossing pollen (male), ovule (female), and outside seed has been suggested (El-Kassaby and Davidson 1990; Wheeler and Jech 1992). Usage of highly polymorphic genetic markers (Brondani et al. 1998) however seems to enable paternity inferences with high confidence even in these types of orchards (Chaix et al. 2003) and, moreover, it may soon be anticipated that the genome-wide scanning using SNPs will fully eliminate the issue of a large parental population size.

1.5.5 Seed germination

Numerous studies have shown that seed germination attributes such as time at germination onset, germination speed, and germination capacity vary substantially both among and within populations. The variation among populations may be due to provenance (Allen 1961), environmental conditions during seed development (e.g. Almqvist et al. 1998; Koller 1962; Sawhney and Naylor 1979), family affiliation (Bramlett et al. 1983; El-Kassaby et al. 1993b), seed pretreatments (Kozlowski and Gentle 1959), seed maturity (Edwards 1980; Edwards and El-Kassaby 1988), or seed size (Dunlap and Barnett 1983; Shoulder 1961). A standard germination test (Int. Seed Test. Assoc. 1985) is usually conducted to obtain germination attributes. Such tests are known to carry the risk of overestimating seedlot field performance due to unrealistically stable and uniform germination conditions (Chaisurisri et al. 1992), but even so, they provide a reasonable approximation of maximum possible seedling production. Although variation in seed germination has no direct effect on genetic quality of seed crops because seed germination follows seed harvest, storage, and possible pre-treatment, it may affect the quality of established forest stands. While variation in germination onset and speed may cause seed from late- and slow-germinating families to produce seemingly sub-standard seedlings which will be culled during the nursery stage and thus unintentionally excluded from future development, variation in germination capacity can substantially divert the genetic parameters from those expected based on clonal seed harvest even with no unintentional selection, because proportional gametic contribution of the families involved will change.

1.6 Methods to Manipulate Genetic Quality

Seed orchard managers have developed a variety of methods for manipulating crop genetic quality. These methods are applied to different stages of the seed orchards' life cycle, but have one common objective: to increase genetic gain while maintaining sufficiently high genetic diversity.

1.6.1 Population size

The number of different genetic entries to be included as parents in a production seed orchard is one of the most fundamental questions to be addressed and there is no universal answer. Theoretically, one single parent with the highest breeding value would generate the highest possible genetic gain in a seed crop, but the seed crop's genetic base would be too limited and the resulting forest stands would lack resilience to unpredictable environmental changes. Furthermore, all the seed produced would be the product of selfing. Lindgren and Prescher (2005) noted that despite the relevance and importance of determining the number of parents in a seed orchard, this subject had not received full discussion in the literature and no journal paper has focused exclusively on this issue. The North Carolina State University Tree Improvement Cooperative (2001) recommended using 20 to 30 parents because a larger number would reduce potential gain; however, the actual numbers used vary widely from as few as 5 to 10 parents (McKeand et al. 2003) to about 90 (Kang et al. 2001a). Lindgren and Prescher (2005) tried to provide a more generalized way of calculating the optimum number of parents, but concluded that the range of optimum values was quite broad; subsequently, 20 parents was chosen as the standard for Swedish orchards. They also pointed out that their suggestion applies only to unrelated, tested parents deployed in equal proportions (which is not optimal anyway), and stressed that the appropriate number of parents may be case specific, i.e., when more relevant information is available, higher or lower numbers may be optimal.

The census number of parents is a crucial starting point to seed orchard establishment, but as mentioned earlier, census number may not be fully reflected in seed crops. Therefore, it was suggested that effective rather than census number of parents should be controlled for each seedlot and used as a criterion for estimating genetic diversity in seed crops because this measure is a better reflection of their actual gametic composition (El-Kassaby and Cook

1994). This suggestion was implemented e.g. in British Columbia where the effective number of any seedlot used for reforestation of Crown land must not drop below 10 (Stoehr et al. 2004; Woods 2005); a value generally considered to capture the majority (95%) of genetic diversity existing in a base population (Nei 1973; Yanchuk 2001). Similarly, Lindgren (1989) originally suggested a value of 8 and in Alberta a more conservative value of 18 has been implemented (Anonymous 2009). Other theoretical and empirical data show that seed orchards with 20 or more parents should represent a similar level of risk as seed collected from the natural population (Johnson and Lipow 2002). Theoretical improvements concerning the effective number of parents were provided by Lindgren and Mullin (1998) and Kang (2001). Funda et al. (2011) (Chapter 6) confirmed the rationale of this approach when they observed a drastic reduction from the census (49) to effective (13.3) number of parents in a seed orchard of Douglas-fir due to a substantial imbalance in reproductive success.

One additional aspect that should be considered is whether the reforestation material from seed orchards will be allowed to regenerate naturally. While it is common practice to clear fell forest stands and exclude natural regeneration, Kjær (1995) in Lindgren and Prescher (2005) suggested that if a species has a long rotation period and is normally naturally regenerated, a higher number of parents should be considered.

1.6.2 Parental representation

Historically, clonal seed orchards have been commonly established using uniform spatial representation of parents across the population, such that each parent was replicated by the same number of ramets. When information on parental breeding values became available, parents above a certain threshold value were preferentially utilized, while those below were rejected. Lindgren (1974) suggested exploiting this information in a more sophisticated way, such that parents with higher breeding values would be planted in higher proportions in order to boost the crops' genetic gain, while inferior parents would only contribute fewer replicates to maintain an acceptably high genetic diversity. Following this idea, Lindgren and Matheson (1986) developed the linear deployment concept, whereby parental representation during orchard establishment would be linearly related to their respective breeding values; latterly Bondesson and Lindgren (1993) modified the concept to guide genetic thinning of existing seed orchards. While these two methods certainly represented a step forward, they

still possessed two major limiting assumptions: firstly, that seed orchard populations consisted solely of unrelated and non-inbred parents, and secondly, that parental reproductive success is a function of parental representation, i.e., production of viable seed and reproductively successful pollen is proportional to the number of ramets representing a particular parent.

Regarding the utilization of the linear deployment concept, it was obvious that as breeding programs advanced, both kinship and inbreeding considerations became inevitable (Kang 2001; Lindgren and Mullin 1997, 1998; Olsson et al. 2001). Lindgren et al. (2009) further extended the original linear deployment algorithm to account for relatedness among parents and proposed an “optimum” deployment. In general terms, Funda et al. (2009) (Chapter 7) proposed an optimization protocol for creating “custom seedlots” from clonally harvested seed which works with both the known female contribution of each parent (e.g. number, volume, or weight of seed cones), and predicted male contribution from pollen-bud production during a given season. While this approach does not enable long-term predictions of parental reproductive success, it has fewer assumptions, making it more accurate (as verified for four conifer species by Funda et al. (2011) (Chapter 6) as well as more flexible because it is capable of accounting for among-year variation in both male and female reproductive output.

Analogously to the concept of genetic thinning, in seedling seed orchards inferior trees are commonly rogued following one of four selection methods, namely individual selection, family selection, family-plus-within-family selection, or combined selection (Magnussen and Yeatman 1990). Their evaluation in terms of balancing genetic gain and diversity was provided by David et al. (2003), who concluded that combined selection should be applied if genetic gain is the priority while individual or family-plus-within-family selections are most suitable when genetic diversity is the priority. Under equal priorities of the two parameters, an optimization approach is suggested that maximizes their values relative to each other by determining the optimum selection intensity (David et al. 2003).

1.6.3 Orchard design

There is a trade-off between the gains obtained by selecting the best (and sometimes related) trees, and the losses associated with inbreeding or the costs associated with attempting to minimize it (Wang et al. 2004; Woods and Heaman 1989). A great variety of seed orchard

designs have been developed and implemented (Bell and Fletcher 1978; El-Kassaby 2003; El-Kassaby et al. 2007; Gietrych 1975), with a common objective to minimize self-fertilization and promote panmixia. These designs are based on the assumption that pollination efficiency is a function of distance, and thus most genetic exchange occurs among neighboring ramets while it decreases between those further apart. To test this assumption, three approaches were used to measure gene dispersal with genetic markers (Adams 1992): firstly, tracking effective pollen or seed dispersal from individual parent trees which carry rare alleles (rare marker approach) (Müller-Starck 1977; Yazdani et al. 1989); secondly, using multilocus markers to infer parentage of offspring (parentage analysis) (Neale 1984); and thirdly, modeling the probability structure of entire offspring genotypes samples and estimating the most likely dispersal parameters that would have generated the sample (model approach) (Adams and Birkes 1989; Schoen and Stewart 1986).

Levin and Kerster (1974) pointed out that the presumption from pollen dispersal patterns in wind-pollinated species that most mating occurs among neighboring plants does not take into account the influence of competing pollen sources. Furthermore, if the amount of pollen produced by near neighbors is small relative to more distant sources, the advantage of proximity may be reduced or eliminated (Adams 1992). In addition, patterns of mating estimated from only pollen dispersal do not account for differential pollen fertility, floral phenology of potential male parents, or their compatibility with specific females (Handel 1983).

The effect of distance and phenology on cross-pollination in a Douglas-fir seed orchard was studied by Erickson and Adams (1989) who estimated the proportion of viable embryos resulting from fertilization by designated male trees. The pollen source was identified using unique allozyme genetic markers that occurred in two of the orchard parents. In 38 mother trees, the proportion of embryos fertilized by the male marker ranged from 0 to 71.4%. Very little pollen was found to be dispersed beyond 30 m, but within 30 m, mating success was only weakly associated with distance. Yazdani et al. (1989) studied pollen dispersal in a Scots pine seed orchard using isozymes markers, and found that only ca. 5% of the genes originated from within 15 m of a given tree. They also developed a mathematical model for pollen distribution, according to which 90% of pollen was distributed further than 15 m from its source.

Burczyk et al. (1996) studied mating system and effective pollen dispersal in a natural stand of knobcone pine using 11 isozyme loci. Analyses were performed by fitting neighborhood and mixed-mating models to multi-locus genotypic arrays of offspring from four mother trees. Neighborhoods consisted of all potential outcrossing males within 11 m of each mother tree (44 on average) and ca 41% of matings were found to be the result of outcrossing within neighborhoods. Distance and direction of individual males from mother trees and the size of males (height) played significant roles in determining outcrossed mating patterns within neighborhoods. Lastly, although males east of mother trees accounted for more offspring than males in other directions, male mating success increased with both proximity and tree size.

Burczyk and Prat (1997) conducted a similar study on a Douglas-fir seed orchard using 11 allozyme loci to analyze progenies from 94 mother trees from the mixed-mating and neighborhood models. The proportion of matings that resulted from outcrossing within neighborhoods (30 m radius for each mother tree) was estimated to be 43%, and the effect of distance and direction of individual males from mother trees, pollen fecundity, and phenological synchronization were all significant in determining patterns of outcrossing within neighborhoods. Additionally, male reproductive success increased with pollen fecundity as well as proximity to and phenological overlap with a given mother tree. For comparison, Burczyk et al. (2002) investigated pollen dispersal in a seed orchard of an insect-pollinated species, *Eucalyptus regnans*. They found that approximately 50% of effective pollen gametes were a product of males more than 40 m distant from the respective mother trees, making insect pollinators efficient promoters of cross-fertilization in this orchard.

Since parents are usually replicated by several ramets in seed orchards, then the aforementioned objectives of all designs (i.e., minimizing selfing and promoting panmixia) have been secured through a specific spatial layout of ramets with respect to their parental affiliation. Probably the most successful design was the permuted neighborhood design (COOL; Computer Organized Orchard Layouts) developed by Bell and Fletcher (1978), which is able to effectively separate ramets of a given parent by a predefined minimum number of positions between them; however, neither this or its subsequent improvement by Chakravarty and Bagchi (1993, 1994) was capable of incorporating relatedness among

parents and inbreeding. As advanced generations orchards became the norm, it was necessary to start considering the issue of relatedness and include it as an additional input factor for designing new orchards; a solution was provided by the minimum inbreeding seed orchard design (Lstibůrek and El-Kassaby 2010).

Recently, El-Kassaby (unpublished) introduced the Randomized, Replicated, Staggered Clonal-Row (R^2 SCR) seed orchard design, which has the advantage of combining randomization with row arrangement to permits equal, linear, and custom deployment options. Rows are staggered such that each parental ramet is surrounded by four different parents and the neighboring parents vary across replications. For any selected deployment option, layouts are compared using the “minimum distance” concept (Lstibůrek and El-Kassaby 2010), the least square difference between the desired and actual clone size, and the number of empty ramet positions the algorithm could not fill. It is expected that as breeding advances and the desire to exploit the differences among parents’ breeding values increases, seed orchard designs will continue to be very dynamic so that orchard managers could balance their crop management activities with gain-diversity goals.

1.6.4 Crown management

Regardless of whether seed orchard trees are seedlings or grafts, they are kept to a manageable size by crown pruning that maintains adequate seed and pollen strobili development (Webber 2000; Zobel and Talbert 1984) and facilitate easy and safe seed cone harvest. Appropriate crown size and shape are also important prerequisites for seed crops’ genetic quality, as tree height was found to be negatively correlated with outcrossing rate in western redcedar (O’Connell et al. 2004), probably due to larger proportions of self-pollen present in the pollen cloud within a tree’s crown. Several studies have focused on within-tree variation in selfing rate relative to crown position, and in Douglas-fir (El-Kassaby et al. 1986) and Sitka spruce (Chaisurisri et al. 1994) higher values were found in lower than upper parts of the crown, although no such trend was reported for western redcedar by O’Connell et al. (2004). Techniques for physical (girdling) and hormonal cone induction are also available, the latter of which most commonly involves a mixture of gibberellins known as GA 4/7 applied either as foliar spray, stem injection, or a combination of both (Ross 1983; Ross and Bower 1989, 1991; Ross et al. 1981). These treatments generally begin at vegetative bud burst, with follow-up treatments for several weeks.

Top grafting (a.k.a. top working) is a low-cost technique that involves grafting young scion material into the branches of existing, reproductively mature trees (Bramlett and Burris 1995), resulting in composites representing three different genotypes: a rootstock, an interstock that was initially grafted onto the juvenile rootstock, and a top-graft. The underlying principle of top-grafting is that health, physiological stability, and reproductive competence of the interstock are transferred to the top-graft, and the top-grafted scion carries the desirable genetic material (Almqvist and Ekberg 2001; Hartman and Kestler 1983). While top grafting has been widely used for many years in fruit trees to widen the production of different fruit varieties (Hartman and Kestler 1983), it was adopted in forest tree improvement programs with the objectives of accelerating breeding by inducing early flowering in grafts and thus shortening intervals between generations (Bramlett and Burris 1995; Goading et al. 2000), as well as integrating new selections into mature seed orchards (Bramlett 1997). While top grafting was certainly successful in meeting these objectives in loblolly pine (Bramlett and Burris 1995, 1998), slash pine (Perez et al. 2007) and to a limited extent also in other species such as European and Japanese larch (Robinson and Wareing 1969), its use to reduce selfing has been surprisingly neglected. Since selfing as the highest level of inbreeding constitutes a major obstruction to both quality and quantity of seed crops (Section 1.5.3), then top-grafting could also be viewed as a possible way of reducing it, especially for trees with larger crowns (see O'Connell et al. 2004 for fine scale estimation of selfing in western redcedar and its correlation with tree height), because pollen grains from one part of a tree fertilizing ovules on other parts of the same tree could effectively still be outcrossing. Results from a study of selfing in one top-grafted western redcedar seed orchard (Chapter 3) suggest that this technique may have promoted outcrossing and thus improved the overall genetic quality of the seed crop. The average selfing rate was estimated to be 7.3%, a value substantially smaller than all others reported for this species both in natural (O'Connell et al. 2008; O'Connell et al. 2001) and artificial populations (El-Kassaby et al. 1994). Unfortunately, direct evidence for this statement is not available because offspring from the top-grafted and control ramets were not directly compared.

1.6.5 Supplemental mass pollination

Pollen is a crucial component of all seed orchards as it contributes one half of gametes to the offspring, provided there is negligible pollen contamination from non-orchard pollen sources.

Since pollen is relatively easy to handle, it represents a powerful tool for seed orchard managers to manipulate the genetic composition of prospective seed crops and specifically increase genetic gain (Askew 1992; Bridgwater et al. 1998). Unlike breeding activities where pollen is applied in small quantities on protected cones to perform controlled crosses among selected individuals, in seed orchards it is commonly applied in larger quantities on unprotected receptive female strobili either for an entire orchard population or on a selected subset of parents or ramets. This latter technique, called pollen augmentation or supplemental mass pollination (SMP; Bridgwater et al. 1993; Bridgwater and Trew 1981; Wakeley et al. 1966) was introduced with the objective of increasing seed crop genetic quality, and it is widely utilized by seed orchard managers to improve seed yields, balance male reproductive success, introduce specific genotypes into the seed orchard, and reduce self-fertilization and pollen contamination (Webber 2000). Before the introduction of molecular genetic techniques, evaluation of SMP effectiveness was limited to traits pertaining to seed development and viability. For instance, Daniels (1978) studied the effect of eight different pollination regimes, including SMP, on filled seed production of one selected Douglas-fir parent, but found no appreciable increase as compared with the wind-pollinated control. In a similar study on Sitka spruce, SMP produced 14 times more filled seeds per cone compared to wind pollination (El-Kassaby and Reynolds 1990). With molecular markers, SMP effectiveness could be evaluated in much more detail. It was proven to facilitate mating among reproductively asynchronous parents (El-Kassaby et al. 1988) and reduce pollen contamination (Askew 1992; El-Kassaby and Ritland 1986b; Stoehr et al. 1998), whilst increasing seed set and outcrossing, particularly for early and late reproductive phenology classes (El-Kassaby et al. 1990; El-Kassaby and Ritland 1986b). Bridgwater et al. (1993) and El-Kassaby et al. (1993a) also concluded that SMP effectiveness greatly depends on timing (see also Owens et al. 1981b) as well as the number and method of pollen applications, but other factors such as pollen contamination rate, complementary management practices such as bloom delay (El-Kassaby and Ritland 1986b), or actual environmental conditions at flowering may also influence its effectiveness. SMP success rate was anticipated around 25% (Stoehr et al. 2004); however, the actual estimates differed markedly from this expectation, reaching 4 and 19% in Scots pine (Eriksson et al. 1995; Yazdani et al. 1986, respectively), 16% in lodgepole pine (Stoehr et al. 2006), 80% in

loblolly pine (Bridgwater et al. 1987) and 15% in Douglas-fir (Lai et al. 2010). When individual strobili were pollinated, average success rates ranged between 66 and 84%, but in an operational study where whole trees were pollinated, the success declined to 7-26% (Eriksson et al. 1994).

1.6.6 Bloom delay

Bloom delay is a method based on work by Silen and Keane (1969), which involves fine water mist application in late winter and early spring to seed orchards with the objective of keeping their immediate environment cooled relative to oncoming spring conditions. The gradual evaporation of water ensures that heat-sum, a cue triggering actions pertaining to reproductive phenology, including reproductive bud development, is accumulated more slowly within seed orchards. As a result, treated trees are temporally isolated from exposure to background pollen contamination. This technique proved to effectively delay (Fashler and Devitt 1980) and compact reproductive phenology, helping to improve panmixia within the studied orchards' populations by balancing male gametic contributions (Fashler and El-Kassaby 1987), as well as reducing pollen contamination (El-Kassaby and Ritland 1986b; Silen 1970) and selfing (El-Kassaby and Davidson 1990, 1991). Another positive by-product reported for bloom delay was a decreased amount of insect infestation in orchard trees (Fashler and Devitt 1980).

1.6.7 Selective harvesting and production of custom seedlots

Another option to balance genetic gain and diversity in seed crops is selective seed-cone harvesting (Kang et al. 2001a; Lindgren and El-Kassaby 1989). While it may be necessary to harvest and utilize a whole cone crop, harvesting more cones from higher-breeding value parents may improve a crop's genetic gain while a more balanced harvesting (Bila et al. 1999; Kang and El-Kassaby 2002) may maintain higher level of genetic diversity if required. The former method relies on the expectation that abundant seed will be produced by all parents and preferential collection will be possible from superior trees, e.g. to mimic linear deployment (see Funda et al. 2009 (Chapter 7) for discussion), although this prerequisite may not be fulfilled due to large variations in seed-cone production, both among parents within and among years; a typical phenomenon in many conifers species. The latter method has an even bigger limitation because the amount of harvested seed-cones is set by the least fecund

parent and thus substantial amounts of potential reforestation material may be abandoned, rendering this method inappropriate from a practical perspective (Kang et al. 2003a). Results of Kang et al. (2001b) and Kang et al. (2005) who compared genetic gain and diversity under several different management strategies, including selective harvesting, genetic thinning, and a combination of both, indicated that the two parameters might vary considerably depending on parental genetic value, selection intensity, pollen contamination, and fertility variation. As anticipated, strong genetic thinning led to the highest genetic gain, while selective harvesting from superior parents maintained the broadest genetic base (Kang et al. 2005). However, neither of these methods takes advantage of data on male reproductive investment, which is significantly correlated with male reproductive success (Funda et al. 2011) (Chapter 6), and assume that male and female reproductive success is equal; therefore, gametic contributions can be improved from seed parents only, unlike both linear deployment of parents and genetic thinning. Furthermore, as mentioned earlier, these methods could not accommodate relatedness among parents and thus are available for first-generations seed orchards only.

Attempts to simultaneously balance genetic gain and diversity in seed orchard crops, including those described in Section 1.6.2, by Son et al. (2003) and by Lindgren et al. (2004), led Funda et al. (2009) (Chapter 7) to develop an optimization protocol which would be capable of utilizing data on both female and male reproductive investment, with the male component serving as a predictor for reproductive success (Funda et al. 2011) (Chapter 6), and accounting for relatedness among parents and inbreeding. This protocol follows the concept of quadratic optimization introduced by animal breeders (Meuwissen 1997; Villanueva and Woolliams 1997) and its original version (Funda et al. 2009) (Chapter 7) was further refined to account for among-family variation in seed germination capacity, whereby the accuracy for estimates of effective number of parents was improved. Among other advantages discussed in Chapters 7 & 8, utilization of the optimization protocol enables the creation of custom seedlots with the maximum possible genetic gain at a predefined level of genetic diversity.

1.7 Genetic Markers

Genetic markers are readily recognizable DNA segments or gene products that can be used for identification of different individuals. As mentioned above, the relatively recent

introduction of molecular markers has enabled scientists across disciplines, including population geneticists and seed orchard managers, to answer a number of important questions. An ideal genetic marker for studies that involve pedigree reconstruction analyses is heritable, highly polymorphic, co-dominant in expression, reproducible, selectively neutral, and with no or low linkage disequilibrium (Selkoe and Toonen 2006; Vendramin and Hansen 2005). Until a few decades ago, most molecular genetic studies relied on isozymes which are multiple molecular forms of an enzyme with similar or identical catalytic activities (Hamrick and Godt 1989). While isozymes still provide a relatively simple and inexpensive way of obtaining genetic information, their use is restrained by the limited number of loci and low variability in some species, as well as the ability to only reveal variation in protein-coding DNA sequences. The development of DNA markers, including restriction fragment length polymorphism (RFLP), randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), simple sequence repeats (SSRs; a.k.a. microsatellites), and single nucleotide polymorphisms (SNPs), has enabled the study of genetic variation with ever-increasing resolution (Butcher et al. 1999; Vendramin and Hansen 2005). In the following paragraphs I will describe microsatellites in more detail because I have employed them as genetic markers throughout this thesis.

Microsatellites (Litt and Luty 1989; Tautz 1989; Weber and May 1989) are co-dominant and highly polymorphic markers (Amos and Pemberton 1992), ubiquitously distributed across both prokaryotic and eukaryotic genomes (Gur-Arie et al. 2000; Hamada et al. 1982). They occur predominantly in non-coding DNA regions (Metzgar et al. 2000; Toth et al. 2000) that have been found to contain ca. 90% of all microsatellites in higher plants (Varshney et al. 2002; Wang et al. 1994) and consist of short DNA sequences (twenty to a few hundreds base pairs) defined by numerous tandem repeats that are one-to-six bases in length (Beckmann and Weber 1992). Microsatellites are easily amplified through the polymerase chain reaction (PCR) using primer pairs designed as the complementary match to their unique flanking sequences. To date they have been successfully utilized in a wide range of fundamental and applied biological and medical research, including forensics, population and conservation genetics, and genetic mapping (Chistiakov et al. 2006; Powell et al. 1996). The high levels of microsatellite polymorphism are the result of a relatively high mutation rate ranging from 10^{-2} to 10^{-6} mutations per locus per generation (Ellegren 2000) compared to that of regular,

non-repetitive DNA experiencing 10^{-9} mutations per locus per generation (Li 2006). Conversely, microsatellite flanking regions are usually highly conserved, making the primers transferable among populations, and sometimes also among species and genera. In spite of several drawbacks such as size homoplasy among alleles of identical size but not identical descent due to convergent mutations (Adams et al. 2004; Estoup et al. 2002), occurrence of artifact stutter bands resulting from polymerase slippage during PCR amplification, preferential allele amplification, and flanking region mutations that produce “null alleles” caused by the failure of allele amplification (Callen et al. 1993; Fisher et al. 1998; Paetkau and Strobeck 1995)), microsatellites are the preferred genetic markers for individual genotyping and gene flow studies, although they will probably soon be superseded by SNPs. Radiata pine was the first forest tree species for which microsatellites were developed (Smith and Devey 1994), but subsequently microsatellites have become available and extensively used for a great number of other species. Microsatellites have also been identified in the chloroplast genome (Marshall et al. 2002; Stoeckert et al. 1998; Vendramin et al. 1996; Vendramin and Ziegenhagen 1997), which is paternally inherited in most gymnosperms, offering opportunities for comparative studies of pollen- versus seed-mediated gene flow (Kent and Richardson 1997) including estimation of pollen contamination levels in seed orchards from background sources (Stoeckert and Newton 2002).

1.8 Parentage Analysis

The development of pedigree reconstruction methods (reviewed by Jones and Ardren 2003) made it possible to determine the genetic relationships among individuals using only molecular markers.

The earliest and conceptually simplest method is exclusion (Hedrick 2005), whereby, based on Mendelian inheritance, incompatibilities between parental and offspring genotypes are used to reject particular parent-offspring hypotheses. Exclusion is not ideal, because issues such as genotyping errors, null alleles, and mutations may cause false exclusions (type II error; see Kalinowski et al. 2007; Oddou-Muratorio et al. 2003 for details; Van de Putte et al. 2006). (Under strict exclusion criteria, a single mismatch is sufficient to exclude a candidate parent regardless of how many loci are scored.) Several statistical methods have therefore been proposed to specify the number of mismatches necessary for a justifiable exclusion

(e.g., Marshall et al. 1998; Sancristobal and Chevalet 1997), making the method more robust (Kalinowski et al. 2007).

Categorical and fractional allocation both assign progeny to non-excluded parents based on likelihood scores derived from their genotypes, and differ only in that the former assigns entire offspring to a particular male (Meagher and Thompson (1986); further developed by Marshall et al. (1998)) while the latter splits each offspring and assigns some fraction, between 0 and 100%, among all compatible males (Devlin et al. 1988).

Parental genotypes can also be reconstructed from progeny arrays of full- and half-sibs, provided that multilocus genotypes for the progeny and one parent are known (Jones 2001; Jones and Avise 1997). This approach may be particularly useful in situations when an entire pool of candidate parents cannot be sampled and, moreover, the software program developed by Jones (2001) is capable of determining the minimum number of parents necessary to explain the given progeny array.

It should be pointed out, however, that the accuracy of all these methods is constrained by a variety of the issues described above, which affect most data and can markedly influence the biological conclusions of a study. A comprehensive review on genotyping errors, their causes, consequences, and suggested minimization strategies, was provided by Pompanon et al. (2005).

In seed orchard research, the increasing availability of highly informative genetic markers and the development of sophisticated pedigree reconstruction methods have provided seed orchard managers and research scientists with opportunities to study seed orchard mating system dynamics, to evaluate the effectiveness of management practices, and verify parental affiliation of ramets with great accuracy.

1.9 Objectives

This thesis is designed to investigate the population dynamics of seed orchards in multiple species (western larch, western redcedar, Douglas-fir, and lodgepole pine) utilizing microsatellite DNA-based pedigree reconstruction, with the following specific objectives:

- Estimate the variation in parental reproductive success and assess its impact on seed crop genetic quality (genetic gain and diversity) (Chapters 2, 3, 4, 5, and 6).
- Compare genetic data obtained from offspring of known maternal parents (family array; partial pedigree reconstruction) versus data obtained from analyzing bulk seed (full pedigree reconstruction), where both maternal and paternal parentages are unknown (Chapter 5).
- Evaluate the reliability of several simplified reproductive investment assessment methods (Woods 2005) developed for predicting actual parental reproductive success (Chapter 6).
- Develop an optimization protocol to maximize seed crop genetic gain at any desired level of genetic diversity that simultaneously considers parental reproductive output, including variation in seed germination capacity, relatedness among parents, and inbreeding (Chapters 7 and 8).

2. Pedigree and Mating System Analyses in a Western Larch (*Larix occidentalis* Nutt.) Seed Orchard

2.1 Introduction

The *Larix* genus is the only deciduous conifer in the *Pinaceae* family, comprising 10 recognized species that encircle the northern hemisphere with variable economic and ecological importance throughout their range (Schmidt 1995; Schmidt and McDonald 1995). Western larch (*Larix occidentalis* Nutt.) is commonly found in valley bottoms and along steep slopes of montane forests in the Upper Columbia River Basin of western North America (Jaquish and El-Kassaby 1998). While it is generally considered to be among the largest of North America's larches, it is the most shade intolerant conifer in its region. However, it attains its large size through rapid growth during early years, thus outgrowing its associates until it gets overtopped by slowly growing and shade tolerant species (Schmidt et al. 1976). The importance of western larch has been recognized not only for its commercial wood products, but also for its outstanding non-timber values such as aesthetics, wildlife habitat, outdoor recreation, and watershed protection (Schmidt and Shearer 1995).

Population genetic studies of western larch and related species began as early as 1980s where measures of genetic variation (Fins and Seeb 1986; Jaquish and El-Kassaby 1998; Khasa et al. 2000; Semerikov and Lascoux 1999; Semerikov and Matveev 1995; Semerikov et al. 1999) and mating system (El-Kassaby and Jaquish 1996) were reported to aid the species' gene resource management, including selective breeding and conservation. During the mid-1990s, the British Columbia Ministry of Forests, Mines and Lands embarked on a western larch genetic improvement program starting with selection of superior phenotypes, breeding and testing, and the establishment of production populations (seed orchards) for seed production (Jaquish et al. 1995).

The genetic quality of the resultant seed crops is of importance since seeds are the conduit for passing breeders' efforts to future forests. Parental male (pollen) and female (seed) gametic contribution, level of inbreeding, and pollen contamination from outside sources collectively determine the seed crops' genetic quality (Stoeckert et al. 2004). The availability of highly informative DNA markers such as microsatellites (Vendramin and Hansen 2005) and the development of sophisticated pedigree reconstruction methods (Jones and Ardren 2003)

made it possible to reconstruct the genetic relationships among seeds/individuals derived from natural open-pollination and classify them into their respective half-sib and full-sib families.

In this study, an example of the utilization of microsatellite DNA markers and a combination of pedigree reconstruction-mating system analyses is provided for estimating variation in parental reproductive success and its impact on genetic gain and diversity estimates in a western larch seed orchard. The level of inbreeding (selfing) and pollen contamination is quantified and discussed.

2.2 Materials and Methods

2.2.1 Seed orchard, plant material, DNA extraction and fingerprinting

The material for this study was provided by a clonal, first-generation western larch seed orchard located near Vernon, B.C., Canada (the Kalamalka Forestry Centre: altitude 480 m, latitude 50°14' N, longitude 119°16' E), established in 1989 from phenotypic selection within south-eastern British Columbia's natural populations. It is one of two physically adjacent, but genetically distinct (no common parents) western larch seed orchards established by B.C. Ministry of Forests, Mines and Lands to provide genetically improved seed to the Nelson (<1,300 m) and East Kootenay (800–1,500 m) seed production units (Woods 2008). From the neighboring orchard it is separated by a road (~ 8m wide) and a row of black cottonwood trees on one side.

The orchard population consists of 1,280 ramets representing 41 parents (average: 31 ramets/parent) and parental breeding values (volume at age 60) were predicted from two series of 10-year-old progeny test trials located within the parental natural range (B. Jaquish, B.C. Ministry of Forests, Mines and Lands, personal communications). Parents and ramets within parents were arranged on the orchard's grid following the permuted neighbourhood design (Bell and Fletcher 1978).

Needle samples were collected from all 41 parents along with wind-pollinated seed from a subset of 14 parents ($N = 40$ /parent). Total DNA from both needles (parents) or germinating embryos (progeny) was isolated using a modified CTAB method (Doyle and Doyle 1990; Khalsa et al. 2000). All samples, parents and embryos (half-sib families), were genotyped using a set of 12 microsatellite markers for pedigree reconstruction and mating system

analyses. Three of the studied markers used (UAKLly_10A, UAKLly_13-1, and UAKLly_13-2) were developed and their Mendelian mode of inheritance was demonstrated by Khasa et al. (2000). The remaining nine markers were developed by Chen et al. (2009) (Appendix A) and their mode of inheritance is illustrated using the sampled 14 half-sib families in Table 2.1.

Table 2.1 Mode of inheritance for western larch microsatellite loci using wind-pollinated seed from 14 parents. G -test for the goodness-of-fit to the 1:1 segregation ratio for pooled data (G_P) and test of heterogeneity among multiple parents sharing the same heterozygous genotype (G_H).

Locus	Mother tree genotype	Number of parents	Observed segregation	G -test	
				G_P	G_H
UBCLXtet_1-22	202/240	2	52:72	3.23	3.25
	220/240	2	73:49	4.75*	6.24*
UBCLXtet_21	142/158	2	85:49	9.79**	10.54**
UBCLXtet_32	188/196	5	218:156	10.32**	20.86***
	168/188	3	68:142	26.64***	30.21***
UBCLXtet_2-12	250/266	3	100:125	2.78	3.36
	Null/250	2	40:31	1.14	9.42**
	Null/260	2	58:69	0.95	1.76
UBCLXtet_2-11	152/164	2	76:45	8.03**	8.31*
	152/170	3	147:87	15.55***	15.6**
UBCLX_1-10	283/287	5	165:157	0.25	4.18
	277/287	3	52:118	26.31***	27.93***
	277/283	2	42:80	12.03***	13.62**
UBCLXA4_1	196/206	3	79:114	6.38*	23.51***
UBCLXdi_21	Null/315	5	124:142	1.21	1.44
	Null/335	3	75:104	4.72*	9.15*
UBCLXdi_16	Null/240	2	32:23	1.48	1.48

*significant at $p < 0.05$; ** $0.005 < p < 0.05$; *** $p < 0.005$

PCR amplifications were performed in 10 μ L total volumes consisting of 1X PCR buffer (10 mM Tris, 1.5 mM $MgCl_2$, 50 mM KCl, pH 8.3) (Roche, Laval, Quebec), 50 ng of total genomic DNA, 0.3 pmol of M13 IRD-labeled primer, 5 pmol of each forward and reverse primers, 0.1 unit of *Taq* DNA polymerase (Roche), and 1 mM of dNTPs (0.25 mM each) using an MJ research PTC-100 thermal cycler (MJ Research, Inc., Watertown, MA). PCR

program comprised of 5 min of pre-incubation at 94°C, 25 cycles of 1 min at 94°C, 1 min at respective annealing temperature (Chen et al. 2009; Appendix A) and 1 min of extension at 72°C, and 3 min of incubation at 72°C as the final step of extension. PCR products were then denatured by adding 2 µL of loading dye (100% formamide, 1 mg/mL pararosaniline basic red 9) and electrophoretically separated using Li-Cor 4200 (Li-Cor Inc., Lincoln, NE) automated sequencer with 5% polyacrylamide (Long Ranger™) gels. No multiplexing was attempted during this study.

2.2.2 Mode of inheritance of microsatellite loci

The wind-pollinated seed collected from the 14 parents (maternal half-sib families with a sample size of 40 seeds per parent) were used to illustrate the mode of inheritance for the 9 SSR markers developed by Chen et al. (2009). The mode of inheritance of segregating loci was restricted to heterozygous parents' progeny array showing segregation at any particular locus, after excluding offspring displaying the same genotype as the respective mother. Individual parents' segregation analyses for heterozygous loci were performed using the *G*-test of independence (Sokal and Rohlf 1981) (not shown). When multiple trees shared the same genotype, the replicated *G*-test of goodness-of-fit (G_P) was conducted on the pooled segregation data (Sokal and Rohlf 1981). Segregation data from multiple trees were also tested for heterogeneity using *G*-heterogeneity test (G_H) (Sokal and Rohlf 1981). In cases where significant G_P or G_H were observed, the cause of either the observed bias towards one allele or heterogeneity among segregating trees were investigated by comparing segregation mode at individual tree level.

2.2.3 Paternity and mating system analyses

The revised likelihood-based paternity inference program CERVUS 3.0.3 (Kalinowski et al. 2007; Marshall et al. 1998) was used to assign the pollen donors for the 14 maternal seed parents' embryo pool. Since the maternal parents' multilocus genotypes were known, this paternity assignment was a partial pedigree reconstruction. The multilocus mixed-mating model of Ritland (2002) was used to estimate mating system parameters, including single-locus (t_s) and multilocus (t_m) outcrossing rates and multilocus correlated matings ($r_{p(m)}$) using the expectation-maximization (EM) procedure of the computer program MLTR 3.1 (Ritland

2002). Standard errors for mating system parameters were obtained from the construction of 1,000 bootstrap replicates.

2.2.4 Assessment of genetic gain and diversity

Average breeding value (genetic gain: ΔG) and male effective population size ($\hat{\sigma}N_e$) were estimated as

$$\Delta G = \sum_{i=1}^N BV_i m_i \quad [2.1]$$

and

$$\hat{\sigma}N_e = \frac{1}{\sum_{i=1}^N m_i^2} \quad [2.2]$$

where N is the orchard's census number ($N = 41$) and BV_i and m_i denote the breeding value and the male gametic contribution of parent i , respectively ($0 \leq m_i \leq 1$, and $\sum_{i=1}^N m_i = 1$). Pedigree-based estimates of ΔG and $\hat{\sigma}N_e$ were compared to those derived from parental representation, i.e., the proportion of ramets representing each parent. Pollen contamination was not considered into $\hat{\sigma}N_e$ calculation.

2.3 Results and Discussion

2.3.1 Microsatellite loci mode of inheritance

Segregation analysis of the 14 wind-pollinated progeny produced excessive departure from the expected 1:1 Mendelian ratio for most of the tested SSRs. This higher than expected segregation distortion could be due to inbreeding depression (Tani et al. 2004), although both pedigree and mating system analyses indicated that inbreeding was minimal (see below), apparent heterozygosity, or a simple mutation occurring within the binding site for a DNA marker (Smith et al. 1997). Genetic analysis of PCR-based markers depends on amplifying the length variability flanking with oligoprimers. Any mutation occurring within the DNA complementary to the oligoprimers may inhibit or completely prevent their binding, leading to an absence of PCR product. These null alleles will not necessarily be recognized when there is a product from the other homologue, and this may lead to an underestimate of marker heterozygosity and to apparent incompatibility of genotypes within a family. Null alleles have long been known for protein polymorphisms, VNTR markers (Chakraborty et al. 1992)

including microsatellites (Liewlaksaneeyanawin et al. 2002) and have been recognized, along with population subdivision, as a major factor in deficiency of observed heterozygosity, compared with that expected on the basis of Hardy-Weinberg equilibrium (Callen et al. 1993). Microsatellite markers' null alleles were frequently observed in coniferous tree species with reported within-population frequencies as high as 35% (radiata pine: Fisher et al. 1998; sugi: Moriguchi et al. 2003; Norway spruce: Yazdani et al. 2003). Null alleles also varied among populations within species, indicating that their presence is population-specific rather than species-pandemic (Tani et al. 2004). Generally, the frequent presence of segregating null alleles found among microsatellites calls for some caution when untested loci are used as genetic markers because they can underestimate heterozygosity and lead to false parentage assignment (Liewlaksaneeyanawin et al. 2002; Pemberton et al. 1995).

In this study, parental genotypes were inferred from the segregation pattern of their offspring for the segregation analysis. The nine tested loci produced 17 segregating allelic combinations, with six following the expected 1:1 segregation ratio and 11 significantly deviating from the expected 1:1 ratio or the G -heterogeneity test (Table 2.1). Homozygous parents' offspring should not segregate and their banding pattern should be uniform and resembling the parental type. However, genotypes of four seemingly homozygous parents were determined to be heterozygous with null alleles (UBCLXtet_2-12: null/250 and null/260, UBCLXdi_21: null/315 and null/335, and UBCLXdi_16: null/240) (Table 2.1), as their offspring segregated and produced an excess of homozygotes as well as genotypes with no products (Table 2.1). For example, at locus UBCLXdi_16, two parents produced a single band at 240 bp, thus they were originally genotyped as 240/240; however, the appearance of two other alleles (238 and 248 bp) from the male side, which resulted into one homozygous genotype of either 238/238 or 248/248, was also indicative of maternal null allele contribution, thus confirming the heterozygous genotype of the mother tree. This consistent pattern indicated that the parental genotypes should have been heterozygous for null/240. Repeated scoring of these two individuals' offspring supported this interpretation and yielded a non-significant deviation from the 1:1 segregation ratio (G -test values of 1.00 ($p = 0.32$) and 0.48 ($p = 0.49$)) (not shown). Furthermore, the G_P and G_H goodness-of-fit tests of the pooled data of these two parents were non-significant (Table 2.1). The same scenario was observed for the null/250 and null/260 genotypes at locus UBCLXtet_2-12 and null/315

genotype at locus UBCLXdi_21 (Table 2.1). Segregation of the pooled data at UBCLXtet_2-12: null/250 followed the expected 1:1 ratio; however, significant G_H was observed, indicating the presence of different segregation patterns among mother trees (Table 2.1). The majority of allelic combination significantly deviated from the 1:1 ratio, in both pooled (G_P) and heterogeneity (G_H) tests (Table 2.1). In most cases, deviations were caused by the unexpectedly higher frequency of one allele over the other (e.g., allele 152 bp at UBCLXtet_2-11) (Table 2.1).

2.3.2 Mating system analysis

High single- ($t_s = 85 \pm 0.023$) and multi-locus ($t_m = 97 \pm 0.013$) outcrossing rates were obtained from the 14-family array, indicating that outcrossing is predominant in this population. Similar outcrossing rate estimates were observed in natural populations of the same species (El-Kassaby and Jaquish 1996). Multi-locus correlated paternity ((r_{pm}) ; correlation of outcross paternity within progeny arrays) was low (0.074 ± 0.013), indicating the presence of multiple parents participating in pollination, and was similar to those reported in the species' natural populations (El-Kassaby and Jaquish 1996). While the mating system analysis is informative, it falls short of providing an estimate of pollen migration to the orchard (contamination). El-Kassaby and Ritland (1986b) highlighted the relationship between outcrossing rate and gene flow and demonstrated that the higher the pollen migration the higher the observed outcrossing, thus it should be stated that most seed orchards' outcrossing rates are inflated by gene flow. The multi-locus correlated paternity estimate also provided an indication of the correlation of outcross paternity within progeny arrays; however, it did not assess which males contributed to this estimate or the different male-female pollination success.

2.3.3 Paternity analysis

A total of 430 out of 551 (78%) embryos were assigned to one of the 41 candidate male parents (Table 2.2). Additionally, a subset of 46 out of the assigned 430 embryos were found to be the product of selfing (mating within a ramet or among ramets of the same parent) resulting into an estimate of outcrossing rate of 91.7%, an intermediate value between the observed single- ($t_s = 85 \pm 0.023$) and multi-locus ($t_m = 97 \pm 0.013$) outcrossing rates obtained from the mating system analysis. The observed variability between the two

estimation methods (i.e., paternity vs. multilocus mixed-mating model) reflects the difference between the “exact” nature of paternity analysis and the estimation protocol of the mating system model that does not account for the presence of null alleles (i.e., offspring produced from gametes with null alleles are considered as the product of outcrossing even if they originate from the same parent).

The remaining unassigned 121 offspring (22%) could be the product of gene flow from the neighbouring orchard or their genotypes were not informative enough for assignment to any of the orchard’s 41 candidate parents (i.e., type II error (false exclusion) (Kalinowski et al. 2007; Oddou-Muratorio et al. 2003; Vandeputte et al. 2006)) or a combination of the two. It is also possible that some of the assigned embryos could be the product of type I error (false assignment) (Kalinowski et al. 2007; Oddou-Muratorio et al. 2003; Vandeputte et al. 2006); however, since the set of embryos used for the mode of inheritance analyses were the same as those used in paternity analysis, it is unlikely that false assignment had a significant contribution. The number of unassigned embryos per female parent varied between two (family #25) and 14 (families #3, 39, and 40), which could be due to reproductive phenology asynchrony between the orchard’s female parents (ovule receptivity) and pollen from background sources and/or due to the lack of informative nature of the embryos’ genotypes.

The pedigree reconstruction produced 221 full-sib families nested within the 14 maternal half-sib families with a range of 12 (family #40) to 22 (family #30) (Table 2.2), indicating the effectiveness of pollination within the studied seed orchard (i.e., availability of many male parents for pollinating receptive females).

Male half-sib family sizes ranged between 0 (male families #29, 31, 37, and 38) and 46 (family #7) (Figure 2.1), indicating the effectiveness of seed-cones to act as pollen traps for providing an accurate assessment of pollination. The observed variation among male half-sib family sizes could also be caused by temporal reproductive separation (Figure 2.1).

Table 2.2 Summary of paternity analysis of 14 western larch half-sib families with known maternal parentage at 95% confidence levels.

Parent ID	Number of offspring			Number of full-sib groups
	Analyzed	Assigned at 95%	Unassigned	
3	40	26	14	15
4	37	31	6	17
7	40	29	11	13
9	40	30	10	13
14	38	31	7	18
19	40	35	5	15
20	40	34	6	15
21	40	29	11	15
24	39	31	8	18
25	39	37	2	21
30	40	34	6	22
34	39	32	7	13
39	40	26	14	14
40	39	25	14	12
Total	551	430	121	221

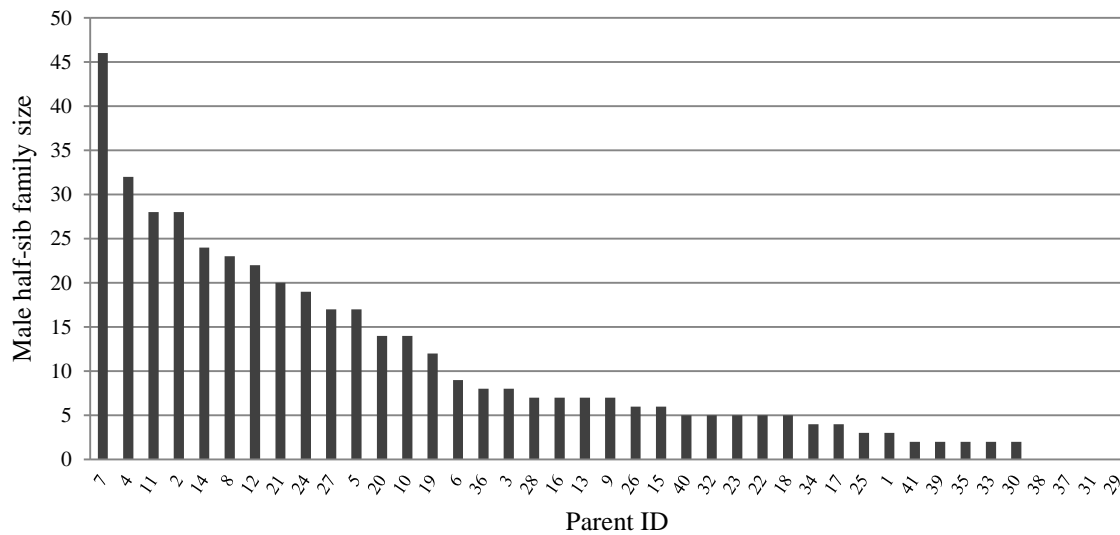


Figure 2.1 Western larch male half-sib family size listed in descending order.

The distribution of full-sib families and their size within and among parents ranged from 1 to 8, with a large proportion of them having a small size (from 1 to 3) (Figure 2.2). This is expected considering the compactness of pollination season in interior British Columbia that

is characterized by its abbreviated spring period caused by increased rate of heat-sum accumulation over very short period of time (a sudden shift from winter to summer seasons; El-Kassaby, personal observation) and the tested seed sample size of 41. Approximately 50, 75, and 90% of successful within-orchard pollen was produced by the top 8 (25%), 15 (37%), and 25 (61%) parents, respectively. Conversely, the remaining 10% was produced by the least successful 16 (39%) parents. The paternity analysis confirmed the presence of male fertility variation within the studied seed orchard as reported previously in sugi (Moriguchi et al. 2004), Douglas-fir (Slavov et al. 2005a), and wild cherry (Mariette et al. 2007), a difficult assessment to make based on pollen-cone count surveys which assume that reproductive investment is equal to reproductive success (Kang 2000; Kang and Lindgren 1998).

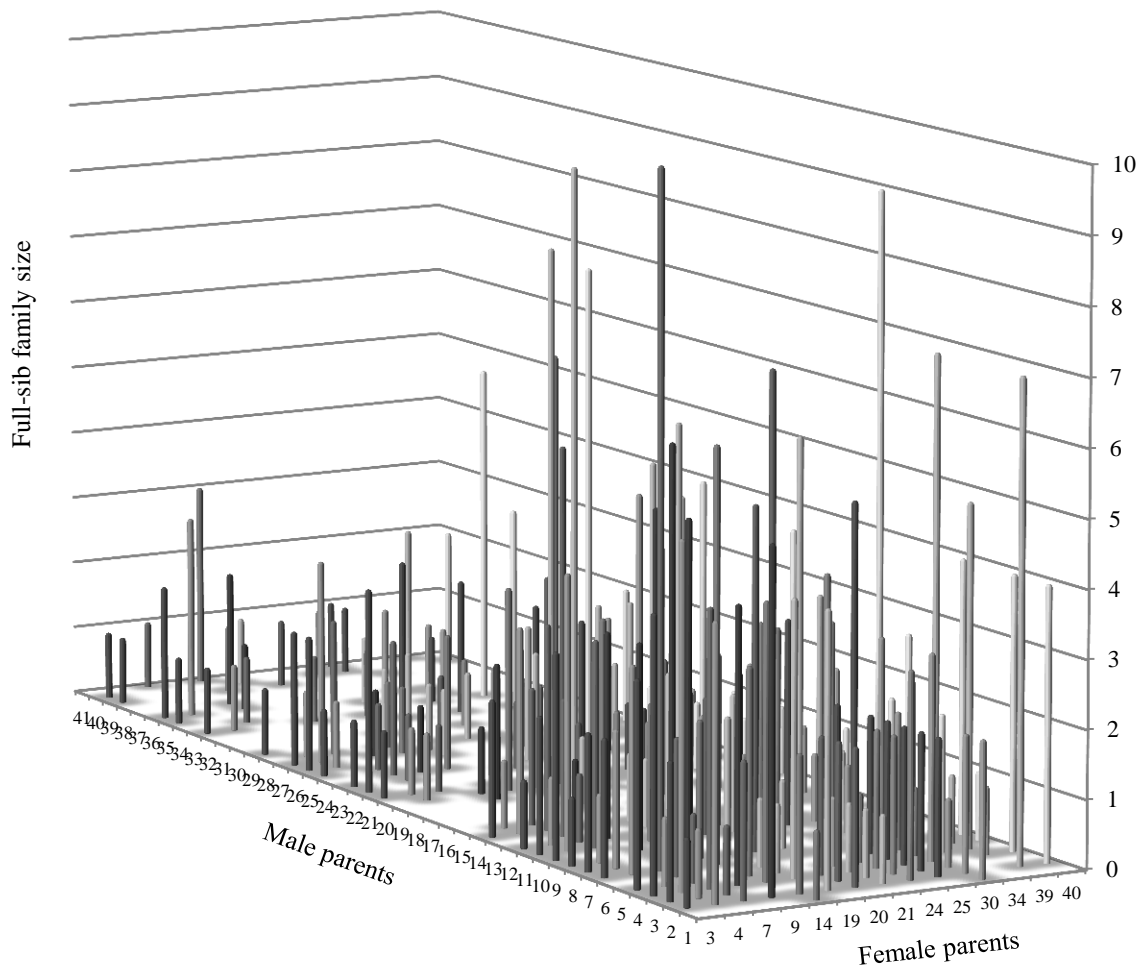


Figure 2.2 The distribution of 221 full-sib families from 14 wind-pollinated western larch seed donors.

The effective number of male parents (\hat{N}_e) was estimated to be 21, a value significantly lower than the orchard's census number of 41 (or 33 if variation in parental representation was considered). Similar results were also reported by Slavov et al. (2005a) who studied male reproductive success in a Douglas-fir seed orchard. The number of seeds fathered in their samples from a test block ranged from 1 to 17, leading to a substantial decrease in the effective number of parents from 40 (the seed orchard's census number) to 20. The impact of the reproductive success variation on genetic gain was not substantial, yielding an estimate of 15.4 as opposed to 15.9 expected from parental representation, respectively. This slight difference can be attributed to a relatively small range of breeding values (10-26) among the orchard's parental population.

Knowledge of pollen dispersal within a seed orchard is essential for the development of new orchard designs and effective management practices (Adams and Birkes 1989; Wheeler et al. 1993). The identification of male reproductive success is considerably more difficult than of its female counterpart (Devlin and Ellstrand 1990). The measure of pollen movement is often tracked by the movement of pollinators, marker dyes (Linhart et al. 1987), or marked pollen (Greenwood 1986; Levin and Kerster 1969). These methods become impractical when large numbers of different pollinators are involved, the host plants are large, or basically the number of the pollen sources is beyond manageable. Moreover, pollen dispersal derived from a direct measurement of physical pollen movement may not always reflect actual fertilization and gene flow (reviewed in Dow and Ashley 1998); therefore, pollen dynamics/dispersal is best assessed using molecular genetic markers (Brown 1989). With the advantage of the high polymorphism and co-dominant mode of inheritance, microsatellite markers provide a high exclusionary power in paternity analyses.

2.4 Conclusion

This study provides an example of the usefulness of microsatellite markers in pedigree and mating system analyses and quantifies the impact of male reproductive success variation on seed crops genetic parameters; namely, male genetic gain and effective number of male parents. The reported genetic evaluation of the studied seed orchard's mating dynamics would not have been possible without the use of highly informative genetic markers and the recent theoretical development of paternity analysis.

3. Mating System Analyses in a Western Redcedar (*Thuja plicata* Donn. ex D. Don) Seed Orchard

3.1 Introduction

Western redcedar (*Thuja plicata* Donn. ex D. Don) is a long-lived, monoecious conifer species native to the coastal and interior regions of western North America (Minore 1990). The species' genetic variation has been studied extensively with molecular genetic markers; namely, isozymes (Copes 1981; El-Kassaby et al. 1994; Yeh 1988), restriction fragment length polymorphism (RFLP) (Glaubitz et al. 2000), and microsatellites (SSRs) (O'Connell et al. 2008; O'Connell et al. 2004), with reported levels ranging from no or very low (isoenzymes and RFLP) to substantial (SSRs). The initial observed lack of variation, unexpected in a long-lived conifer, was attributed to a combination of inbreeding (El-Kassaby et al. 1994) and one or more population bottlenecks that the species had undergone during the most recent glacial period (Critchfield 1984; O'Connell et al. 2008), leaving it genetically depauperate, and phenotypic plasticity was considered to be the species' evolutionary strategy for survival rather than genetic variation (El-Kassaby 1999).

While many plant species have evolved as self-pollinators, many others have developed a variety of pre- and/or post-pollination mechanisms to minimize or avoid selfing, such as self-incompatibility in angiosperms (reviewed by de Nettancourt (1997)), dioeciousness or physical separation of male and female strobili within a tree in monoecious species (Park and Fowler 1984), asynchrony of male and female reproductive phenology (El-Kassaby et al. 1984; Erickson and Adams 1990; Eriksson et al. 1973), or polyembryony (Sorensen 1982). It is interesting to note that the pollination mechanism of western redcedar appears to be predisposed to selfing (Colangeli and Owens 1990) as within-tree pollen shed and female cone receptivity are synchronous, with distinctive pollination drops secreted by female cones' nucellar tips, facilitating the immediate delivery of pollen to the micropyle regardless of its source, i.e. whether it is self or outcross (Owens et al. 1990). However, unlike other conifers, e.g. Douglas-fir (Owens et al. 1981b) or Engelmann spruce (Owens et al. 1987), western redcedar cannot effectively accumulate pollen in the micropylar region over a period of time, as the ovule ceases to be receptive for any pollen arriving after the delivery of the initial pollen (Owens and Molder 1980). Furthermore, while the species has archegonial

polyembryony, no preferential selection favoring outcross pollen was observed in western redcedar (O'Connell and Ritland 2005).

Although conifers are predominantly outcrossing (mean selfing = 16.5% for 52 different species (O'Connell 2003)), species of the genus *Thuja* display a relatively high selfing (e.g. Chinese arborvitae (25%; Xie et al. 1991) or northern whitecedar (49%; Perry and Knowles 1990; Lamy et al. 1999)). In natural populations, higher selfing can be associated with a variety of factors such as clonal structure (O'Connell et al. 2008), tree size (O'Connell et al. 2004), or crown position (Chaisurisri et al. 1994; El-Kassaby et al. 1986) and population-level selfing rates ranging from 22 to 36% were estimated for western redcedar (O'Connell et al. 2001). However, a surprisingly high selfing of 68% – in fact, one of the highest ever reported for a conifer species – was estimated in a seed orchard (El-Kassaby et al. 1994), a population specifically designed and intensively managed to promote outcrossing where ramets are randomized and trees are pruned to ensure effective pollen flow among trees.

Evidence for western redcedar's propensity to self was provided by Owens et al. (1990) based on a study of the species' reproductive attributes: pollen germination, fertilization, number of filled seeds per cone, and pollination and seed efficiencies were all higher in self- than cross-pollinated individuals, indicating that the species does not suffer from inbreeding depression, at least for these traits. Almost no inbreeding depression was detected for seed production and seedlings' early nursery growth between western redcedar's selfed and outcrossed families, although it reached about 3 to 10% at circa 2 m in height, with very little subsequent development (Russell et al. 2003).

The species' predisposition and high tolerance of selfing play a negative role in seed orchards where inbreeding is undesirable. Pollen augmentation (a.k.a., supplemental mass pollination (Wakeley et al. 1966)), which involves applying pollen from known external and/or internal unrelated and desirable pollen donors to receptive female strobili, is commonly practiced to promote outcrossing. However, top-grafting, a management practice where scions of specific parents are grafted in crown tops of other, unrelated parents with the aim to induce early flowering in the grafts and thence shorten breeding cycles (Bramlett and Burris 1995; Goadin et al. 2000), has not been evaluated in this regard yet although one can assume that it has also a positive effect on outcrossing, as it substantially reduces space for self-fertilization within a tree crown by splitting it into two different genetic entries.

The objectives of this study are to detect mislabeled ramets and alien genotypes, determine male reproductive success variation, estimate gene flow to the orchard (pollen contamination), and estimate the selfing rate after the implementation of top-grafting treatment to one third of ramets in a western redcedar seed orchard.

3.2 Materials and Methods

3.2.1 Study seed orchard

Seed and foliar samples from a first generation, clonal western redcedar seed orchard located on the Saanich Peninsula, southern Vancouver Island, British Columbia (48°35' N, 123°24' W, 50 m) provided material for this study. The orchard was established in 2000 and the parental population at the time of seed collection (2008 seed crop) consisted of 27 parents (338 ramets) planted on an 18 × 30 grid with 3 and 6 meters between trees within rows and between rows, respectively. In order to promote outcrossing and thereby the overall genetic quality of seed crops, approximately one third of the orchard (98 out of 338 ramets) was top-grafted using scions from other parents in this orchard; thus, the top-grafted tress consisted of two different genotypes from the same orchard's population. While the orchard is generally located in an area devoid of western redcedar's background pollen, it is adjacent to another western redcedar orchard with 96 parents, representing a significant source of potential gene flow. It should be noted that 11 replicated parents are common in both orchards, a possible source of cryptic gene flow.

Seed were collected from a subset of ramets representing 15 of the 27 parents, processed individually on a maternal half-sib family basis, and stored at +4°C. Germination was conducted without pre-treatment at ~25°C and embryos were excised from their seed coats after reaching circa 3 cm in length. Fresh foliage was collected in May 2011 from all ramets present in the studied orchard (128 ramets including all interstocks, plus additional two top-grafts), as well as from five ramets in the neighboring orchard representing the external pollen donor. Since the orchard was rogued in fall 2009, samples from 14 missing parents were collected from a clonal archive at Cowichan Lake research station also in May 2011. Both the germinated embryos and foliar tissue were stored at -80°C until further use.

3.2.2 DNA extraction, PCR and genotyping

DNA from both embryonic and foliar tissue was isolated using a modified protocol by Doyle and Doyle (1990) and a set of eight polymorphic nuclear microsatellites (Table 3.1) were amplified using primer pairs designed by O'Connell and Ritland (2000). Each reaction mix (10 uL) comprised of 1X PCR buffer (10 mM Tris, 1.5 mM MgCl₂, 50 mM KCl, pH 8.3; Roche), 1 mM dNTPs (0.25 mM each), 5 pmol of each forward and reverse primers, 0.3 pmol of M13 IRD-labeled primer, 1 unit of Taq DNA polymerase (Roche), and 100 ng of genomic DNA. PCR was performed in a GeneAmp 9700 thermal cycler (Perkin-Elmer, Foster City, CA) following optimization by O'Connell and Ritland (2000). After PCR, 3 uL of loading dye (100% formamide, 1 mg per mL pararosaniline basic red 9) were added to each reaction.

DNA amplicons were electrophoretically separated in a 6% polyacrylamide (Long Ranger™) gel using LiCor 4300 (LiCor Inc., Lincoln, NE) automated sequencer and gel images were genotyped in software program SAGA™ (LiCor Inc., Lincoln, NE). Allele sizes were determined with the aid of 50–350 base pairs sizing standards (LiCor Inc., Lincoln, NE) along with maternal genotypes accompanying each half-sib family (2–4 families per gel).

3.2.3 Paternity analysis

A total of 771 offspring from 15 maternal seed-producing parents were analyzed (average 51; range 2 to 121). Male parentage was assigned to each offspring using the software program CERVUS 3.0.3 (Kalinowski et al. 2007) where all known parents were set as candidate fathers. The resolving power of the eight microsatellite loci and critical delta values (delta is the difference in log-likelihood (LOD) scores between the most likely candidate parent and the second most likely candidate parent where the LOD score is the likelihood that the candidate parent is the true parent divided by the likelihood that it is not the true parent) were computed based on the assignment of 1,000,000 simulated offspring. Minimum number of loci typed on each individual was set to 4 and mistyping error rate of 0.03, estimated by the simulation analysis, was allowed as error rate in the corrected likelihood calculations (Kalinowski et al. 2007). Delta score at 95% confidence level served as a criterion for assignment while all unassigned male gametes were considered to have originated from non-

orchard parents (i.e., alien ramets, parents in the neighboring orchard, or background pollen sources) or to be a product of insufficiently informative genotypic profiles.

3.3 Results and Discussion

3.3.1 Mislabeling and alien genotypes

Parental tree genotyping revealed the presence of three mislabeled ramets, all of which perfectly matched the multilocus genotype of another parent (#440) (Figure 3.1, indicated by M). Additionally, three alien genotypes matching none of the orchard's parents (one shown in Figure 3.1, indicated by C) were detected and assumed to be the result of overgrown rootstocks, increasing the number of candidate fathers in the paternity analysis to 30.

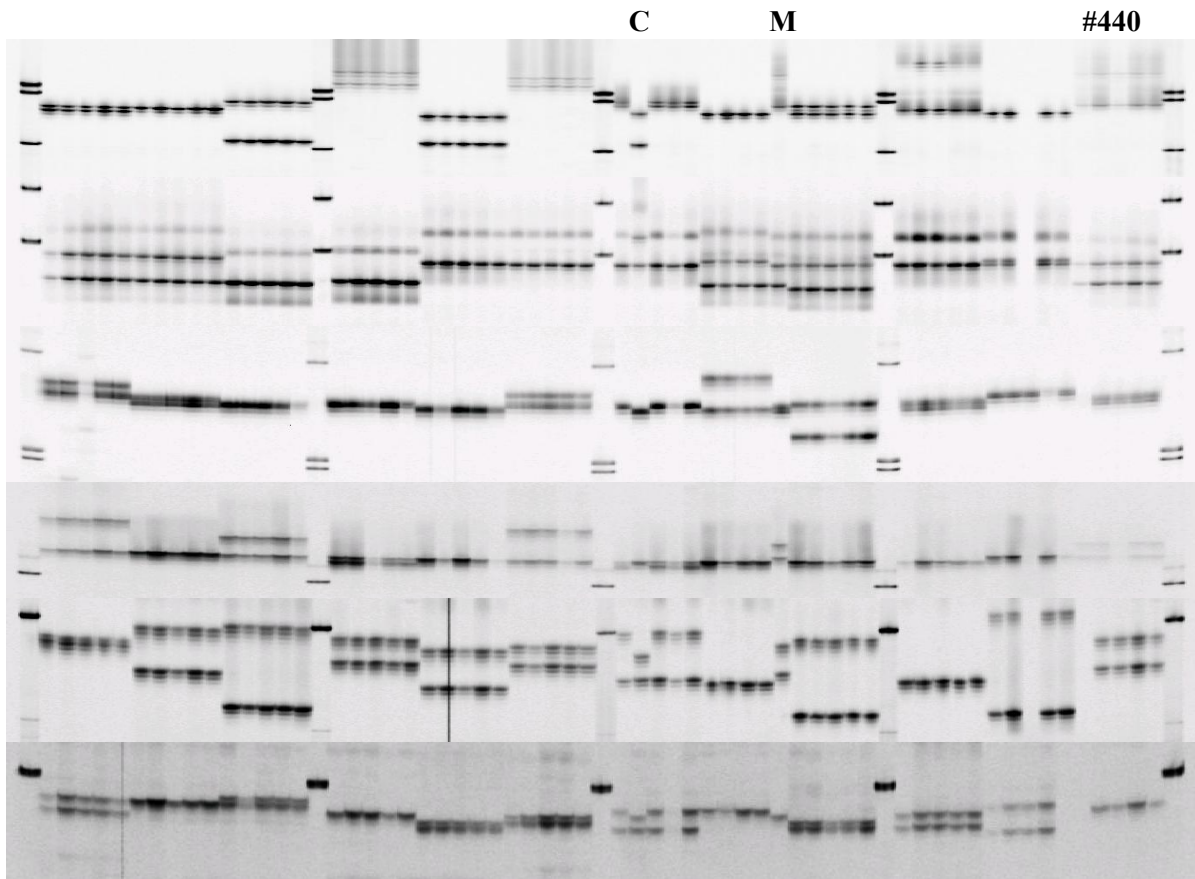


Figure 3.1 Verification of parental affiliation of ramets in the western redcedar seed orchard using microsatellite DNA markers. Genotypic profiles are shown for 5 ramets of 12 parents (total 60 ramets) on loci TP3, TP4, TP11, TP 7, TP9, and TP5 (top to bottom) and allelic standards surround each group of 3 parents (15 ramets). One incident of clone contamination (C) and mislabeling error (M) are indicated above.

3.3.2 Paternity analysis

A total of 449 out of the studied 771 individuals (58.2%) were successfully assigned to one of the 30 candidate fathers (i.e., within-seed orchard matings) while the remaining 322 (41.76%) were sired by outside pollen sources, representing a substantially higher than expected pollen contamination. One can argue that the discrimination power of the used panel of SSR loci may be insufficient to detect the true father; however, given their population allele frequencies (Table 3.1), CERVUS successfully assigned 99.3% of one million simulated offspring to the true father, making the expected erroneously unassigned offspring (i.e., false contamination) acceptably low. The high contamination observed in this study can, to a large extent, be explained by the proximity of the neighboring orchard, its larger area and population size (96 parents with 528 ramets), as well as the species' well-known synchronous and compact male and female reproductive phenology (El-Kassaby 1999; Owens and Molder 1980). All of these factors, individually or in concert, had created a reproductive overlap leading to the observed substantial gene flow. In fact, these successful contamination events are outcrossing events (El-Kassaby and Ritland 1986b). It is noteworthy to mention that ramets from the neighboring orchard, representing the 11 common parents, could have also contributed to cryptic gene flow; thus, the actual gene flow might in fact have been underestimated (Devlin and Ellstrand 1990). Interestingly the number of alleles per locus detected across all studied loci was much higher in the offspring than the parental population (average of 16.88 vs. 10.13), supporting the observed substantial gene flow (Table 3.1).

3.3.3 Male reproductive success variation

Male reproductive success was relatively well balanced in the studied seed crop (Figure 3.2). Excluding pollen contamination, the six most successful male parents (22%) sired 50% of seed, an imbalance similar to that anticipated based on a number of previously published works on other conifer species (e.g., white spruce: Schoen and Stewart (1986), Norway spruce: Xie and Knowles (1992), Japanese black pine: Goto et al. (2002), western larch: Funda et al. (2008) (Chapter 2), or Douglas-fir: Lai et al. (2010)).

Table 3.1 Characterization of parental and offspring populations in a western redcedar seed sample using eight nuclear microsatellite loci.

Locus	PIC	$NE-2P$	f_{null}	Parents			Offspring			
				k	H_O	H_E	k	H_O	H_E	$H-W$
TP1	0.617	0.564	0.089	8	0.548	0.649	14	0.546	0.655	***
TP3	0.806	0.332	0.007	10	0.677	0.795	18	0.820	0.826	*
TP4	0.631	0.547	0.016	8	0.710	0.725	12	0.638	0.660	ns
TP5	0.734	0.540	0.034	8	0.700	0.805	9	0.722	0.772	ns
TP7	0.621	0.555	0.075	6	0.613	0.634	11	0.552	0.651	***
TP8	0.850	0.267	0.178	17	0.581	0.899	25	0.603	0.862	-
TP9	0.896	0.192	0.016	16	0.903	0.916	35	0.875	0.904	ns
TP11	0.648	0.539	0.011	8	0.516	0.713	11	0.691	0.695	ns
Mean	0.726	0.0007		10.13		0.767	16.88		0.753	

PIC , polymorphic information content (Botstein et al. 1980); $NE-2P$, combined non-exclusion probability (second parent); f_{null} , frequency of null alleles; k , number of alleles; H_O , observed heterozygosity; H_E , expected heterozygosity.; $H-W$, excess of homozygotes relative to Hardy–Weinberg expectations (ns not significant, * significant at $\alpha = 0.05$, *** significant at $\alpha = 0.001$, - not calculated; significance levels include a Bonferroni correction).

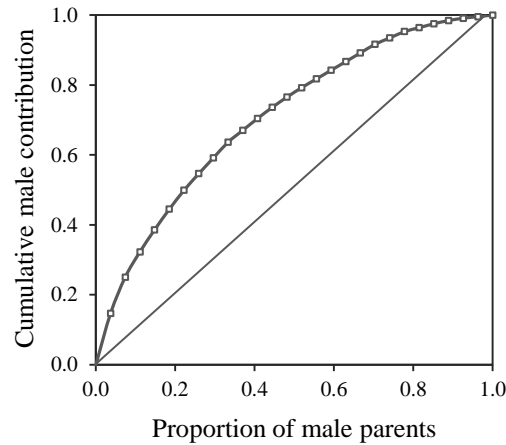


Figure 3.2 Cumulative reproductive success of 27 western redcedar male parents determined from microsatellite-DNA-based partial pedigree reconstruction. The diagonal line represents a hypothetical scenario with equal success among parents.

3.3.4 Selfing

A low selfing rate of 7.3% was estimated from the studied sample and is attributable to the high gene flow ($\sim 42\%$) and top-grafting that promoted within-tree outcrossing. The within tree outcrossing is effectively the result of first-pollination primacy hypothesis in conifers (Franklin 1974). Franklin (1974) hypothesized that the first pollen to land is the first to be

taken in with increased chances of fertilization success. This hypothesis was experimentally tested in Douglas-fir (Owens and Simpson 1982; Webber and Yeh 1987) and, more specifically, Webber and Yeh (1987) have shown, using genetically distinctive pollen sources that were applied in tandem with different time intervals, that the effectiveness of the second pollination is much lower than the first pollination, even when the time difference between the two applications is as close as 5 minutes. Thus, if unrelated pollen is captured by the pollen drop first, then the chances for outcross-fertilization increases. This is a very distinctive feature of the western redcedar active pollination mechanism where the pollen drop acts as a trap and then the trapped pollen is actively directed towards the ovule.

Contrary to the high selfing rate detected in western redcedar's natural (27-29%; (O'Connell et al. 2008; O'Connell et al. 2001)) and artificial (68%; El-Kassaby et al. 1994) populations, my estimate of 7.3% is exceedingly low. The species' predisposition to high selfing (Owens et al. 1990) and the lack of preferential selection against selfed seeds (O'Connell and Ritland 2005) made it the exception rather than the norm in terms of conifers (O'Connell 2003). In natural populations, western redcedar rarely exists in pure stands and often grows with other associates effectively restricting its pollen flow (Minore 1990), it has a very limited pollen dispersal due to the relatively small wing surface (Fowells 1965), and may occur in cohorts of genetically identical individuals (O'Connell et al. 2008); thus, in spite of the physical exchange of pollen an apparent outcrossing is in fact still selfing. However, in seed orchards, which are designed and managed to maintain inbreeding at minimum, the observed low selfing rate could be the product of randomization and top-grafting; besides, the role of pollen contamination should not be overlooked. At individual-parent (maternal) level, selfing markedly differed among trees and ranged from 0.0 to 31.7%, similar to differences reported by O'Connell et al. (2004) (0.0 to 78%), indicating that there is variation among western redcedar individuals in their propensity to self.

3.4 Conclusion

Although direct evidence supporting the hypothesis that top-grafting actively increased the outcrossing rate is not available, because comparison between top-grafted and control ramets was not conducted, the observed low mean selfing suggests that top-grafting may have positively contributed to its reduction and ultimately improved the overall genetic quality of the seed crop.

4. Female Reproductive Success Variation in a Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) Seed Orchard Revealed by Pedigree Reconstruction of a Bulk Seed Collection

4.1 Introduction

Reproductive success is the vehicle of genetic transmission between parental and offspring generations. In forest tree populations, reproductive success determines the breadth of genetic diversity of future generations and their resilience to unpredictable environmental contingencies. While female reproductive success seems to be easy to quantify through estimating the number of reproductively active individuals and their reproductive investment, such as the number of strobili, flowers, or seed (Bila et al. 1999; Hirayama et al. 2004; Kang 2000; Kang et al. 2004; Savolainen et al. 1993; Stoehr 2000; Wittwer et al. 1997), greater efforts have been made to studying the male component (Adams 1982; Burczyk 1996; Burczyk and Prat 1997; Goto et al. 2002; Gömöry et al. 2003; Moriguchi et al. 2004; Roberds et al. 1991; Schoen and Stewart 1986, 1987; Slavov et al. 2005a; Xie and Knowles 1992).

Virtually all family-array mating system and gene flow studies were based on drawing an equal number of seeds per maternal parent, thus only providing unbiased inferences pertaining to the male component (Burczyk 1996) while any female-related inferences suffered from this bias. For example, in a Douglas-fir seed orchard mating system study, Slavov et al. (2005a) used a “bulk” seed sample constituted by mixing an equal number of seed from each reproductively active ramet. This sampling method accounted for variation in parental representation but did not consider the within-parent (among-ramet) fecundity variation and, again, the derived inferences were valid for the male component of reproductive success only (Slavov et al. 2005a). In order for the female component of the reproductive success to be estimated without bias as well, a random sample of offspring (e.g. seed crop) reflecting actual female gametic contributions should be used.

In forest tree natural and experimental populations, female reproductive success variation has often been equated to variation in parental representation (i.e., number of ramets per parent) or variation in parental reproductive investment, approximated e.g. by seed-strobili or seed-cone production (Byram et al. 1986; El-Kassaby et al. 1989; Eriksson et al. 1973; Griffin

1982; Jonsson et al. 1976; Kang 2000; O'Reilly et al. 1982; Savolainen et al. 1993; Schmidting 1983; Schoen et al. 1986; Varnell et al. 1967; Ying et al. 1985). Furthermore, discrepancies between female reproductive investments at different developmental stages (such as the seed-cone production versus filled-seed yield (Chaisurisri and El-Kassaby 1993; El-Kassaby and Cook 1994; Reynolds and El-Kassaby 1990) and the filled-seed yield versus germinated seed (Chaisurisri et al. 1992; El-Kassaby and Edwards 1998; El-Kassaby et al. 1992; Krakowski and El-Kassaby 2005) were reported, demonstrating that female reproductive investment and success may not fully correspond.

In this study, microsatellite-DNA based full pedigree reconstruction was utilized to analyze a random sample of 801 seeds drawn from a bulk seed of a Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) seed orchard crop to obtain the first reported unbiased estimates of female reproductive success and mating system parameters, demonstrate the impact of female reproductive success variation on the extent of genetic diversity of the filial generation, and investigate if parental representation correlates with actual female reproductive success.

4.2 Materials and Methods

4.2.1 Seed orchard population and seed sampling

The same seed orchard and plant material as described in Lai et al. (2010) were used. During fall 2005, dormant vegetative buds were collected from the orchard's parental population and a random sample of 801 bulk seed with unknown male and female parentage, representing the entire orchard's 2005 seed crop, was drawn. All samples were immediately stored at -80°C and +4°C, respectively, until DNA extraction.

4.2.2 Microsatellite genotyping

Seeds were germinated and dissected to separate the diploid embryos and their corresponding haploid megagametophytes. DNA was extracted from all samples (vegetative buds and embryo-megagametophyte pairs) following the protocol of Doyle and Doyle (1990), amplified using the methods listed in Slavov et al. (2004), and scored with eight nuclear microsatellites as described in Lai et al. (2010). The megagametophyte (1n) and its corresponding embryo (2n) were concurrently scored for each seed.

4.2.3 Data analysis

The multilocus haploid genotype of each megagametophyte was independently used to determine the female parent using a computer program written in Microsoft VisualBasic® and the male parent of each embryo was assigned using the program CERVUS 3.0.3 (Kalinowski et al. 2007) with 95% confidence and allowance for 1% genotyping errors. Results from the parentage assignment were used to estimate female and male gametic contributions, selfing rate, and the rate of pollen contamination. Genetic diversity of the filial generation was estimated using effective status number as described in Lindgren and Mullin (1998).

4.3 Results and Discussion

The combination between determining the female parents of the 801 seeds using their megagametophytes' haploid multilocus genotypes and the assignment of their male parents through pedigree reconstruction analysis was proven to be successful. After determining female and male parentage of each seed, parameters such as actual female, male and parental reproductive success, individual-parent selfing rate, and gene flow were easily estimated through direct count. Of the 49 orchard parents, 42 and 46 effectively contributed to seed and pollen pools, respectively, indicating that 7 and 3 parents did not participate as females and males in the formation of the filial generation. Female, male, and parental reproductive success estimates substantially varied among the 49 studied parents. The magnitude of this variation was assessed using Griffin's (1982) parental cumulative gametic contribution curves which indicated that approximately 80% of the orchard's female (seed) and male (pollen) gametes were produced by only 23 and 45% of parents, respectively, highlighting a greater variation in female than male reproductive success (Figure 4.1).

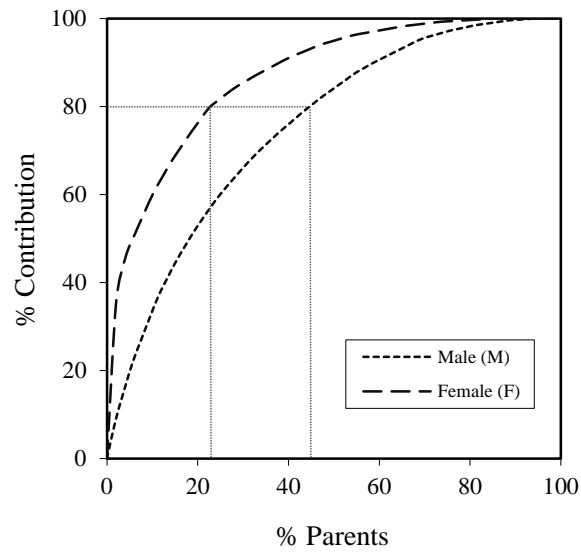


Figure 4.1 Female and male cumulative reproductive success in a Douglas-fir seed orchard showing a greater female than male gamete contribution distortion.

Two parents dominated the production of successful female gametes, contributing as much as 35.7 and 10.0% of the orchard's total seed crop, while male gametic contributions were less distorted. Although the two components were significantly correlated ($r = 0.626$, $p < 0.01$, $N = 49$) (Figure 4.2), the high coefficient of variation (76.8%) indicates that one sex's reproductive success should not be inferred from the other. The top two males produced 9.2 and 7.3% of successful gametes, with only one of them dominating both female and male gamete contribution.

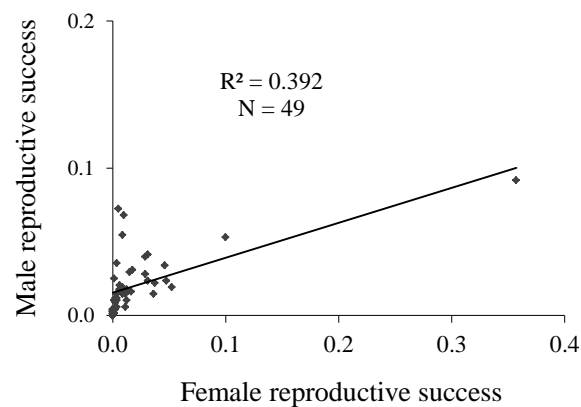


Figure 4.2 Pearson's correlation between female and male reproductive success for 49 parents in a Douglas-fir seed orchard.

The observed male reproductive success variation is not surprising and is similar to that reported recently using molecular markers on other seed orchard populations (sugi: Moriguchi et al. (2004); Douglas-fir: Slavov et al. (2005a); Nordmann fir: Hansen and Kjær (2006); western larch: Funda et al. (2008) (Chapter 2)). The use of a random sample of seeds from the entire orchard's seed crop enabled the detection of unbiased female reproductive success variation, a situation that could not have been attained if an equal number of seeds had been used from each parent. Even if individual parents' seed crops were harvested separately, female reproductive success variation would still be difficult to ascertain since seed production is not an indication of successful germination (Chaisurisri et al. 1992; El-Kassaby and Edwards 1998; El-Kassaby et al. 1992; Krakowski and El-Kassaby 2005).

A direct estimate of selfing rate of 15.2% was determined for the seed orchard population. This value is surprisingly higher than that previously reported for several Douglas-fir natural and seed orchard populations (El-Kassaby and Davidson 1991; Neale and Adams 1985; Ritland and El-Kassaby 1985; Shaw and Allard 1982; Slavov et al. 2005a). However, it should be pointed out that none of these studies factored female fertility variation during seed sampling. This selfing rate is even more surprising under the estimated external gene flow of 10.4%. Since every successful pollination event resulting from gene flow is an outcrossing event (El-Kassaby and Ritland 1986b), then the observed higher-than-usual selfing rate can only be attributable to the female fertility variation. In fact, the role of female fertility variation was confirmed by the observed high correlation between female reproductive success and selfing rate ($r = 0.956$, $p < 0.01$, $N = 42$) (Figure 4.3a). This relationship may seem to be driven by only a few exceedingly high-producing females; however, this correlation persisted even after the removal of the most productive one ($r = 0.895$, $p < 0.01$, $N = 41$). Notwithstanding the intrinsic mechanisms for selfing avoidance, namely, genetic load and polyembryony, selfing did occur, indicating that competition among selfed embryos was not effective (Bishir and Namkoong 1987; Bishir and Pepper 1977; Bramlett and Popham 1971; Koski 1971, 1973; Namkoong and Bishir 1987; Savolainen et al. 1992; Sorensen 1969; Sorensen 1982; Williams et al. 1999; Williams and Savolainen 1996). Similarly, a positive relationship was also found between male reproductive success and selfing rate ($r = 0.639$, $p < 0.01$, $N = 46$) (Figure 4.3b); however, its magnitude was much lower. The observed relationship between male reproductive success and selfing rate is

similar to that reported for a white spruce seed orchard (Denti and Schoen 1988; Schoen and Stewart 1986, 1987) and Engelmann spruce and subalpine fir natural populations (Shea 1987). The correlation between selfing and collective female-male (parental) reproductive success was almost identical to that of female ($r = 0.943$, $P < 0.01$, $N = 46$) (Figure 4.3c), indicating that the role of the female component was instrumental in sustaining the observed strong relationship.

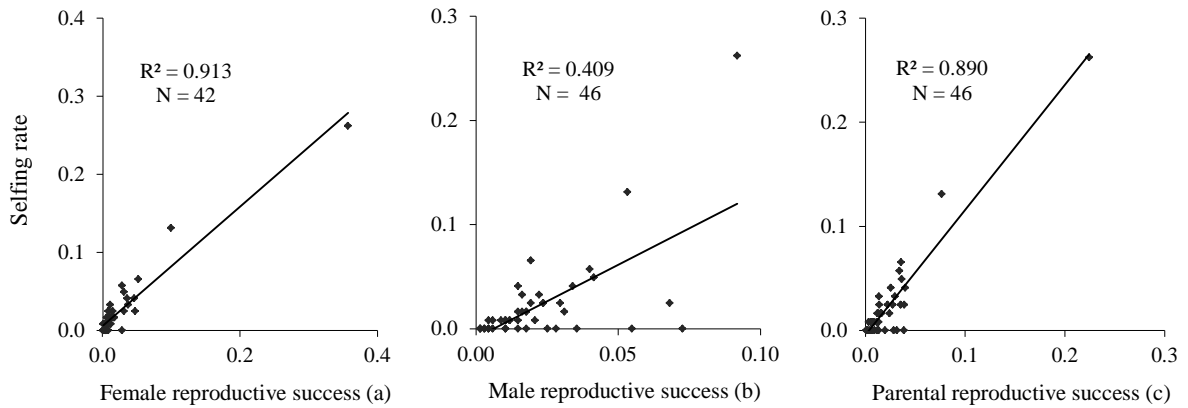


Figure 4.3 Pearson's correlations between female (a), male (b), and combined male and female (parental; c) reproductive success and selfing rate within a Douglas-fir clonal seed orchard.

It is also expected that the observed female reproductive success variation will be manifested in the observed number of embryos sired by outside pollen sources (i.e., gene flow). This was confirmed by the significant correlation between parental seed production and the number of contaminant pollen siring a particular parent ($r = 0.959$, $p < 0.01$, $N = 42$), indicating that contamination rates were nearly equal among all female parents in the orchard. The positive and significant correlations between female reproductive success on one side and selfing and gene flow on the other may seem contradictory; however, it should be pointed out that the relationship between outcrossing pollen sources (within-orchard and contaminating pollen) is exclusively independent.

Variation in female reproductive success is also expected to affect the level of genetic diversity in the filial generation. This was confirmed by the surprisingly low effective number of female parents (6.50) for a sexually active population of 42. It is important to note that the presence of the few related parents in the orchard (four parents representing two full-sib families and four parent-offspring) had a negligible effect on the resultant effective

number of females. Under no relatedness assumption, this value would still be as low as 6.54, emphasizing the role of female reproductive success variation.

Female reproductive success has often been approximated by parental representation, i.e., the number of ramets per parent (Bondesson and Lindgren 1993; Kang et al. 2001a; Lindgren and Matheson 1986; Prescher et al. 2008). In this study, this approximation was proven questionable due to a non-significant correlation between the two measures ($r = 0.270$, $p = 0.08$, $N = 42$), demonstrating that this approach requires further consideration (see Chapter 6 for a detailed study on the comparison between parental reproductive investment and success in four different conifer species seed orchards).

4.4 Conclusion

The results presented in this study are the first to demonstrate the importance of considering female reproductive success variation and its role in determining future generation's genetic diversity as affected by the mating system pattern and gene flow.

5. Partial Versus Full Pedigree Reconstruction in a Lodgepole Pine (*Pinus contorta* Dougl. ex Loud.) Seed Orchard

5.1 Introduction

The combined use of molecular genetic markers (reviewed by Vendramin and Hansen 2005; Vignal et al. 2002) and sophisticated statistical methods (e.g. Butler et al. 2004; Jones and Wang 2010; Kalinowski et al. 2007; Koch et al. 2008) enabled researchers across study areas to infer genealogical relationships among individuals through pedigree reconstruction using molecular data. Pedigree reconstruction can be accomplished either by assigning parentage (male, female, or both) to a given offspring (reviewed in Jones and Ardren 2003; Pemberton 2008) or by assigning a group of individuals into sib-ships (reviewed in Blouin 2003). While both of these approaches possess several caveats that obscure their proper functioning such as non-Mendelian transmission of alleles (occurrence of null alleles, mutations, or genotyping errors (Pompanon et al. 2005)) or the presence of non-sampled candidate parents (Koch et al. 2008), they have been widely employed as a powerful tool to resolve pedigrees in both natural and experimental populations (Chistiakov et al. 2006; Massah et al. 2010).

In forest tree populations – and in artificial production populations such as seed orchards in particular – pedigree reconstruction has often been limited to inferring the male component of parentage, as seed collected from each known female parent is analyzed separately (Burczyk and Prat 1997; Goto et al. 2005; Hansen and Nielsen 2010; Moriguchi et al. 2007; Schoen and Stewart 1986). While such studies offer the opportunity to estimate male reproductive success, population and individual selfing rate, and gene flow (both from outside the target population and among its members) as each ovule acts as an effective pollen trap, they fail in providing unbiased estimates of the female reproductive success because the family array cannot reflect actual gametic production (and, strictly speaking, the relative reproductive success) of the respective female parents. El-Kassaby et al. (2010) (Chapter 4) demonstrated in their Douglas-fir seed orchard study that female reproductive success can only be accurately inferred using a random sample of offspring (seed) from the populations' bulk seed crop. Unlike the family array in which male parentage can be successfully assigned using nuclear genome markers, bulk seed samples require additional procedures, either separate genotyping of the megagametophyte (maternal haploid tissue) and

its corresponding diploid embryo, which doubles the amount of laboratory work and genotyping efforts and thus is relatively expensive (El-Kassaby et al. 2010), or the utilization of uniparentally inherited, chloroplast or mitochondrial genomes. In most Pinaceae species, chloroplast and mitochondrial genomes are exclusively inherited from female and male parents, respectively (Neale and Sederoff 1989; Neale et al. 1986; Szmidt et al. 1987; Szmidt et al. 1988; Wagner et al. 1987), thus allowing an independent investigation of pollen- and seed-mediated gene flow (Kent and Richardson 1997).

Microsatellite DNA markers (SSRs), the highly polymorphic, co-dominant markers with ubiquitous distribution across nuclear genomes, extensively used in a wide range of fundamental and applied research in biology and medicine (Chistiakov et al. 2006; Powell et al. 1996), have also been identified in organelle genomes in a number of species. While mitochondrial SSRs (Soranzo et al. 1999) have had limited impact in forest genetics, chloroplast SSRs (Powell et al. 1996) have been employed commonly (Marshall et al. 2002; Navascues and Emerson 2005; Stoehr et al. 1998; Vendramin et al. 1996; Vendramin and Ziegenhagen 1997). Plomion et al. (2001) have proven their capability of effective separation of the female and male component of parentage without the need for additional DNA isolation, as the chloroplast haplotype composition of the megagametophyte tissue corresponded with the chloroplast haplotype of the female parent. Unlike their nuclear counterparts, however, cpSSRs possess very low levels of polymorphism, inadequate for creating multiple unique haplotypes required for the identification of every parent in a studied population, especially in production populations (seed orchards) comprising a higher number of parents.

Seed orchards are artificial populations of forest tree species established for producing genetically improved seed. They are expected to transmit their genetic superiority and diversity to the offspring; however, there are a variety of factors influencing their mating dynamics, such as variation in the reproductive phenology synchrony among parents and their relative reproductive success and gene flow from background pollen sources (pollen contamination).

In this study, the mating system dynamics of a lodgepole pine (*Pinus contorta* Dougl. ex. Loud. ssp. *latifolia* Engelm.) seed orchard was investigated. Lodgepole pine is a fast-growing, short-lived and fire-adapted hard pine species, which represents an essential

component of forest ecosystems in western North America and is the most important commercial pine species in British Columbia (Critchfield 1980). A combination of nuclear and chloroplast SSR markers was used to evaluate two approaches of pedigree reconstruction analyses: 1- offspring (seed) collected from an array of known female parents and 2- offspring randomly selected from a bulk collection of all parents' seed crops, thus with unknown male and female parentage. Results from the two approaches were compared to ascertain their capability in discerning the parental reproductive success, selfing rate, gene flow from background pollen sources, and effective number of male and female parents.

5.2 Materials and Methods

5.2.1 Seed orchard population and plant material sampling

A clonal lodgepole pine seed orchard (Pacific Regeneration Technologies; located near Armstrong, British Columbia; 50°23' N, 119°17' E, 470 m a.s.l.; established in 1994) provided material for this study. At the time of seed collection (2007) its population consisted of 71 replicated parents with a total of 1,047 ramets (13.9 ± 7.0 replicates per parent). Dormant vegetative buds from two ramets of each of the 71 parents, resulting in a total of 142 sampled ramets, a seed array of 11 maternal half-sib families with varying sample size (56.3 ± 7.3), and a random sample of 635 bulk seed with unknown male and female parentage (i.e., mixed over all seed-producing ramets and thus representing the entire population's reproductive output) were collected. The dormant buds were stored at -80°C until DNA extraction while the seed were stored at +4°C until germination.

5.2.2 DNA extraction and SSR genotyping

Genomic DNA was isolated from both the vegetative buds and germinated seeds (ca 2–3 cm long, 7-day-old embryos) using a modified Doyle and Doyle's (1990) CTAB method. All samples were genotyped using an array of nine polymorphic nuclear microsatellite loci developed originally for loblolly pine and successfully cross-amplified on lodgepole pine, following the protocol described in Liewlaksaneeyanawin et al. (2004). In addition, parents and bulk seed were genotyped using an array of six chloroplast microsatellite loci (Stoehr and Newton 2002), grouped into two multiplexing sets (G2R1 + 69FR + 9/87 and I1A2 + 10FRR + L2T1) and amplified following the protocol by Stoehr and Newton (2002). Each reaction (10 µL) consisted of 100 ng of total genomic DNA, 0.5 µM of each forward and

reverse primer, 0.2 mM of dATP, dCTP, dGTP, and dTTP, 1X PCR buffer (10 mM Tris-HCl, 1.5 mM MgCl₂, 50 mM KCl, pH 8.3; Roche, Laval, QC), 0.25 U of *Taq* polymerase (Roche, Laval, QC), and 0.5 pmol of M13 infrared label (LiCor Inc., Lincoln, NE). PCR amplification consisted of 5 min denaturing at 94°C, 25 cycles of 30 sec at 94°C, 30 sec at 55°C, and 60 sec at 72°C, and 5 min of final extension at 72°C. Amplicons were electrophoretically separated on 5.5% Long RangerTM polyacrylamide gels using a LiCor 4200 automated sequencer (LiCor Inc., Lincoln, NE) and visually scored according to their molecular weight using software program SAGATM.

5.2.3 Mendelian inheritance tests

To determine the mode of inheritance of the microsatellite markers used, the maternal genotypes were inferred by assigning a specific color code to each allele within a progeny array of a heterozygous tree. Based on the expected Mendelian inheritance, all offspring within a progeny array should display at least one maternal allele (discerned from the co-dominant segregation of alleles). The observed maternal allele ratios at each locus were tested against the expected 1:1 Mendelian segregation ratio using log-likelihood *G* test (Sokal and Rohlf 1981). Individual *G*-statistics was calculated from segregation data of each heterozygous parent. *G*-test of goodness-of-fit (*G_P*) and *G*-heterogeneity test (*G_H*) were also conducted on the pooled segregation data from multiple trees. Significance of *G_P* and *G_H* indicates deviation from the expected Mendelian segregation and heterogeneity among segregating trees, respectively. All offspring with an identical genotype to that of the female parent were excluded from the analyses, as in these instances the maternally transmitted allele(s) could not be parsed out from those contributed by fathers.

5.2.4 Parentage analyses

Parentage analyses were performed using CERVUS 3.0.3 (Kalinowski et al. 2007), a likelihood-based paternity inference method with a known level of statistical confidence and accounting for genotyping errors. Two types of parentage analysis were conducted: paternity analysis for the family array and parent-pair analysis with unknown sexes of the candidate parents for the bulk seed sample. The pool of sampled candidate fathers (pollen donors) comprised of the 71 orchard parents plus three detected alien genotypes (see below) was assumed to represent 95% of all possible candidate fathers involved in the production of the

offspring generation. Genotyping error rate estimated from the known mother-offspring genotypes across the nine nuclear loci was 0.01. The parentage analyses were conducted based on 10,000 simulations with 74 sampled candidate parents, genotyping error rate of 0.01, and with 95% (strict) confidence level. The log-likelihood (LOD) score, the likelihood that the candidate parent is the true parent divided by the likelihood that the candidate parent is not the true parent, was calculated for each putative parent. The delta score, the difference in LOD scores of the two most likely candidate parents, was used as a criterion for parentage assignment at 95% confidence level.

To identify the paternal parentage of each offspring for the bulk seed sample, the identity analysis with cpSSRs was conducted in CERVUS 3.0.3. Since CERVUS cannot directly analyze haploid genetic data, dummy genotypes (homozygous offspring) were created. To account for genotyping errors at loci I1A2, 10FRR, and 69FR, in which alleles differed in size by one base pair only, one or two mismatches were allowed between the candidate male parent and a given offspring. When a male parent was assigned with mismatching loci, then these loci were excluded and the identity analysis was repeated in order to increase its accuracy. For each offspring, the male parent determined by the identity analysis was compared with the two parents identified by the parent-pair analysis. The identity analysis fully corresponded with the parent pair analysis, as for each of the analyzed offspring the assigned candidate paternal parent was the same as one of the two most likely parents determined by the parent-pair analysis (in total 545 offspring). Finally, the maternity analysis with known fathers was conducted for these 545 offspring, using the same parameters described earlier.

5.2.5 Reproductive success and genetic parameters estimates

Results from the parentage analyses served to estimate the male, female, and parental (joint male and female) gametic contributions (reproductive success), selfing rate, and the rate of gene flow from background sources (pollen contamination). Genetic diversity was approximated by pedigree-based effective number of parents, N_e , which describes the proportion of parents involved in the production of the next generation, as

$$N_e = \frac{1}{\sum_{i=1}^N p_i^2} \quad [5.1]$$

where p_i is the gametic contribution of parent i as male (m), female (f), and combined male and female parent (p) and $0 \leq p_i \leq 1$, $\sum_{i=1}^N p_i = 1$, and $p_i = \left(\frac{m_i + f_i}{2}\right)$.

5.3 Results and Discussion

5.3.1 Mendelian inheritance

Maternal genotypes were inferred from the segregation pattern of their offspring at the nine nuclear SSR loci and used for Mendelian inheritance analyses. Out of 84 individual G -tests, 74 showed no deviation from the expected 1:1 Mendelian segregation ratio. For the purpose of adequate statistical power, G -test of goodness-of-fit (G_P) and G -heterogeneity test (G_H) were also conducted on the pooled segregation data from multiple parents sharing the same genotype at all of the SSR loci, except for PtTX3049 and LOP5 (Table 5.1). These two loci had no heterozygous female parents sharing the same genotype; nevertheless, individual G -test confirmed the mode of inheritance for both markers. Only one out of 19 individual G -tests showed a significant deviation from the expected 1:1 segregation ratio.

There were no significant differences in segregation analyses for both the pooled and heterogeneity G values at PtTX2146, PtTX3011, PtTX3029, PtTX3127, and PtTX4054, confirming that they segregated in accordance with Mendelian expectations (Table 5.1). The segregation of the pooled data at PtTX3025 and PtTX4058 showed a significant G_P , indicating an excessive departure from the expected 1:1 segregation ratio. Significant G_H was only observed at one segregating genotype at PtTX4058 (164/168), indicating the absence of different segregation patterns among female parents. The significant deviations from the expected segregation ratio at PtTX3025 and PtTX4058 were due to the unexpected higher frequency of one allele over the other (e.g. alleles 286 and 168 bp at loci PtTX3025 and PtTX4058, respectively) (Table 5.1).

Table 5.1 Log-likelihood *G*-test on segregation ratio of nine microsatellite loci for the same heterozygous genotype of two parents from a lodgepole pine seed orchard.

Locus	Maternal genotype	Number of trees	Observed ratio	Pooled G^1	p	Heterogeneity G^2	p
PtTX2146	209/215	2	47:34	2.10	0.15	1.51	0.22
PtTX3011	169/178	2	49:41	0.71	0.40	1.33	0.25
PtTX3025	286/292	3	61:35	7.13	0.03	0.24	0.89
PtTX3025	277/286	3	29:69	16.81	0.00	2.41	0.30
PtTX3029	280/286	2	49:52	0.09	0.76	1.17	0.28
PtTX3127	199/208	2	28:38	1.52	0.22	0.02	0.89
PtTX4054	297/307	2	47:42	0.28	0.57	2.87	0.09
PtTX4058	168/172	2	54:34	4.59	0.03	0.78	0.38
PtTX4058	164/168	2	39:61	4.88	0.03	4.44	0.04

¹ Pooled *G* values indicate the overall deviation from 1:1 ratio. ² Heterogeneity *G* values indicate the amount of heterogeneity in the segregation.

5.3.2 Genetic diversity analysis

Genetic polymorphism of the nine nuclear SSR loci was estimated based on the 74 orchard parents using CERVUS. All nine loci exhibited a high level of polymorphism with a mean number of 15.2 ± 4.0 alleles per locus and an average expected heterozygosity of 0.84 ± 0.08 (Table 5.2). The average null allele frequency was estimated to be 0.02, with the highest value of 0.10 for PtTX3049 (Table 5.2). Microsatellite loci with null alleles (Callen et al. 1993) have been identified in numerous species (Varshney et al. 2005) and can lead to false parentage exclusions, thus reducing the rate of assignment (Dakin and Avise 2004). Three loci (PtTX3011, PtTX3049, and LOP5) exhibited relatively high null allele frequencies (over 0.05); however, since their genotyping error rate was estimated to be relatively low (max 0.05 for PtTX3011) (Table 5.2), I still included them in the analyses.

The six chloroplast SSR loci produced a mean number of 4.8 ± 1.3 alleles per locus, ranging from 4 (9/87, 10FRR, 69FR, L2T1) to 7 (G2R1), and provided in total 51 multilocus haplotypes among the 74 studied parents. Of these 51 haplotypes, 38 were unique to only one orchard parent while the remaining ten, two, and one haplotypes were common for two, six, and three parents, respectively. The observed level of polymorphism at these loci was similar to a previous study by Stoeckert and Newton (2002).

Table 5.2 Number of alleles (A), observed (H_o) and expected (H_e) heterozygosities, frequency of null alleles (f_n), and genotyping error rate (e) for nine microsatellite loci based on genetic analysis of 74 parents in a lodgepole pine experimental population.

Locus	A	H_o	H_e	f_n	e
PtTX2146	13	0.836	0.843	0.001	0.000
PtTX3011	20	0.797	0.909	0.063	0.053
PtTX3025	8	0.703	0.702	0.000	0.000
PtTX3029	15	0.892	0.852	0.000	0.000
PtTX3049	18	0.730	0.891	0.100	0.031
PtTX3127	11	0.757	0.733	0.000	0.003
PtTX4054	16	0.889	0.915	0.009	0.000
PtTX4058	16	0.795	0.837	0.022	0.002
LOP5	20	0.785	0.915	0.075	0.000

5.3.3 Verification of parental affiliation

Benefiting from the multilocus genotypic profiles of the 142 screened ramets, two mislabeling errors and three alien, off-population genotypes were identified within the studied population. While the mislabeled individuals' profiles perfectly matched those of other two parents and thus could be corrected for their actual affiliation, the alien genotypes – probably the result of overgrowing root stocks – were subsequently treated as additional parents in the parentage analyses, adjusting the final number of sampled candidate parents to 74. Mislabeling of ramets, which may result in biased estimates of seed crop genetic parameter as well as incorrect application of controlled crosses among desired parents for the production of advanced-generations breeding, is a commonly occurring issue in commercial production populations. It has been reported e.g. by Harju and Muona (1989), Kawauchi and Goto (1999), Moriguchi et al. (2005), Slavov et al. (2004), Wheeler and Jech (1992).

5.3.4 Parentage analysis and gene flow

For the analysis of the array of the 11 maternal half-sib families, male parentage was successfully assigned to 528 out of 619 offspring (85.3%) while for the bulk seed sample, both male and female parentage was successfully assigned to 522 out of 635 offspring (82.2%) (Table 5.3).

Table 5.3 Summary of parentage analysis of 11 maternal half-sib families (known female parentage) and bulk seed (unknown male and female parentage) at 95% confidence levels.

Parent ID	Number of offspring			Number of full-sib families
	Analyzed	Assigned at 95%	Unassigned	
10	42	36	6	22
15	59	28	31	18
20	60	53	7	27
34	59	53	6	27
37	41	35	6	17
38	59	50	9	27
46	59	52	7	24
52	60	54	6	31
53	60	52	8	22
61	60	58	2	29
64	60	57	3	24
Family	619	528	91	268
Bulk	635	522	113	446

Failure to assign paternity and/or maternity to the remaining offspring was either due to the production of insufficiently informative genotypes to match the candidate parents with 95% confidence, the presence of offspring (seed) sired by unsampled parents from outside the studied population, or a combination of both. However, since the nine nuclear SSRs used in this study are highly polymorphic and possess low null allele frequencies, and since most of the unassigned offspring had mismatches on at least two loci, then it is unlikely that the loci failed to provide enough statistical power. The utilization of the additional six uniparentally inherited chloroplast markers improved the assignment rate on the male side, with 545 offspring being successfully assigned to one of the candidate males (85.8%), indicating that a small fraction of the bulk seed had been collected from additional alien mother trees which were not sampled during dormant bud collection of the 142 ramets and thus represent seed contamination. The unassigned offspring on the male side are most likely a product of gene flow from non-sampled candidate male parents that occurred in the vicinity of the studied orchard one year prior to seed collection (i.e., pollen contamination) as pine species possess a two-year reproductive cycle after pollination. Pollen contamination rate of 14.7 and 14.2% was estimated in this studied seed crop from the family array and bulk seed sample, respectively, values slightly higher as compared with the 8% previously reported for another

lodgepole pine seed orchard (Stoehr and Newton 2002). In a similar study, an independent analysis of embryonic and megagametophytic cpDNA was successfully applied to estimate pollen contamination in a bulk seed collection from a maritime pine (*Pinus pinaster* Ait.) seed orchard (Plomion et al. 2001). In this study, however, the combination of parent-pair and identity analyses enabled the assignment of both male and female parentage to the bulk seed sample without the need to genotype the megagametophyte.

5.3.5 Pedigree reconstruction and parental reproductive success

Pedigree reconstruction of the family array revealed 268 full-sib families nested within the 11 maternal half-sib families, ranging in number from 17 (family #37) to 31 (family #52) and in size from 1 (56.3% of all families) to 15 (one family) (Figure 5.1 and Table 5.3). Pedigree reconstruction of the bulk seed revealed 446 full-sib families, ranging in size between 1 (86% of all families) and 4 (2 families) (Figure 5.2 and Table 5.3). As expected, bulk seed provided a better representation of the seed orchard population, capturing 65 out of the 74 candidate mothers (87.8%), and consequently revealed a higher number of full-sib families. The paternal half-sib family sizes ranged from 0 (4 families) to 58 (1 family) and from 0 (5 families) to 28 (1 family) for the family array and bulk seed, respectively, with a positive correlation of 0.61 between them (Figure 5.3). The variation in the paternal half-sib family sizes between these two approaches might have been due to the sampling technique of the seed assayed, because seed representing each maternal half-sib family was only collected from one single ramet (i.e., one position) while the bulk seed sample was taken from a mixture of seed collected from numerous ramets representing the entire seed-producing population. It could also be due to variation in reproductive phenology, dividing the parental population into a number of smaller, mutually non-overlapping groups with no or limited gene flow among them (El-Kassaby 1995).

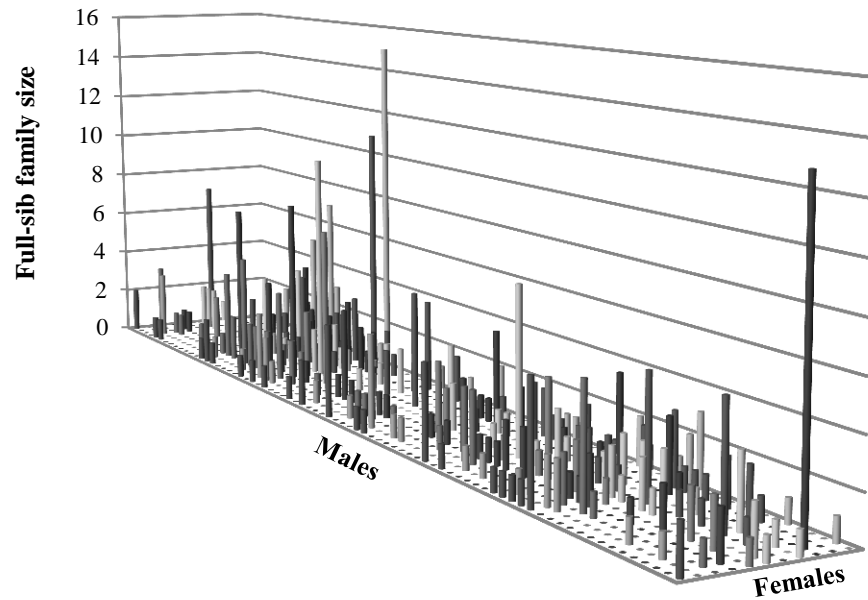


Figure 5.1 Distribution of naturally occurred crosses (mating design) in a lodgepole pine seed orchard revealed by partial pedigree reconstruction of 11 wind-pollinated maternal half-sib families using nine nuclear microsatellite loci.

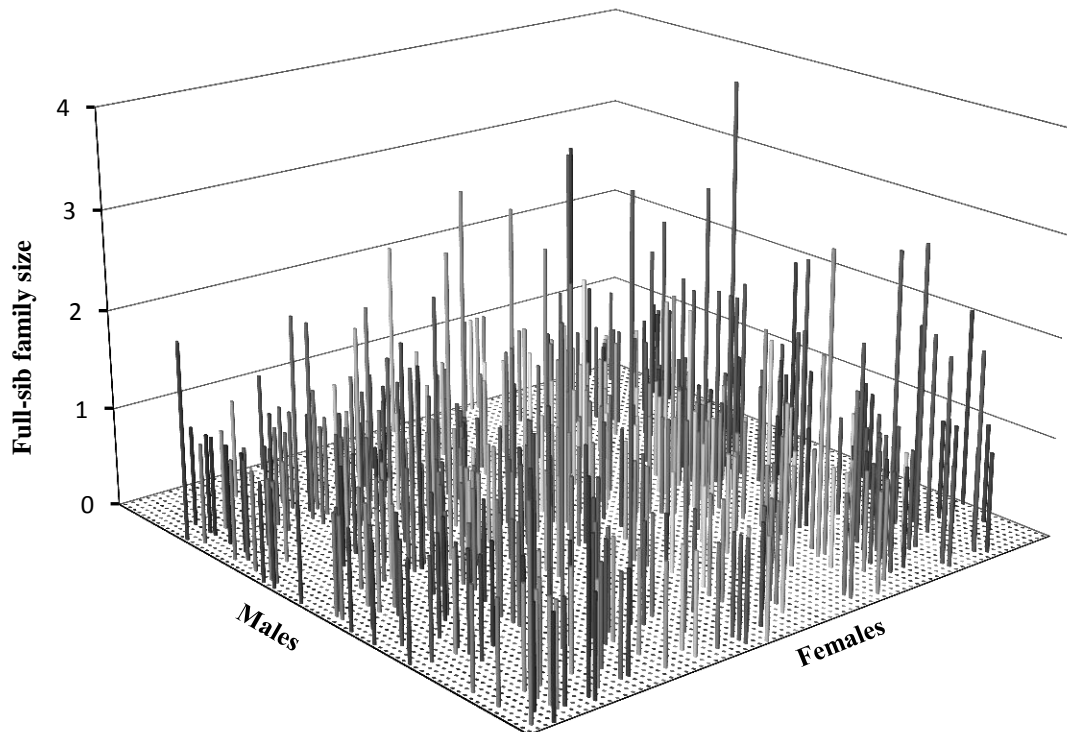


Figure 5.2 Distribution of naturally occurred crosses (mating design) in a lodgepole pine seed orchard revealed by full pedigree reconstruction of offspring with unknown male and female parentage using a combination of nine nuclear and six chloroplast microsatellite loci.

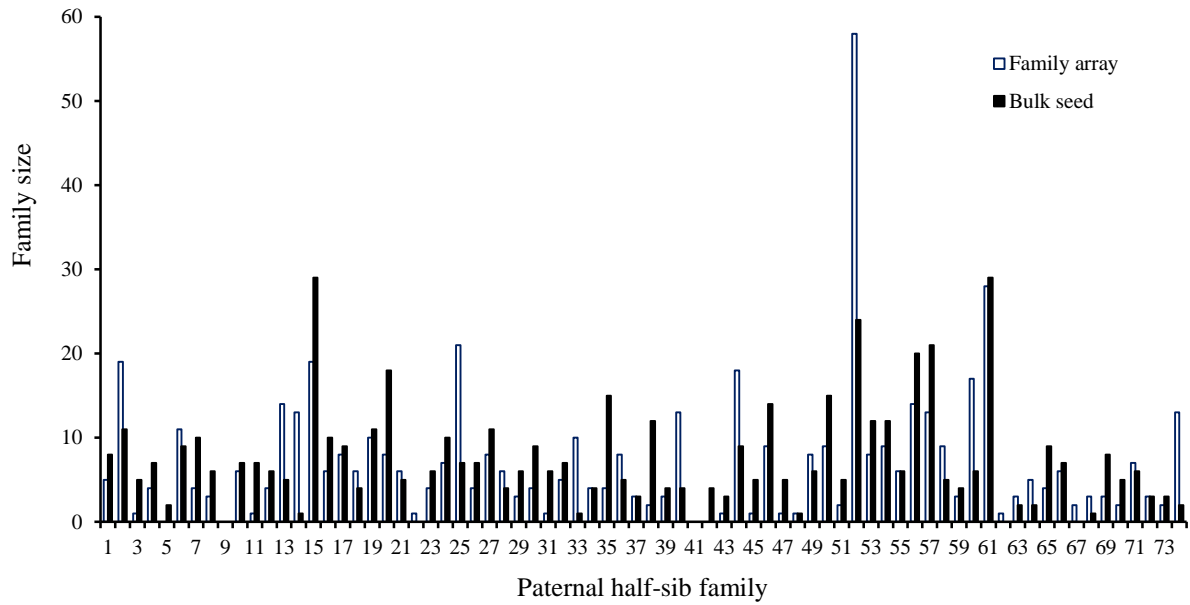


Figure 5.3 Comparison of paternal half-sib family sizes obtained from partial (family array) and full (bulk seed) pedigree reconstruction of offspring from a lodgepole pine seed orchard.

As to parental gametic contributions, it has been commonly observed that pollen- and seed-cone production is dominated by a small fraction of the parental population (El-Kassaby and Askew 1991; El-Kassaby et al. 1989; Roberds et al. 1991; Schoen and Stewart 1986). For the female component, this phenomenon was conveniently termed the “20/80” rule (Anonymous 1976), meaning that 20% of a population produce 80% of the seed-cone crop. Whereas this rule was derived from phenotypic assessment of female fecundities only, El-Kassaby et al. (2010) (Chapter 4) and Lai et al. (2010) verified such a distortion experimentally using full pedigree reconstruction of a bulk sample of a Douglas-fir seed crop, in which a single mother (out of 49, ~2%) contributed as much as 35.7% of the total seed crop and the most successful ten mothers (~20%) as much as 77%. In this study, the distortion was not so substantial, as 80% of the seed crop was produced by 41% of mothers (Figure 5.4; right). Male contributions based on both the family array and bulk seed displayed a slightly better balance, with 80% of the successful pollen being contributed by 45 and 53% of fathers, respectively (Figure 5.4; left). The lower value for the family array could again be explained by the sampling method, as seed collected from a single ramet is more likely sired by neighboring individuals than more distant ones (Burczyk et al. 1996; Erickson and Adams 1989) and in turn may possess lower variation in paternity. Combined

male and female contribution was similar to that of male, with 80% of all successful gametes produced by 53% of the orchard's population (Figure 5.4; right).

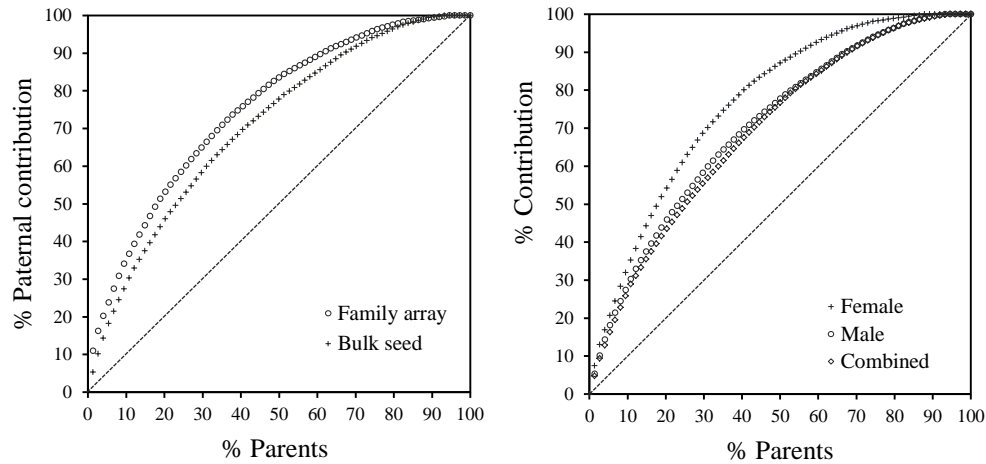


Figure 5.4 Comparison between cumulative male reproductive success from partial (family array) and full (bulk seed) pedigree reconstruction (left) and between male, female, and combined (parental) reproductive success from full pedigree reconstruction (right) of offspring from a lodgepole pine seed orchard. The diagonals represent equal contribution among parents.

5.3.6 Selfing

In both family array and bulk seed, 10 offspring were assigned to the same seed and pollen parent, resulting in a similar level of selfing of 1.6%. Comparable estimates were obtained in another lodgepole pine seed orchard (2%: Stoeher and Newton (2002)) as well as in natural populations (1–5 %: Epperson and Allard (1984); Perry and Dancik (1986)). Low effective selfing rates in conifers are mainly due to the abundance of wind-borne pollen and the high inbreeding depression at the seed stage, which filters out most of the selfed seed by abortion (Husband and Schemske 1996). Specifically in lodgepole pine it may also be due to protandry, a mechanism securing that pollen shedding peaks a few days prior to female receptivity on the same tree (Owens et al. 1981a), as well as polyembryony, which may provide opportunity for outcrossed embryos to outcompete their selfed counterparts (Ledig 1998). Since in this study the selfing rate was inferred from germinated seeds (i.e., effective selfing), then this estimate was probably downward-biased because none of the selfed seed that did not reach the embryonic stage were included in the selfing rate estimation. However, this parameter is more important from the evolutionary as well as forest management

perspectives because the selfed embryos, provided they would reach sexual maturity, would be involved in the production of the next generation as inbred trees.

5.3.7 Effective number of parents

While the allelic richness and expected heterozygosity as shown in Table 5.2 provide insight into the genetic diversity of a studied population, the concept of effective number of parents is commonly employed to express the relative reduction in population size from its census number due to non-random union of gametes. As anticipated, the pedigree-based N_e was lower than the census number of parents in the studied population for both the family array and bulk seed. For the male component, the family array and bulk seed provided an estimate of 32.0 and 44.0, respectively. The latter was closer to what can be considered to be the “actual” value of 40.6 obtained using a larger sample size, composed of both family array and bulk seed samples (Funda et al. 2011) (Chapter 6). Female N_e was only provided by the bulk seed and an estimate of 35.4 was obtained. The combined parental (male and female) N_e provided an estimate of 47.4, thus 64.1% of the census size of the studied population.

5.4 Conclusion

Unlike the widely applied analysis of family array that can only infer male-related genetic parameters (male reproductive success, selfing rate, gene flow, and effective number of fathers), the full pedigree reconstruction of bulk seed proved to be effective in providing information on the female component as well, thus offering a more complex picture of the mating dynamics in a studied population. It should be highlighted that the sampling nature of the family array restricted the generated information only to the selected 11 female parents as well as to the limited spatial locations of their replicates within the seed orchard population, thus with restricted pollen pools. On the other hand, bulk seed provided an unbiased sampling of the entire population, enabled the concurrent estimation of both male and female reproductive success, and was superior in the number of sampled parents, the number of assembled full-sib families and the effective number of parents. The close correspondence between results from the partial and full pedigree reconstruction further indicates that genetic parameters of the offspring population (such as seed crop) can be inferred with no prior knowledge of the female parentage. The results of the parent-pair analysis using nuclear genetic markers provided information on the two most likely parents (whose sexes were not

specified) while the identity analysis using uniparentally inherited, organelle markers (chloroplast SSRs in this study) helped in determining the male parent. Additionally, the full pedigree reconstruction using individuals with unknown male and female parentage enabled the posterior assemblage of naturally occurring crosses among the population's members, resulting in the creation of a mating design in the extent that would otherwise only be accomplishable by controlled pollination with extremely high costs and labor efforts (Figure 5.2). This is the premise of the novel tree breeding concept known as breeding without breeding (El-Kassaby and Lstibůrek 2009).

6. Congruence Between Parental Reproductive Investment and Success Determined by Microsatellite DNA-based Pedigree Reconstruction in Conifer Seed Orchards

6.1 Introduction

Forest tree recurrent selection programs attain their goals through two distinctive streams: breeding and production. Breeding with its classical three steps – crosses, testing, and selection – leads to the identification of elite individuals for further breeding or the establishment of production populations (seed orchards) (Namkoong et al. 1988). Seed orchards are exclusively dedicated to packaging the genetic gain captured from breeding, maintaining acceptable level of genetic diversity, and delivering both gain and diversity to future forests (El-Kassaby 2000). As breeding programs advance, higher gain is attained and greater focus is directed towards the management of genetic diversity to counterbalance the inevitable build-up of co-ancestry and the unavoidable reduction in genetic diversity associated with the domestication process (El-Kassaby 2000). Seed orchards are expected to function as closed, panmictic populations meeting Hardy-Weinberg expectations so that seed crops' allelic and genotypic frequencies reflect those of the parental populations (Eriksson et al. 1973). This ideal condition is rarely met; however, seed orchard populations have proven to be robust and substantial gains have been captured even under these deviations (El-Kassaby 1989; Xiang et al. 2003). Estimating seed orchards' reproductive success (male and female), the rate and genetic quality of pollen contamination, and, when practiced, the effectiveness of SMP (El-Kassaby et al. 1993a) is essential for calculating crops' genetic parameters, namely, gain and diversity (Stoehr et al. 2004).

Evaluation of female reproductive success is commonly approached using visual surveys of various intensities, applied on individual parent basis, that include number, volume, or weight of seed-cones and/or number of seeds or filled seeds per cone (Byram et al. 1986; El-Kassaby 1989; Griffin 1982; Kang and El-Kassaby 2002; Stoehr et al. 2004); on the other hand, male reproductive success is much more difficult to assess using visual surveys (Denti and Schoen 1988; Kang and El-Kassaby 2002; Roberds et al. 1991; Savolainen et al. 1993; Schmidting 1983; Schoen et al. 1986; Schoen and Stewart 1986, 1987). It should be pointed out that all visual survey methods assume that reproductive investment equals to reproductive

success, an assumption requiring validation (El-Kassaby and Cook 1994; El-Kassaby et al. 2010; Reynolds and El-Kassaby 1990; Schoen and Stewart 1986, 1987).

The exact determination of parental reproductive success, especially its male component, involves the use of elaborate DNA fingerprinting techniques and parentage assignment, requiring financial resources and technical expertise (Hansen and Kjær 2006; Lai et al. 2010; Moriguchi et al. 2004; Slavov et al. 2005a). Even when these methods are employed, the results are retrospective and cannot be used to manipulate the genetic potential of developing seed crops. Comparative studies between DNA-based reproductive success and survey-based reproductive investment are needed to determine the magnitude of their congruence. In addition to the time and resource savings, the development of representative assessment methods would allow orchard managers to be responsive to crops' genetic profiles and provide opportunities for their genetic gain – diversity optimization (Funda et al. 2009) (Chapters 7 & 8) as such methods would have the advantage, compared to DNA-based pedigree reconstruction analyses, to predict the crops to be harvested a few months later.

In this study I addressed the following objectives: to compare five female and three male reproductive investment survey methods proposed by Woods (2005) with those based on parental representation (i.e., the number of ramets per parent), evaluate their congruence with the actual reproductive success obtained from DNA fingerprinting and parentage determination, and select the most appropriate ones for conifer seed orchard management.

6.2 Materials and Methods

6.2.1 Seed orchard populations

Clonal seed orchards of lodgepole pine (*Pinus contorta* Dougl. ex. Loud. ssp. *latifolia* Engelm.; Pacific Regeneration Technologies' first-generation orchard located near Armstrong, British Columbia; 50°23' N, 119°17' E, 470 m a.s.l.; 70 parents, established in 1994), Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco; Western Forest Products' second-generation orchard located on the Saanich Peninsula, southern Vancouver Island, British Columbia; 48°35' N, 123°24' W, 50 m a.s.l.; 49 parents, established in 1990), western larch (*Larix occidentalis* Nutt.; Kalamalka Forestry Centre's first-generation orchard located near Vernon, British Columbia; 50°14' N, 119°16' E, 480 m a.s.l.; 41 parents, established in 1989) and western redcedar (*Thuja plicata* Donn. ex D. Don; Western Forest Products' first

generation orchard located next to the Douglas-fir orchard; 27 parents, established in 2000, one third of ramets top-grafted) provided material for this study.

6.2.2 Reproductive investment

Parental reproductive investment (female and male) was assessed using multiple simplified visual survey methods proposed by Woods (2005), progressing with more details to provide an assumed concomitant increasing level of accuracy. The assessment was conducted in 2005 (Douglas-fir and western larch), 2006 and 2007 (lodgepole pine; male and female fecundity, respectively), and 2008 (western redcedar). The male fecundity was assessed on each ramet of all parents as follows:

- lodgepole pine: visual estimate of the number of pollen clusters,
- Douglas-fir and western larch: visual estimate of the volume of pollen cones,
- western redcedar: based on the abundance of male buds, ramets were classified into five rank classes (from 0 – no buds present to 4 – buds highly abundant), which were subsequently converted into biologically meaningful categories of potential pollen production, such that classes 0, 1, 2, 3, and 4 were assumed to produce 0, 1, 6, 18, and 40 units of pollen, respectively.

The survey methods applied for the studied species are listed in Table 6.1 while their detailed description, provided in Woods (2005), is summarized as follows:

6.2.2.1 Female reproductive investment

F1: visual estimate of seed-cone count on every ramet for all parents,

F2: measure of harvested seed-cone volume by parent,

F3: measure of harvested seed-cone weight by parent (weighing is done either within 2 days of cone harvest or following cone drying),

F4: seed-cone count calculated from weight of a standard volume (5 L) by parent (requires estimating the total seed-cone weight for each parent), and

F5: seed-cone count calculated from the standard volume by parent (requires estimating the total seed-cone volume for every parent).

For Douglas-fir, reproductive investment data from 17 parents were excluded from survey methods F4 and F5 due to inadequate seed-cone production (standard volume) and thus the

R^2 -based comparisons were restricted to within groups of equal sample sizes (F1, F2, and F3 vs. F4 and F5) (Table 6.1).

6.2.2.2 Male reproductive investment

M1: parental representation (i.e., number of ramets per parent) followed by crown volume adjustment,

M2: visual assessment of pollen-cone production on at least 50% of ramets (conducted prior to pollen-cone fall), and

M3: visual assessment of pollen-cone production on every ramet of all parents (i.e., 100% sampling, conducted prior to pollen-cone fall).

Furthermore, prediction of male reproductive success from female reproductive investment was evaluated.

6.2.2.3 Parental representation

Additional two methods, denoted as F0 and M0 for female and male reproductive investment, respectively, were tested in this study. These methods assume that parental reproductive success is equal to parental representation, i.e., that successful pollination events and production of viable seeds are a function of the number of ramets representing a particular parent. Unlike M1, however, they do not involve any crown volume or age adjustment. Note that parental representation was adjusted for the presence of top-grafts in western redcedar such that the inter-stock and top-graft were given weight of 2/3 and 1/3, respectively.

6.2.3 Reproductive success

The actual parental reproductive success (female and male) was estimated using nuclear (all species) and chloroplast (lodgepole pine) microsatellite DNA markers and a combination of partial and full pedigree reconstruction (see Chapters 2, 3, 4, & 5 for details on western larch, western redcedar, Douglas-fir, and lodgepole pine parentage analyses, respectively). Plant material consisted of samples of young vegetative buds or fresh foliage representing all parental genotypes in the four orchards as well as random samples of clonally harvested (lodgepole pine, western larch, and western redcedar) and bulk seed (lodgepole pine and Douglas-fir). Bulk seed samples were a representative of the entire orchards' seed crops.

All seeds were soaked in water for 24 hours, stratified for 3 weeks in +4°C (this pre-treatment was skipped for western redcedar), germinated in an incubator at alternating temperatures (30°C/20°C for 8 hours light and 16 hours dark, respectively), and embryos 1–3 cm in length were collected for DNA analysis. In addition, vegetative buds were collected from a 2-year-old western larch progeny test representing 15 open-pollinated families from the same 2005 seed crop in order to increase the sample size (El-Kassaby et al. 2011). Genomic DNA was isolated using a modified CTAB method (Doyle and Doyle 1990) and microsatellite DNA amplicons (GeneAmp 9700 thermal cycler; Perkin-Elmer, Foster City, CA) were electrophoresed on polyacrylamide gels (Long Ranger™) with a LiCor 4200 or 4300 automated sequencer (LiCor Inc., Lincoln, NE) and visually scored using a software program SAGA™ (LiCor Inc., Lincoln, NE). When bulk seed were used, the female component in Douglas-fir was discerned using the corresponding megagametophyte tissue of each embryo while an additional six cpDNA SSRs (Stoeckert and Newton 2002) were employed along with their nuclear counterparts for the identification of the male parent in lodgepole pine. Parentage was assigned using CERVUS 3.0.3 (Kalinowski et al. 2007) with 95% confidence level as a criterion for correct assignment throughout the analyses.

The number of parents per orchard and gametes assigned to either the female or male parents are summarized in Table 6.1. All unassigned gametes, which were either a product of contamination or insufficiently informative genotypes, were excluded from the analyses. The final numbers of male and female gametes used in this study following this adjustment are summarized in Table 6.1.

6.2.4 Survey methods evaluation

The DNA-based female and male reproductive success estimates determined from the pedigree reconstructions served as the benchmark for evaluation of the various visual survey methods where the former and the latter were considered as dependent and independent variables in subsequent linear regression analyses, respectively. The proportional female and male gametic contributions (the dependent variables), were transformed following Eq. 6.1 in order to meet the regression analysis assumptions

$$y' = \arcsin \sqrt{y} \quad [6.1]$$

and the statistical computation was performed using SAS 9.1.3 (SAS Institute Inc., Cary, NC). Both within female and male survey comparisons were based on the coefficient of determination (R^2) and, when sample size differed across survey methods (e.g., when some parents' seed-cone production did not reach the standard volume of 5 L), they were based on the square root of the error mean square ($RMSE$) (Table 6.1). Furthermore, 95% prediction intervals, i.e., 95% confidence intervals for individual observations (Hahn and Meeker 1991) were calculated.

6.2.5 Genetic worth and diversity

The seed crop genetic worth (volume gain at age 60) and diversity were determined using the best combination of female-male reproductive investment survey methods in terms of fit statistics (Table 6.1) and were compared to those based on pedigree reconstruction. The genetic diversity was expressed as effective status number (N_e ; Lindgren et al. (1996)) and was calculated as

$$N_e = \frac{1}{2\Theta} \Leftrightarrow N_e = \frac{2}{\sum_{i=1}^N \sum_{j=1}^N (f_i + m_i)(f_j + m_j) c_{ij}} \quad [6.2]$$

where N is the size of the parental population, Θ is the average co-ancestry among parents, including themselves (i.e., self co-ancestries) (Cockerham 1967), f_i and m_i are proportional female and male gametic contributions of parent i to a seed crop, respectively, such that $0 \leq f_i \leq 1$, $0 \leq m_i \leq 1$, $\sum_{i=1}^N f_i = 1$, and $\sum_{i=1}^N m_i = 1$, and c_{ij} is the co-ancestry coefficient between parents i and j . The status effective number thus describes the proportion of parents involved in the production of the next generation (i.e., a seed crop), considering the genetic relationships among the parental population. The co-ancestry coefficients are 1/2, 1/4, 1/4, 1/8, and 0 for self co-ancestry, full-siblings, parent-offspring, half-siblings and individuals with no genetic relationship, respectively. Of the studied orchards, co-ancestry other than self co-ancestry was only present in the Douglas-fir seed orchard which consisted of four pairs of parent-offspring and two full-sib families, each represented by two parents. The genetic worth (ΔG) was calculated as

$$\Delta G = \sum_{i=1}^N BV_i \frac{(f_i + m_i)}{2} \quad [6.3]$$

where BV_i , f_i , and m_i denote breeding value and female and male gametic contribution of parent i to the seed crop, respectively. Parental breeding values (volume gain at age 60 determined from an extensive multi-site progeny test programs) were provided by the B.C. Ministry of Forests, Mines and Lands.

6.3 Results and Discussion

6.3.1 Survey methods evaluation

All linear regression models for predicting female and male DNA-based reproductive success (actual gametic contributions) from reproductive investment (field surveys) were highly significant ($p < 0.001$), indicating that the surveys reflect the respective reproductive success profiles (see Table 6.1 and Figures 6.1 & 6.2 for female and male survey methods, respectively). Note that clonally collected seed with known female parentage (i.e., family array) can only provide information pertaining to the male component (western larch and western redcedar) while that from orchard's bulk permits inferences on both male and female components (Douglas-fir and lodgepole pine).

6.3.1.1 Female component

Results from survey methods F1, F2, and F3 in both lodgepole pine and Douglas-fir suggest that predictions of female reproductive success based on seed-cone volume (F2) and weight (F3) are superior to those based on cone-count (F1). In lodgepole pine, F2, and F3 produced similar and higher R^2 (0.651 & 0.657) as well as a lower $RMSE$ (0.036) as compared with F1's 0.607 and 0.039, respectively, indicating a better data fit and a lower error. In Douglas-fir, within the first group with the same sample size (F1, F2, and F3), the highest R^2 values (0.888 and 0.870) were also produced by F2 and F3, respectively, reflecting similarity between these two methods, while F1 produced a moderately lower R^2 value (0.823) (Table 6.1). Within the second group (F4 and F5), the difference between R^2 values was negligible (0.804 vs. 0.816; Table 6.1); however, the $RMSE$ -based across-group comparison shows that these two methods introduce a substantially higher variability and confirms the suitability of F2 and F3. It is important to note that the first three methods (F1, F2, and F3) rely on 100% sampling (i.e., the entire orchard) whereas F4 and F5 are based on a subset of the orchard's production (standard volume). Methods F2 and F3 therefore appear to be good indicators of female reproductive success as long as the number of seeds produced within a seed-cone is

proportionate to its size and weight, respectively. In this case, larger and heavier seed cones are assumed to yield more seed than their smaller and lighter counterparts while the seed yield of the latter is compensated by their increased number per unit volume. This is expected to hold true if pollen sources, irrespective of their origin (i.e., within or outside the orchard), are adequate to sire all receptive strobili and in the absence of differential insect-induced seed loss among parents. Although a discrepancy may exist between parental seed-cone crops and production of filled-seeds (referred to as reproductive energy and reproductive success, respectively, in El-Kassaby and Cook (1994)), the high congruence between seed-cone production (count, volume, and weight) and DNA-based reproductive success obtained in this study suggests that this issue is probably marginal and the impact on crop genetic parameter estimates is negligible. This statement can be supported by the high correlation between the visual assessment of cone count (F1) and the total number of filled seed as estimated from a random sample of 50 cones per parent in western larch ($r = 0.909$). Using the measurable seed-cone volume (F2) or weight (F3) seem to be more reliable than the subjective visual counts of F1 and one should also take into consideration the additional cost associated with the entire orchard's cone counting. Furthermore, although F2 and F3 produced similar results in terms of R^2 and $RMSE$ in the two studied species, the former can be considered more practical since volume is more easily determined during seed-cone harvest as compared to the additional weighing required for F3. The preceding section also indicates that the suitability of F4 or F5 is questionable unless all parents produce the minimum required volume of 5 L and thus their utilization will be limited to mature seed orchards and moderate or heavy seed years. These two methods will therefore not be further elaborated in this study.

In addition to the above-described methods proposed by Woods (2005), I wanted to investigate whether reproductive success could be predicted solely from parental representation (i.e., number of ramets per parent) without any prior adjustment for ramet crown volume or age (denoted as F0 here), an assumption that is the basis to a variety of approaches such as linear deployment of parents (Lindgren and Matheson 1986) and its subsequent modifications (Bondesson and Lindgren 1993; Lindgren et al. 2009). The results show that F0 produced a markedly worse data fit for both lodgepole pine and Douglas-fir as compared to any other tested method (Table 6.1). F0 yielded low R^2 (0.204) for lodgepole

pine as opposed to that produced by F2 (0.651) while this difference was even more pronounced in Douglas-fir (0.188 vs. 0.888) (Table 6.1). The regression for lodgepole pine further shows that prediction intervals are much wider than those based on survey methods (Figure 6.1a), a result of the relatively higher error associated with F0. Note that no figure is shown for Douglas-fir as the transformed data failed to conform to the assumption of normal distribution of residuals.

6.3.1.2 Male component

Predicting male reproductive success is expected to be difficult as compared to that of female as it is assumed to be a function of pollen-cone production. For the four studied species, the regression models significantly improved when information on full survey's number or volume of pollen cones (method M3) was utilized ($R^2 = 0.581, 0.545, 0.731$, and 0.606 for lodgepole pine, Douglas-fir, western larch, and western redcedar, respectively) as compared with the situation where male success was predicted solely from parental representation ($R^2 = 0.314, 0.420, 0.275$, and 0.084 , respectively) (Table 6.1). This finding is consistent with several other studies conducted on white spruce (Schoen and Stewart 1986), Douglas-fir (Burczyk and Prat 1997), Japanese black pine (Goto et al. 2005) and sugi (Moriguchi et al. 2007). As a result of lower *RMSE*, prediction intervals were also markedly narrower for M3, especially in western larch and western redcedar (see Figure 6.2a & Figure 6.2d for comparison), pointing out that its prediction power is also higher. Regression from method M2 where only a subset of ramets was scored (every other ramet) further indicates that a full pollen survey (method M3) may not even be necessary. High Pearson's product-moment correlations between M2 and M3 for lodgepole pine, Douglas-fir, western larch, and western redcedar ($r = 0.977, 0.930, 0.970$, and 0.948 , respectively) suggest that scoring just 50% of all ramets may provide reasonably accurate estimates of the total pollen production in seed orchards.

Similar to that reported for the female component, findings from this study imply that the assumption of successful pollination events being a function of parental representation (i.e., assumption that more trees produce more successful pollen) should be applied with care. On the other hand, it is worthwhile to note that a simple crown-volume adjustment (method M1) in western larch substantially improved both data fit and accuracy of the prediction as compared with M0 (R^2 : 0.499 vs. 0.275 and *RMSE*: 0.060 vs. and 0.072 ; Table 6.1).

Although no such records were collected on the other three species, the improved regression advocates the utilization of this approach as a suitable alternative in situations where no male fecundity data is available. It should be also noted that M1 was not prescribed for western larch by Woods (2005) and thus its considerably poorer result relative to M2 and M3 is not surprising.

Predicting of male reproductive success from female reproductive investment was found to be less reliable than from male fecundity as well. Fecundity scores from method F2 were chosen for illustration and while they provided an acceptably good fit in Douglas-fir ($R^2 = 0.323$), they were very unreliable in lodgepole pine ($R^2 = 0.071$), indicating that this approach is inconsistent and that genders should be considered separately. It is probably due to the sexual allocation of resources in hermaphroditic species (Charlesworth and Charlesworth 1981; Charnov 1982; Savolainen et al. 1993), causing weak or even negative correlation between male and female fecundities (Caron and Powell 1989; Kang and Lindgren 1999; Kang et al. 2004; Kjær 1996; Nikkanen and Velling 1987).

6.3.2 Genetic diversity

Effective status number is a good representation of the extent of genetic diversity (Lindgren et al. 1996; Lindgren and Mullin 1998). For its importance, it was integrated in some forestry jurisdictions such as British Columbia, where a minimum acceptable level was set at a value of 10 (Yanchuk 2001), as crops with smaller values were expected to lead to genetic vulnerability and lower resilience of established plantations. In this study, female (N_{ef}), male (N_{em}), and joint parental (N_{ep}) effective number of parents were determined using estimates of actual pedigree-reconstruction-based reproductive success and were used to evaluate the robustness of those derived from the variety of reproductive investment survey methods (Table 6.1).

Without exception, all female survey methods produced inflated N_{ef} estimates (Table 6.1). While the survey-based N_{ef} ranged from 39.49 to 46.00 and 11.55 to 15.15 for lodgepole pine and Douglas-fir, respectively, the actual DNA-based N_{ef} was estimated to be as low as 35.17 and 6.50. One possible explanation for this discrepancy is that the former is an approximation while the latter is a more accurate estimate. Any N_e estimates based on survey methods and thereafter predictions of the actual reproductive success are based upon the assumption that parents are in reproductive synchrony. However, an extremely early or

late reproductively active parent can be ranked as a strong contributor to a seed crop on the basis of its male strobili/cone production while its actual reproductive success may be low due to reproductive phenology asynchrony with other parents (El-Kassaby and Askew 1991; Roberds et al. 1991). This could be an issue in areas with slow heat sum accumulation where phenological differences tend to be exaggerated (Fashler and El-Kassaby 1987); however, reproductive phenology was singled out by Xie et al. (1994) as the factor with the lowest impact on crops' genetic parameters, specifically with greater phenology overlap. In the Douglas-fir seed crop, the most successful female produced 35.7% of seed (Lai et al. 2010) while it only produced 15.2, 18.5, and 17.2% of cones (F1), cone volume (F2), and cone weight (F3), respectively. This way almost 20% of orchard's actual seed production (the difference between the pedigree- and survey-based estimates of the reproductive success assigned to the most productive single female) was redistributed among other parents, which unintentionally balanced female gametic contributions estimated based on survey methods and inflated respective N_{ef} estimates. It is noteworthy to mention that the reproductive investment-success regressions for Douglas-fir (Figure 6.1b), in spite of their statistical significance, could not fully elucidate this situation. The drastic overestimation of the survey methods' N_{ef} observed in Douglas-fir, which more than doubled the actual DNA-based value (15.15, 13.74, and 14.36 for F1, F2, and F3 vs. 6.50; Table 6.1), indicates that the survey-based results should be interpreted with caution. On the other hand, one could assume that this situation will only occur when just one or a few parents dominate the seed-cone production. To support this hypothesis, I looked at lodgepole pine where the most successful female (7.5% DNA-based) produced 5.9 and 5.6% of the total seed-cone volume and weight, respectively, and observed that the DNA-based N_{ef} was only overestimated by 12.2 and 16.1% using methods F2 and F3 (Table 6.1).

However, no clear pattern could be identified for N_{em} bias. While method M3 (total orchard surveys) produced N_{em} estimates similar to those based on DNA analyses in Douglas-fir, western larch, and western redcedar (25.25 vs. 24.52; 16.96 vs. 18.55; 19.41 vs. 16.80, respectively), it inflated this estimate by 32.6% in lodgepole pine (53.83 vs. 40.59) (Table 6.1). In spite of this incongruity, considering the difficulty of predicting male reproductive success from the production of pollen buds or strobili, these results appear to be surprisingly good and support the recommendations by Woods (2005). Partial pollen surveys (M2)

produced similar or identical N_{em} estimates to those from M3 across the studied species (Table 6.1). This finding, along with the good fit statistics attained by the regression analyses, supports the notion that M2 could become the method of choice in estimating seed crops' genetic parameters as it provides a reasonable balance between the accuracy of results and costs associated with data collection.

As expected, F0 and M0, which assume no fertility variation among parents, produced highly inflated N_e estimates for both male and female components in all situations. In Douglas-fir, for instance, the full pedigree reconstruction proved that this assumption was drastically violated as two most reproductively successful females produced nearly half of the entire seed crop in 2005 (Lai et al. 2010) while they only represented 5.7% of the orchard's population, which resulted in the F0's estimate being nearly six fold bigger than that based on DNA analysis (38.51 vs. 6.50) (Table 6.1). As mentioned earlier, however, these two methods were only introduced with the aim to provide an assessment of the correspondence between their estimates with those based on pedigree reconstruction whereby the commonly assumed equality between parental representation and reproductive success could be evaluated. My findings show that they are least reliable of all of the tested methods and should not be applied to any of the four studied species.

When female and male reproductive investment estimates were considered collectively in determining parental N_e , the best female-male combination (F2-M3) produced relatively higher values (51.70 and 20.86) as compared to those of DNA analysis (46.45 and 13.30) (Table 6.1). This overestimation needs to be considered when crops' genetic diversity is calculated.

6.3.3 Genetic worth

Utilizing the best survey-combination's gametic contributions (F2-M3) produced similar genetic worth estimates to those based on DNA (16.01 vs. 14.84 and 9.84 vs. 10.01, respectively), indicating that this parameter is more robust than genetic diversity (Table 6.1). It should be noted, however, that data on pollen contamination were not utilized when the survey-based estimates were calculated because none of the tested survey methods has the capability of quantifying it. On the other hand, it was included in calculations of the DNA-based estimates because its rate was known from the pedigree reconstruction. (It was estimated to be 14.4% and 10.4% for lodgepole pine and Douglas-fir, respectively, and was

assigned a breeding value of 0.) Since pollen contamination reduces seed crops' genetic worth, then a penalization could be applied to the survey-based estimates when crops from seed orchards with a high potential risk of contamination are qualified.

To evaluate the impact of the commonly observed among-year fluctuations in parental gametic contributions (e.g., El-Kassaby et al. 1989; Kang 2000) on the sensitivity of the genetic worth estimates, the contributions were artificially varied while holding breeding values and pollen contamination rate constant for both lodgepole pine and Douglas-fir. This process was repeated 50 times and the seed crops' genetic worth was determined. For the F2–M3 combinations, the genetic worth was found to vary substantially in both lodgepole pine and Douglas-fir and averaged 14.69 (range 13.26–16.21) and 11.77 (range 9.84–13.46), respectively, indicating that year-to-year fertility variation may strongly affect genetic worth estimates. When the same process was repeated for DNA-based genetic worth estimates, values averaged 14.77 (range 13.08–15.95) and 11.83 (range 8.96–15.27), respectively. Lodgepole pine produced smaller ranges of genetic worth values than Douglas-fir (both survey- and DNA-based), which is due to the fact that the former had more evenly balanced gametic contributions across the parental population (see Table 6.1 for comparison with actual genetic worth estimates).

6.4 Conclusion

Methods F2 (seed-cone volume) and M3 (full pollen survey) provided the most reliable proxies to actual female and male components of parental reproductive success, respectively. A high correlation between M2 (partial pollen survey) and M3 suggests the use of the former because it does not compromise the estimates' accuracy and is associated with reduced efforts. My results indicate that seed crops' genetic worth estimates based on the combined F2-M3 survey method are representative of those based pedigree reconstruction; however, the observed differences from the gametic contribution manipulations show that they are year-specific and therefore should be assessed on a yearly basis. Additionally, the results highlight the impact of gametic contribution scale differences on genetic diversity estimates, specifically in the presence of extreme reproductive success distortion (i.e., the role of p_i and p_i^2 in estimating genetic worth and effective number of parents, respectively).

Table 6.1 Fit statistics (R^2 : coefficient of determination and $RMSE$: root mean square error) for linear regressions between female (F) and male (M) reproductive investment assessment methods and reproductive success as determined by microsatellite-DNA-based pedigree reconstruction and crops' genetic parameters (N_e : status effective number; GW : genetic worth; m : male; f : female; p : joint male and female (parental)) for Douglas-fir, lodgepole pine, western larch, and western redcedar seed orchards.

Species / Seed orchard	Gender	Sample size	No. of parents	Survey code	R^2	$RMSE$	N_{ef} N_{em} (survey)	N_{ef} N_{em} (DNA)	N_{ep} (survey)	N_{ep} (DNA)	GW_p (survey)	GW_p (DNA)
Lodgepole pine	♀	520	70	F0	0.204	0.055	57.81	35.17	51.70 ^{F2/M3}	46.45	16.01 ^{F2/M3}	14.84
				F1	0.607	0.039	46.00					
				F2	0.651	0.036	39.49					
				F3	0.657	0.036	40.83					
	♂	1030	70	M0	0.314	0.038	58.18	40.59				
				M2	0.532	0.031	54.35					
				M3	0.581	0.029	53.83					
Douglas-fir	♀	801	49	F0	0.188	0.097	38.51	6.50	20.86 ^{F2/M3}	13.30	9.84 ^{F2/M3}	10.01
				F1	0.823	0.045	15.15					
				F2	0.888	0.036	13.75					
				F3	0.870	0.039	14.36					
	♂	718	49	F4	0.804	0.052	12.07	24.52				
				F5	0.816	0.051	11.55					
				M0	0.420	0.053	38.51					
Western larch	♂	1848	41	M2	0.476	0.051	23.13	18.55				
				M3	0.545	0.047	25.25					
				M0	0.275	0.072	32.82					
				M1	0.499	0.060	30.48					
Western redcedar	♂	771	27	M2	0.478	0.056	18.65	16.80				
				M3	0.606	0.049	19.41					
				M0	0.084	0.075	23.61					

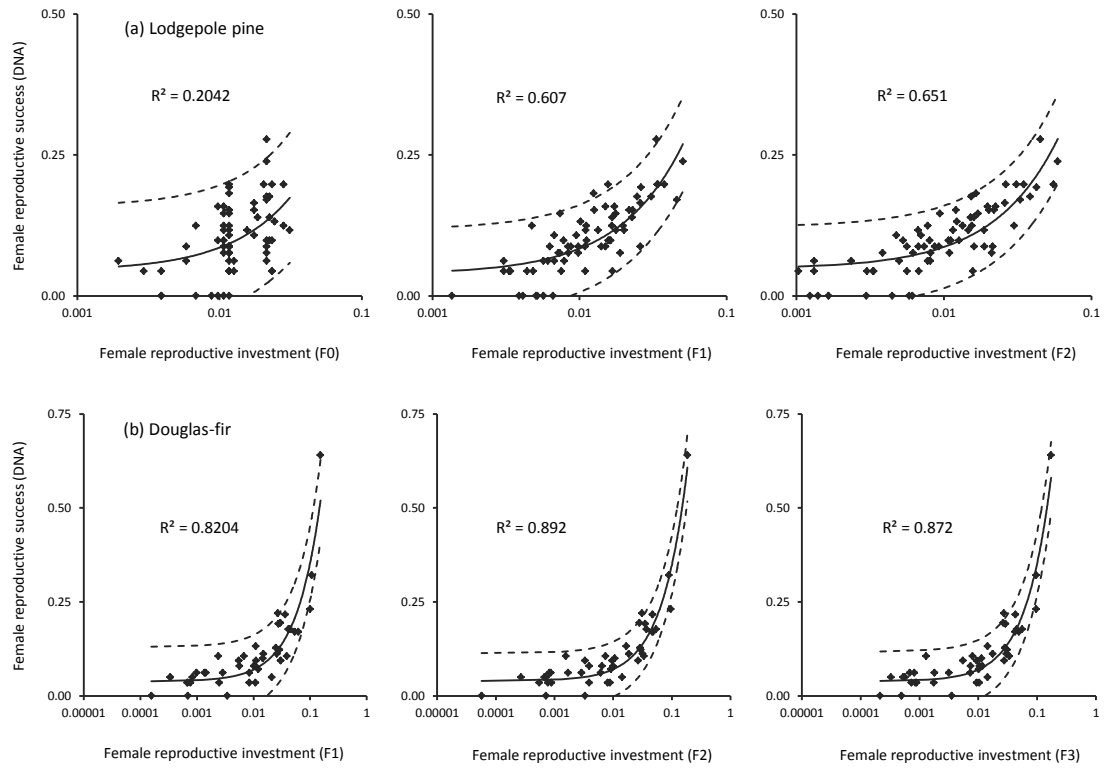


Figure 6.1 Linear regressions between female reproductive investment (survey methods F0–F3; see Section 6.2.2.1 for details) and transformed female reproductive success (microsatellite-DNA-based pedigree reconstruction) for lodgepole pine (a) and Douglas-fir (b). The three curves represent the linear regression line (solid) and 95% confidence limits for individual predictions (dashed).

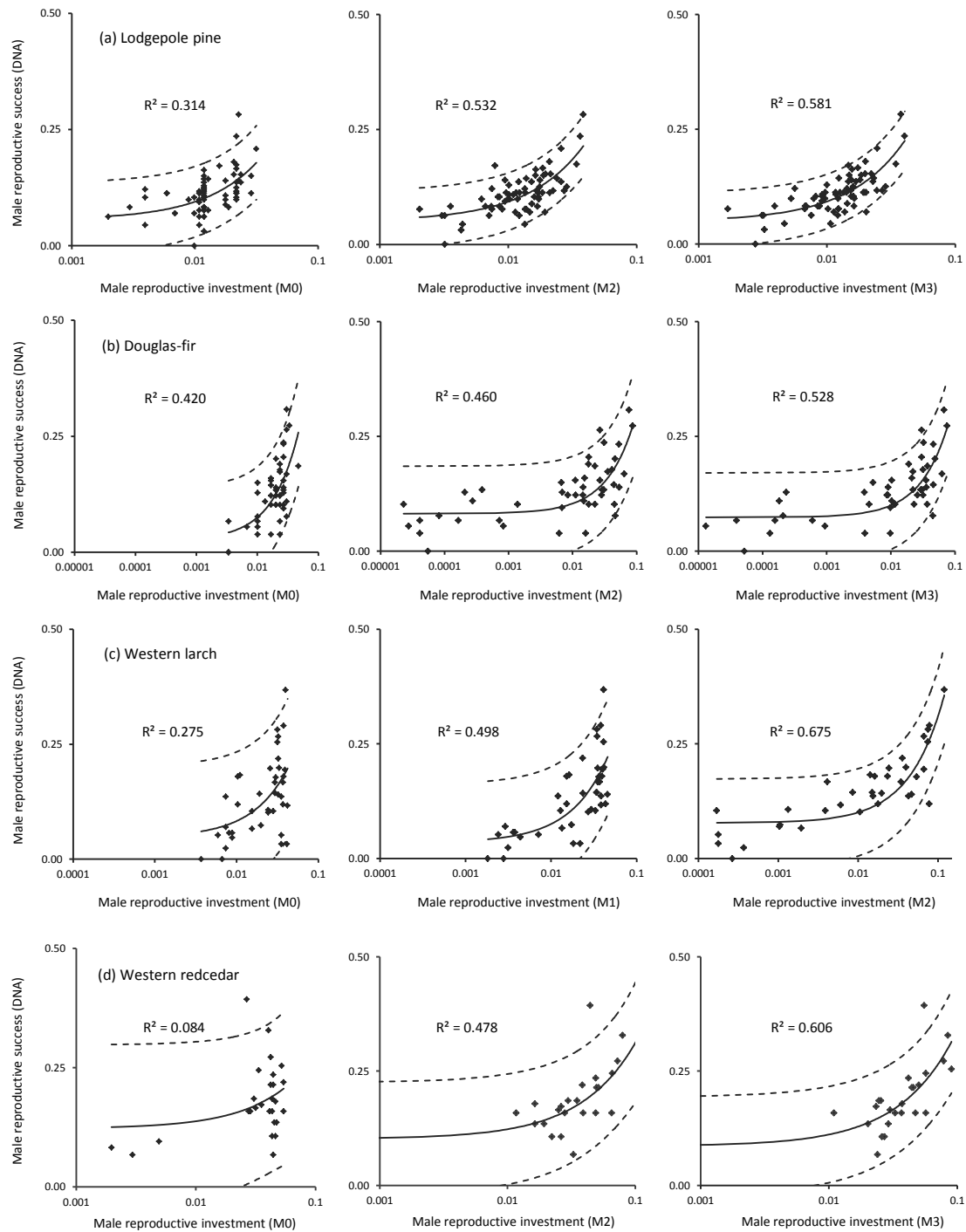


Figure 6.2 Linear regressions between male reproductive investment (survey methods M0–M3; see Section 6.2.2.2 for details) and transformed male reproductive success (microsatellite-DNA-based pedigree reconstruction) for lodgepole pine (a), Douglas-fir (b), western larch (c), and western red cedar (d). The three curves represent the linear regression line (solid) and 95% confidence limits for individual predictions (dashed).

7. Optimization of Combined Genetic Gain and Diversity for Collection and Deployment of Seed Orchard Crops

7.1 Introduction

Seed orchards are the main delivery vehicle for coniferous species' genetic improvement programs, in which balancing gain and diversity is of major concern. The level of genetic gain of a seed orchard's crop depends primarily on the genetic superiority of the selected parents, their actual gametic contribution to the resultant seed crop, and the level of gene flow from outside pollen sources and their respective genetic quality (Lindgren et al. 2004; Slavov et al. 2005a; Stoehr et al. 1998). The genetic diversity of a seed crop is greatly influenced by the magnitude of parental fertility variation (Burczyk and Chalupka 1997; Kjær 1996; Xie et al. 1994) as well as by the level of kinship among the orchard's parental population (Lindgren and Mullin 1998). Theoretically, maximum genetic gain and diversity for a given population of unrelated and non-inbred individuals would be attained if seed from only the best parent were collected and if all parents contributed equally to the seed crop, respectively.

The reproductive choreography – the timing, duration, and extent of reproductive activity of an orchard's parents – is of great importance because it affects the pollination patterns both within the orchard and between its parents and background pollen sources, thus affecting parental reproductive success and ultimately the genetic composition of resultant seed crops (Burczyk and Chalupka 1997; El-Kassaby et al. 1984; El-Kassaby and Ritland 1986b; Matziris 1994).

Variation in parental reproductive success is a commonly occurring phenomenon in seed orchards. It can be approximated using various simplified phenotypic assessment methods including visual scoring of pollen- and seed-cone production (Funda et al. 2011 (Chapter 6); Woods 2005) or estimated using a combination of DNA fingerprinting and pedigree reconstruction (Chapter 5; Funda et al. 2008 (Chapter 2); Lai et al. 2010; Moriguchi et al. 2004; Roberds et al. 1991; Slavov et al. 2005a; Xie and Knowles 1992). Since it may lead to over-representation of most productive parents (Kjær 1996) and in turn to accumulation of co-ancestry and ultimately to genetic erosion (Lindgren et al. 1996), it should be taken into consideration whenever seed crops' genetic parameters are estimated (Kang and El-Kassaby

2002). While it can be caused naturally by reproductive phenology asynchrony among parents or the simple variation in parental fecundities as observed in many forest tree species (see Section 1.5.1), it can also be manipulated deliberately e.g. by unequal deployment of parents (Lindgren et al. 2009; Lindgren and Matheson 1986), genetic thinning (Bondesson and Lindgren 1993), or by practicing selective seed-cone harvest (Kang et al. 2001b; Lindgren and El-Kassaby 1989). It is important to mention, however, that effectiveness of these practices will be greatly affected by the congruence between parental representation and reproductive investment on one side and success on the other as well as by the direction and strength of the correlation between their female and male components.

While parental representation was found to be the weakest predictor of parental reproductive success, parental fecundities (namely, full pollen survey and volume of clonally harvested seed-cones for the male and female component, respectively (Woods et al. 2005)), provided a reasonable approximation (Funda et al. 2011) (Chapter 6). Regarding the relationship between male and female success, most seed orchards' studies have shown that these two components are either weakly correlated (Caron and Powell 1989; Kang 2000; Kjær 1996; Nikkanen and Velling 1987) or independent (Kang and Lindgren 1999; Kang and Lindgren 1998; Kang et al. 2004; Savolainen et al. 1993). While the latter studies were based on visual assessment of fecundities only, a weak correlation, although positive, was also obtained using microsatellite-DNA-based full pedigree reconstruction (Chapters 4 & 5), indicating that the common practice of estimating the genetic crops' parameters based on one sex alone (e.g., seed production) (Kang and Lindgren 1999) may be inaccurate. Furthermore, due to the well-known fluctuations in parental fecundities among years (Kang 2000), the genetic composition of any given year's seed crop is likely to be unique, suggesting that the assessment of parental fecundities and manipulation with reproductive output should be conducted on a yearly basis.

The objective of this study is to develop a model that maximizes the genetic value of a selected portion of an existing orchard crop through the optimization of parental contributions. It is designed to account for relatedness among parents, an issue that is becoming increasingly important as advanced-generation seed orchards are established. The need to account for relatedness has been demonstrated in animal breeding (e.g., Grundy et al.

1998; Meuwissen 1997) and it is prudent to consider it also during the deployment of reproductive material from forest-tree production populations.

Reproductive output data (visual seed-cone count and pollen-cone volume scored on every ramet of each parent; methods F1 and M3, respectively (Woods et al. 2005)) from a western larch (*Larix occidentalis* Nutt.) seed orchard were used to illustrate the proposed approach and three different scenarios were tested: a) equal male and female fecundities (i.e., presence of a perfect correlation between female and male reproductive investment (see Lindgren et al. 2004; Prescher et al. 2006)); b) male contribution proportional to parental representation (number of ramets per parent); and c) actual male (M3) and female (F1) fecundities. A new approach that combines the above-mentioned elements in a single optimization step is presented. The optimization leads to determining optimum female contributions to new seedlots, after considering actual male and female reproductive output, parental breeding values, co-ancestry, and the desired level of genetic diversity, expressed as effective status number (Lindgren and Mullin 1998).

7.2 Materials and Methods

7.2.1 Seed orchard and fecundity assessment

The proposed optimization approach is illustrated using total male and female reproductive output records collected over three consecutive years (2004–2006) in the western larch clonal seed orchard (57 parents) described in Chapter 2. The genetic composition of the orchard maintained a dynamic state as a result of replacing lower-breeding-value parents by higher-ranking ones following progeny testing of the parental population. Parental breeding values (volume at age 60; provided by the B.C. Ministry of Forests, Mines and Lands) were predicted based on two series of 10-year-old progeny test trials located within the parental natural range (East Kootenay, B.C.). Female and male contributions to the orchard's total seed crop were predicted based on the number of seed-cones and volume of pollen-cones produced by every ramet of every parent, respectively (i.e., 100% survey). All parents in the orchard were assumed to be unrelated and non-inbred; however, this assumption was intentionally relaxed in several instances so that issues concerning build-up of co-ancestry in advanced-generation seed orchards could be addressed (details are provided at the end of this section).

Pearson's product-moment correlation coefficients were calculated for male and female fecundities within and among years to illustrate the commonly observed among-year variation in parental fecundity (Kang 2000) and thereby the difficulty of predicting future seed crops (Kroon et al. 2008; Lindgren et al. 2007). Several parents had to be excluded from the correlation analysis due to changes in the orchard's population structure over time. Correlation analysis was therefore only based on 27 parents (with the average of 35.6 ramets per parent) that were present in the orchard over the three consecutive study years.

7.2.2 Mathematical model

Let N be the number of parents in a seed orchard with corresponding breeding values provided in a vector \mathbf{X} . The average breeding value (genetic response) is

$$\Delta G = \sum_{i=1}^N x_i p_i \quad [7.1]$$

where p_i denotes the contribution of parent i to the orchard crop, and $0 \leq p_i \leq 1$, and

$$\sum_{i=1}^N p_i = 1 \quad [7.2].$$

The gametic contribution of parent i is split into its respective female (f_i) and male (m_i) components such that

$$p_i = \left(\frac{f_i + m_i}{2} \right) \quad [7.3]$$

where $0 \leq f_i \leq 1$, $0 \leq m_i \leq 1$, $\sum_{i=1}^N f_i = 1$, and $\sum_{i=1}^N m_i = 1$.

Genetic diversity of a seed orchard crop can be approximated by the effective number of parents, which describes the proportion of parents involved in the seed crop production. Considering related or inbred individuals, the concept of effective status number (N_e) (Lindgren et al. 1996) was used as

$$N_e = \frac{1}{2\Theta} \quad [7.4]$$

where Θ is group co-ancestry (Cockerham 1967). Group co-ancestry is defined as the average co-ancestry of all pairs of population members including individuals with themselves while co-ancestry is the probability that any two alleles sampled at random (one from each individual) are identical by descent (Malécot 1948).

Let us assume N parents are contributing to the resulting gametic pool (seed orchard crop). Selecting alleles randomly from the gametic pool, with replacement, the probability that the first allele originates from genotype i is p_i , and the probability that the second originates from genotype j is p_j . The likelihood that these two alleles are IBD is c_{ij} , which is the coefficient of co-ancestry between genotypes i and j (an element in the $N \times N$ co-ancestry matrix \mathbf{C}). Note that this coefficient is also referred to as coefficient of kinship (Falconer and Mackay 1996), which increases with the level of relationship. (In outbred populations, it equals to 1/8 for half-sibs, 1/4 for full-sibs and parent–offspring, and 1/2 for selfing.) The probability (Θ) that a pair of alleles sampled from the gamete pool is IBD is given by adding over all possible probabilities (Lindgren and Mullin 1998), thus

$$\Theta = \sum_{i=1}^N \sum_{j=1}^N p_i p_j c_{ij} \quad [7.5]$$

Assuming that all individuals are known starting from the founder population, then matrix \mathbf{C} can be recursively calculated from a given lineage (e.g. Emik and Terrill 1949). Any \mathbf{C} , if correctly specified and thence internally consistent, is a positive definite matrix with diagonal elements between 0.5 and 1 (the diagonal elements represent individuals' self co-ancestries), which is a prerequisite for its inclusion in the proposed optimization protocol.

7.2.3 Model constraints

The optimization's objective is to maximize the genetic gain in any given seed orchard crop (i.e., the maximization of the function in Eq. 8.1). Obviously, this would be reached if seed from the best parent or a limited number of parents with the highest breeding values were collected. However, in such a case, N_e of the resulting mixture would be unreasonably low and the seed crop could not be utilized due to its genetic vulnerability. Therefore, the minimum level of N_e was set as a constraint. This was derived out of Eqs. 8.4 and 8.5 as

$$\Theta \leq \frac{1}{2N_{e \min}} \Leftrightarrow \sum_{i=1}^N \sum_{j=1}^N p_i p_j c_{ij} \leq \frac{1}{2N_{e \min}} \quad [7.6]$$

where $N_{e \min}$ stands for the minimum desired N_e . This value is provided as an input to the optimization. If male contributions differ from those of females, Eq. 8.6 can be adjusted correspondingly (using Eq. 8.3) such that

$$\Theta \leq \frac{1}{2N_{e \min}} \Leftrightarrow \frac{1}{2} \sum_{i=1}^N (m_i + f_i) \sum_{j=1}^N (m_j + f_j) c_{ij} \leq \frac{1}{N_{e \min}} \quad [7.7]$$

A second constraint to the solution was given in Eq. 7.2.

7.2.4 Limits to gametic contributions

Furthermore, limits to female gametic contributions were set to reflect the amount of seed produced by each parent. Let s be a fraction of the total amount of seed crop collected in a given year (reflecting the actual seed need and the minimum desired level of diversity in this subset), where $0 \leq s \leq 1$ (value of 1 represents a situation where the entire seed crop is selected). The upper bound of the female gametic contribution of parent i was therefore set as the total amount of seed collected from the same parent (denoted as b_i) relative to the total amount of seed available in the entire orchard in a given year ($\sum_{i=1}^N b_i$):

$$0 \leq f_i \leq \min \left(1, \frac{b_i}{s \sum_{i=1}^N b_i} \right) \quad [7.8]$$

7.2.5 Optimization

One could study individual relationships encapsulated within the equations presented earlier; however, my goal was to combine all of them into a single model. This leads to a rather elaborate problem with many variables (each parent constitutes a variable) and corresponding dimensions. Due to the complexity of the problem, a mathematical programming approach was employed, whereby the objective function was maximized, subject to all variables' constraints.

The optimization software MOSEK[®] ApS (Anonymous 2002) was used to search for the optimum solution after declaring the problem in a correct mathematical programming format. Throughout the study, the default optimality tolerance level was used, meaning that the solutions presented and the actual optima could be considered identical. To assess the sensitivity to violations of various assumptions often used in similar studies (e.g. Kang et al. 2003b; Kang et al. 2001b; Lindgren et al. 2004; Prescher et al. 2006), three scenarios were compared. In Scenario A, male and female contributions were assumed to be equal (i.e., perfect correlation between male and female reproductive output) and only actual female seed-cone counts were used. In this case, the male component cancels out and p_i equals to

the seed production of the i^{th} parent. Note that this scenario is mathematically identical to the concept of “genetic thinning” implemented by Bondesson and Lindgren (1993). Next, under Scenario B, male contributions were considered to be a function of parental representation (the total number of ramets per parent) while the female contributions were based on actual seed-cone count. Finally, under Scenario C, actual estimates of both male and female gametic contribution were used.

Reproductive output data from one seed orchard of a single species was used to illustrate the benefits of this new optimization approach. It should be stated that although high correlation is assumed between reproductive investment (production of seed-cones) and reproductive success (production of living offspring) in this study, the optimization protocol does not rely on this assumption at all, as it can handle any reproductive output data available such as the number of seeds, filled seeds, germinated seeds or seedlings, depending on the resources invested during fecundity variation assessment.

In this study, $N_e \text{ min}$ was set to a value of 10 and 15 for illustration; however, N_e of 10 is generally considered to be the minimum N_e required to capture the majority of genetic diversity in a population (Yanchuk 2001). It is up to seed orchard managers to determine any minimum level of genetic diversity they desire, but this must be considered along with the crops’ genetic breadth. Since the study orchard consists of unrelated and non-inbred parents, a certain level of relatedness was assigned to 30 parents whereby a hypothetical advanced-generation orchard with co-ancestry was simulated (see the relationships below). This was conducted to demonstrate the impact of the relatedness on the optimum parental contributions and N_e estimates. Furthermore, the original first-generation breeding values were retained. Two examples with the total number of 44, 40, and 47 parents were tested (the seed orchard’s census numbers in 2004, 2005, and 2006, respectively), 30 of which were related (various levels of co-ancestry) and the remaining ones were unrelated. The first example consisted of 26 half-sibs and 5 full-sibs (moderate relatedness) while the second consisted of 13 half-sibs, 7 full-sibs and 15 parent-offspring (strong relatedness). All variables employed by the optimization protocol are listed in Table 7.1.

Table 7.1 List of variables used in the optimization protocol.

	Variable	Symbol
Input	Breeding values ¹	X
	Upper bounds set by parental seed production ¹	B
	Male gametic contributions ¹	M
	Co-ancestry among parents ²	C
	Selected fraction of total seed crop	<i>s</i>
	Minimum desired effective population size*	$N_{e\ min}$
Output	Optimum female gametic contributions (forming custom seedlots) ¹	F

¹ vector $N \times 1$, ² matrix $N \times N$ (N = number of parents), * N_e expressed as effective status number (Lindgren et al. 1996).

7.3 Results

7.3.1 Correlation

Over the three years, with the exception of female (2005–2006), all observed correlation coefficients between male, female, and male-female reproductive investment were positive and significant (Table 7.2). Correlations between male investments (Table 7.2, above diagonal) were, on average, slightly higher than those between their female counterparts (Table 7.2, below diagonal). These values imply that yearly male and female reproductive outputs are similar; however, close observation of the data on the individual ramet level indicated the presence of greater variability (data not shown). Male-female reproductive output correlation varied among years (Table 7.2), indicating that yearly crops are unique and should be evaluated on a yearly basis.

Table 7.2 Pearson's correlation coefficients between female (below diagonal), male (above diagonal), and between male and female reproductive investment (diagonal).

Year	2004	2005	2006
2004	0.62**	0.50**	0.60**
2005	0.36*	0.48**	0.77**
2006	0.83**	0.28 ^{ns}	0.62**

$N = 27$, r critical value = 0.330 ($\alpha = 0.05$) and 0.453 ($\alpha = 0.01$) (^{ns}not significant, *significant, **highly significant)

7.3.2 Genetic gain and diversity estimates

The genetic gain estimates of the orchard's crops (100% sampling) varied slightly over the three years, depending on the scenario applied (Figure 7.1). When only the female component was considered (i.e., assuming perfect positive correlation between male and female gametic contribution), the genetic gain estimates were either under- or over-estimated. The observed differences were caused by the relatively higher production of pollen from either high- or low-breeding-value parents, respectively (Figure 7.1: 100% sampling). The effective population size of the entire seed crop was found to be the highest in Scenario B (male reproductive output set as a function of parental representation), ranging from 27.3 to 33.8 for the three years, whereas the other two scenarios that take the male component into account reached similar but substantially lower values (Figure 7.1). The former approach is rather unrealistic because it assumes male reproductive output equality among ramets within parents which inflates N_e . Using Kang and Lindgren's (1999; 1998) method for estimating male fecundity variation among parents (the concept of sibling coefficient); Scenario B produced estimates of 1.23, 1.24, and 1.51 while Scenario C produced estimates of 2.50, 2.36 and 2.84 for 2004, 2005 and 2006, respectively. These estimates are high and indicative of extensive male reproductive output variability.

7.3.3 Optimization

The proposed protocol was illustrated by performing the optimization in a stepwise manner that excluded the cone crop in an increment of 10% while maintaining a minimum N_e of 10 (Figures 7.1 & 7.2a) and 15 (Figure 7.2b). An appreciable increase in genetic gain was attained by sacrificing some genetic diversity through relaxing the effective population size constraint. As expected, exclusion of seed from lower-breeding-value parents was associated with a steady increase in genetic gain across the three scenarios over the study years.

Under no relatedness (Figure 7.1), Scenario A produced a relatively steeper gain curve compared to Scenarios B and C. This result can be attributed to the fact that genetic gain calculations for the former were restricted to the subset of parents from which seed had been selected while male contributions from the remaining parents were not taken into account. It is important to note that N_e of 10 was attained with fewer seed parents under Scenarios B and C due to the inclusion of all pollen contributors. That means that the genetic diversity

reduction associated with focusing on fewer females is counteracted by the inclusion of many parents as males. For instance, for the 2006 crop, seed from as few as seven parents would secure the attainment of $N_e = 10$ when 20% of the crop was selected (Figure 7.1), thus demonstrating the benefits of including male reproductive output in the genetic diversity calculations.

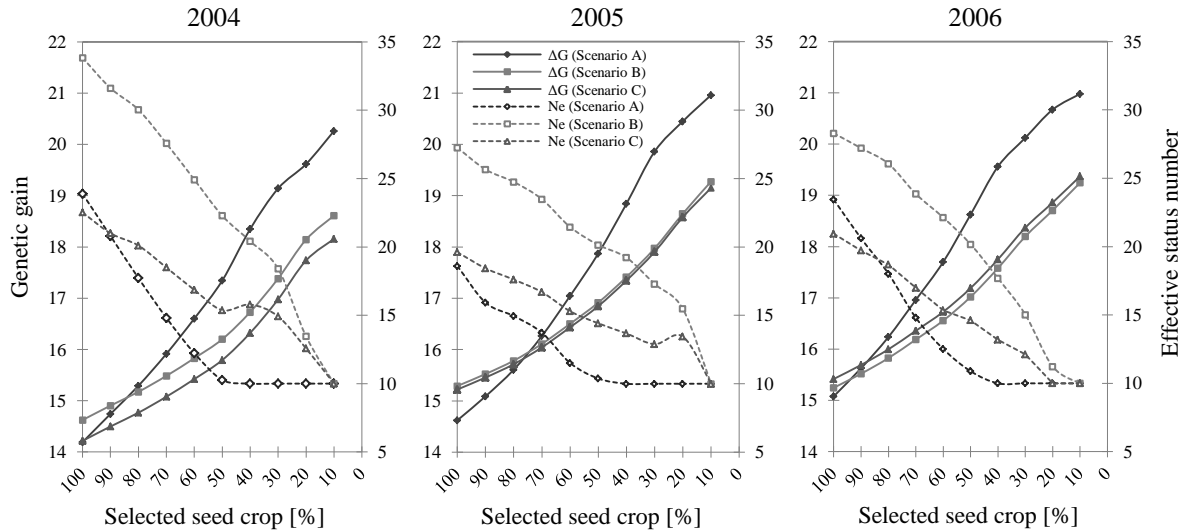


Figure 7.1 Estimated genetic gain (left axes, solid markers) and corresponding effective status number (right axes, open markers) in the western larch seed orchard over three years (2004–2006) after a stepwise exclusion of lower-breeding-value seed constrained by $N_{e\min} = 10$. The three pairs of curves in each year represent results from the three different scenarios applied (A- male and female reproductive equality, B- male as a function of parental representation, and C- actual data).

Relatedness among parents drastically affected the crops' genetic gain and diversity estimates (Figures 7.2a & 7.2b), although it should be noted that the genetic gain is only linked to this quantity indirectly. As expected, the higher the genetic relatedness among parents, the greater the reduction in the expected diversity estimates. For instance, when strong relatedness was assigned to the orchard's parents as defined in Section 7.2.5, N_e of the entire crop (100% sampling) dropped from 22.5, 19.6 and 20.9 to 15.5, 12.8, and 14.5 for 2004, 2005, and 2006, respectively. Further reduction was associated with the crop proportion selected by the optimization and N_e eventually reached a value of 10 and 15 (Figures 7.2a & 7.2b, respectively), depending on the constraint set prior to optimization. Conversely, the genetic gain estimates increased with the decrease in the crop proportion

selected (i.e., the inclusion of seed from fewer good parents). When the constraint of the effective population size was further relaxed ($N_{e\ min} = 10$), the effect of relatedness on the attained gain was only restricted to small proportions of the selected seed ($s \cong 0.1$; Figure 7.2a), i.e., the optimum proportions were equal regardless of relatedness present in the orchard. However, under higher N_e ($N_{e\ min}=15$; Figure 7.2b), the genetic gain reduction became apparent along the whole range of the X-axis (curves became more divergent), indicating that the increase in genetic diversity forced the inclusion of additional female parents with lower breeding values. Note that under strong relatedness, when N_e was set to 15, no solutions could be attained when more than 70% (2005) and 90% (2006) of total seed crops was retained. This situation generally arises when all options of manipulating parental proportional representations are attempted without reaching any combination yielding $\geq N_{e\ min}$ (i.e., the preset N_e exceeds a particular crop's biological limits) (Figure 7.2a).

As expected, a gradual selection of smaller subsets of seed led to a steady increase in gain in all cases tested, i.e., across all scenarios and years (Figures 7.1, 7.2a & 7.2b). However, although the genetic gain and diversity are known to be antagonistic, an increase in the former does not necessarily have to be associated with a decrease in the latter (see Figure 7.2a). For instance, the 2004 crop shows a situation where N_e fluctuates across different seed-retention ratios. Although this may seem surprising, it can be simply explained as a result of the gametic composition of the “best” 40% of seed, where the effective population size is just a by-product of the optimization process, as long as it meets the preset constraint. The described examples are strictly intended to illustrate the optimization procedure, and the results are merely a representation of this particular orchard with its simulated co-ancestry, its observed reproductive output for the three studied seed crops, and the assumptions made (i.e., scenarios applied).

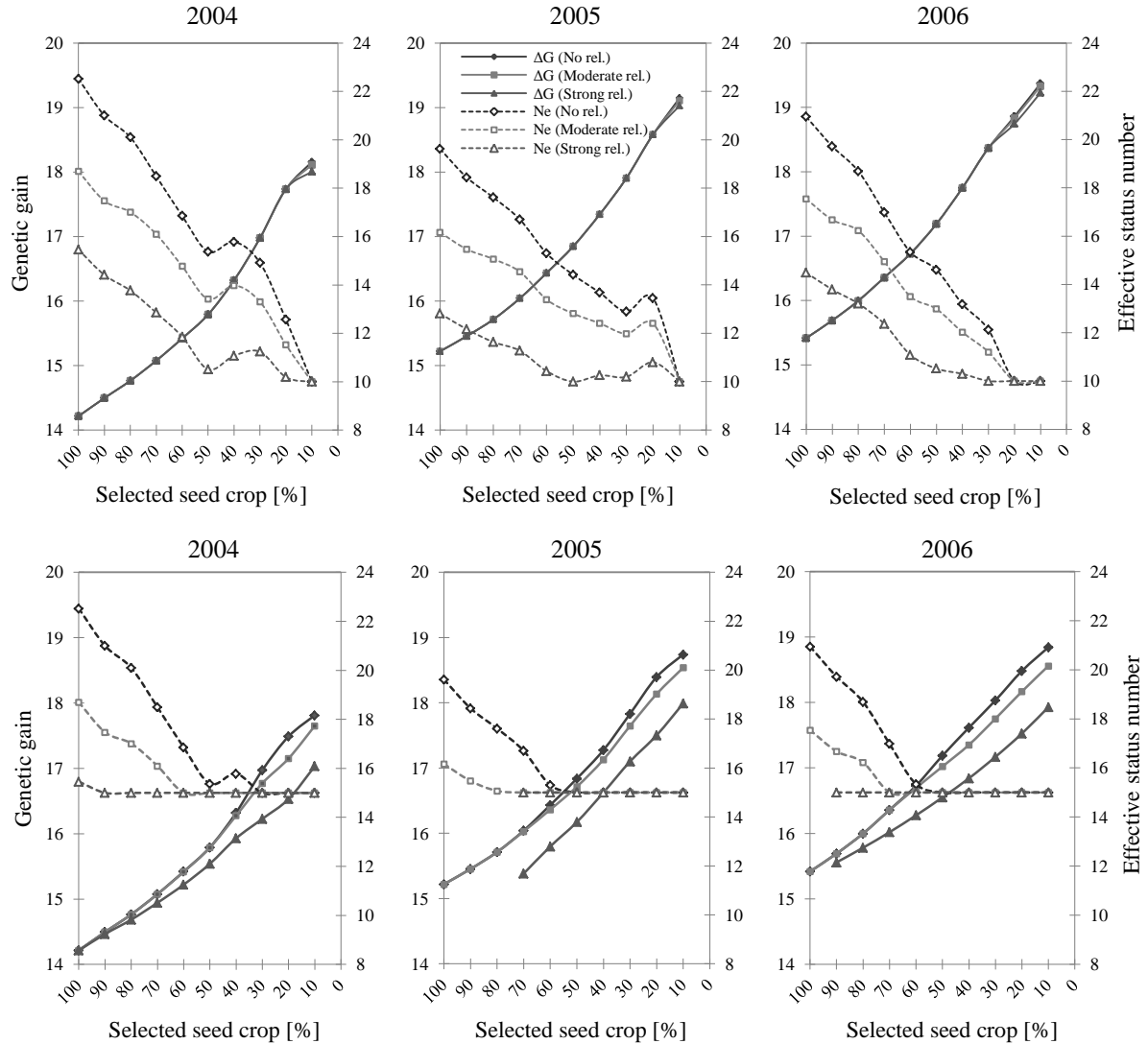


Figure 7.2 Estimated genetic gain (left axes, solid markers) and corresponding effective status number (right axes, open markers) in a western larch seed orchard over three years (2004–2006) after a stepwise exclusion of lower-breeding-value seed considering breeding values and relatedness following Scenario C and constrained by $N_{e\min} = 10$ (top) and $N_{e\min} = 15$ (bottom). The three pairs of curves represent no, moderate, and strong relatedness among parents as specified in Section 7.2.5 (hypothetical scenario).

7.4 Discussion

The simultaneous optimization of genetic gain and diversity is the subject of intense research in animal breeding circles (e.g., Meuwissen 1997). Forest geneticists recognized this concept and efforts have been made to simultaneously optimize gain and diversity of seed orchard

crops (Lindgren et al. 2004; Son et al. 2003). These efforts include linear deployment of parents proportionally to their breeding values in newly established orchards (Lindgren and Matheson 1986) or the application of genetic thinning for removal of lower-breeding-value parents from established orchards in a manner that approximates linear deployment (Bondesson and Lindgren 1993; Prescher et al. 2008).

Another option to balance genetic gain and diversity is selective seed-cone harvesting (Kang et al. 2001b; Lindgren and El-Kassaby 1989). This approach has a positive impact on the genetic gain attained in the resulting seed crops and it is also expected to maintain high diversity levels since no control is imposed on pollen donors. When selective seed-cone harvesting is implemented to mimic linear deployment, the amount of seed collected from each parent is linearly proportional to its breeding value, provided the following assumptions are met: (1) all parents are unrelated and non-inbred (i.e., the co-ancestry matrix is of the form $0.5\mathbf{I}$, where \mathbf{I} is an identity matrix), (2) mating is random, (3) reproductive output is proportional to parental representation, and (4) parental male and female gametic contributions are equal. In other scenarios (parents related and/or inbred, unequal parental male/female gametic contributions, and presence of constraints due to variation in seed production), the optimum contributions become progressively more non-linear as these factors are cumulatively added to the model. The generalized deployment method presented here thus covers both linear and non-linear solutions.

Bila et al. (1999) and Kang and El-Kassaby (2002) suggested constructing crops with equal female contribution (i.e., equal harvesting) to maximize the genetic diversity. It should be noted, however, that a crop genetic diversity is at maximum only if parents are unrelated and non-inbred, or strictly speaking, if there is no variation in self or pair-wise co-ancestries. Furthermore, this method is limited by the seed production of the least fertile parent, thus making it inappropriate because substantial amount of seed from more fertile parents would have to be abandoned (Kang et al. 2003b). It should be stated that both the selective seed-cone harvesting by Lindgren and El-Kassaby (1989) and attempts to equalize parental female reproductive contributions (Bila et al. 1999; Kang and El-Kassaby 2002) can be considered as special cases of the proposed generalized optimization approach; however, neither of them takes advantage of information on parental reproductive investment.

As mentioned earlier, the commonly observed parental imbalance among seed orchard parents (Askew 1988; Bila et al. 1999; Bilir et al. 2002; El-Kassaby and Cook 1994; El-Kassaby et al. 1989; El-Kassaby and Reynolds 1990; Eriksson et al. 1973; Griffin 1982; Hansen and Kjær 2006; Kang et al. 2004; Kjær and Wellendorf 1998), coupled with the observed fecundity variation over the orchard's developmental stage, environmental conditions, and management practices (El-Kassaby et al. 1989; Lindgren et al. 2004), led me to the conclusion that no two crops from one orchard are genetically the same. The observed among-year reproductive output correlations in this study are consistent with Kang's (2000) findings in a Japanese red pine seed orchard and show that reproductive output assessments made on existing orchard for any particular year or even over multiple years are unreliable in predicting the genetic composition of future seed crops.

To demonstrate the concept of the presented optimization model and its features, reproductive output data from the studied orchard (2004–2006) were used to produce new optimum seedlots (portions of the entire crop) which encompass the maximum possible genetic gain given the genetic diversity constraints under three modes of relatedness (no, moderate, and strong relatedness). Note that while the optimization produces various proportional contributions on the female side (i.e., the amount of seed from a particular parent entering the seedlot), the male component is also simultaneously considered. Furthermore, although the proposed approach results in the utilization of only a subset of the seed crop based on a preset culling criterion, the unused portion of the crop is still usable and its properties in terms of the genetic gain and diversity could be estimated. For instance, the remaining seed after selective harvesting through optimization attained the required minimum N_e of 10 in nearly all possible cases (results not shown). If seed selection is done to attain high genetic gain relative to the original crop, then the remaining seed portion will be substantial and its genetic diversity will be high due to its large number of contributing female and male parents.

As to comparing the three approaches tested in the present study, there seems to be a substantial difference between Scenario A on one hand and Scenarios B and C on the other (see Figure 7.1 for details). Notwithstanding the significant male-female reproductive output correlation, it appears that consideration of the male and female gametic contribution into the optimization is relevant for obtaining more reliable estimates of genetic gain and diversity.

The observed small differences between Scenarios B and C in 2005 and 2006 may suggest that estimating pollen production solely based on parental representation seems to be sufficient and relatively accurate for these two years; however, it can be misleading as in 2004, highlighting the caveat of this assessment method (Figure 7.1).

The optimization protocol utilizes data on parental reproductive investment that is used as a proxy to reproductive success. Since it is developed to maximize gain at a desired diversity when crops are harvested by parents, then any discrepancy between reproductive investment and success is confined to the male component. Evidence of the correspondence between pollen production and siring success are becoming available when genetic markers are used (Burczyk and Prat 1997; Funda et al. 2011 (Chapter 6); Goto et al. 2005; Hansen and Nielsen 2010; Schoen and Stewart 1986).

Furthermore, the protocol assumes no pollen contamination. However, even if the rate of pollen contamination is known, then its consideration as input into optimization only appears relevant if parents show large phenological differences. Xie et al. (1994) evaluated the effect of reproductive attributes in seed orchards on estimating N_e and concluded that male and female outputs had greater importance than other attributes, including phenology. The authors also point out that the role of reproductive phenology becomes pronounced only in small orchards and under substantial asynchrony. When pollen contamination is expected to be uniform across all parents (El-Kassaby et al. 2010), the resultant optimum contributions to any designed seedlot remain unchanged and the genetic gain is calculated *a posteriori*. It must be stated, however, that the genetic gain obtained by the optimization protocol is still at the maximum possible value under any pollen contamination rate.

Most studies concerned with estimating genetic gain and diversity of seed orchard crops have been limited to first-generation seed orchards. However, it has recently become obvious that omitting genetic relatedness among parents is no longer sustainable in seed orchard management and that an improved theoretical basis for dealing with advanced-generation material is needed (Prescher 2007). For instance, Bila et al. (1999) demonstrated a rapid build-up in group co-ancestry over generations in a simulation study on a teak stand where fertility variation differed across scenarios but was kept constant over generations. The steepest increase in group co-ancestry was observed when both female and male fertilities

varied, while the slowest was observed when fertilities were assumed to be equal across all parents in the stand, resulting into a considerably steeper decline in N_e in the former instance.

7.5 Conclusion

This optimization approach employs a reduced number of assumptions compared with other approaches by predicting genetic gain and diversity of seed orchard crops from parental reproductive investment rather than from parental representation. It is suitable for both first- and advanced-generations seed orchards, as it is capable of limiting individual female and male gametic contributions, while accounting for relatedness among parents and inbreeding. Owing to its ability of creating custom seedlots with multiple gain levels, I expect that it will be particularly advantageous in situations where seed crops exceed need or where more extensive site management will be applied (matching specific sites productivity). The surplus unused seed can still be utilized if it meets minimum genetic diversity requirements, and in situations where mixing of surpluses from multiple years is possible. In these cases, the same optimization protocol can be used. Furthermore, this protocol can be applied in situations where a given seed crop does not meet the minimum desired level of genetic diversity. This is likely to occur in rogued or advanced-generation seed orchards where the number of parents is limited and where group co-ancestry and fertility variation are high. An added advantage of this approach is the possibility to design seedlots with levels of genetic diversity exceeding that of the total seed crop (N_e of 10 vs. 15; Figures 7.2a & 7.2b). For instance, under strong relatedness, the entire seed crop's N_e of 12.8 was increased to 15 after excluding as little as 30% of the total crop (Figure 7.2). Finally, this optimization could extend the lifespan of older orchards through the formation of seed crops with genetic gain exceeding orchards' averages.

8. Optimization of Genetic Gain and Diversity in Seed Orchard Crops Considering Variation in Seed Germination

8.1 Introduction

The genetic quality of seed orchard crops is determined by the genetic gain and diversity they harbor. A concept of quadratic optimization was introduced by animal breeders (Meuwissen 1997; Villanueva et al. 2006; Villanueva and Woolliams 1997) to balance these two parameters in populations of related individuals, i.e., maximize genetic gain at a desired diversity level. Following this concept, Funda et al. (2009) (Chapter 7) developed an optimization protocol for maximizing genetic gain in seed orchard seedlots that collectively considers within-orchard actual female and male reproductive output, co-ancestry among parents, and inbreeding. This model assumes that parental reproductive success is a function of reproductive investment, i.e., that the production of seed and their number sired by a particular male parent is correlated with their respective production of seed and pollen cones (Funda et al. 2011) (Chapter 6), and that the rate of pollen contamination is uniform across the parental population (El-Kassaby et al. 2010).

Since the female reproductive output is usually assessed as seed-cone or seed production, then one could expect that even if the seed yields are assessed correctly, the common among-parent variation in seed germination (Bramlett et al. 1983; El-Kassaby et al. 1993b) could add a new dimension to the proper derivation of the female reproductive success and thus could influence the estimates of seed crops' genetic gain and diversity. El-Kassaby (2000) and El-Kassaby and Thomson (1996) have clearly shown that germination capacity (% germination) and speed (number of days needed for 50% of the viable seed to germinate) were the most important factors affecting parental genetic representation and consequently the genetic gain and diversity of orchards' seedlots.

The objective of this study is to investigate the effect of among-parent variation in seed germination on the optimum solutions produced by the optimization protocol and to elucidate the benefits of this extension to seed orchard management and nursery practice.

8.2 Materials and Methods

8.2.1 Seed orchard and parental fecundity assessment

The same seed orchard as described in the previous chapter provided material for this study. Parental (female and male) reproductive success was approximated from seed- and pollen-cone production (volume and count, respectively) scored in April 2005, using methods F2 and M3 as described in Woods (2005). Seed cones representing each of 37 seed-producing parents were collected and seed was extracted following common practices.

8.2.2 Seed germination

Parental seed germination was determined following the Int. Seed Test. Assoc. (1985) prescriptions (pre-treatment consisted of 24 hours water soak, followed by 22 days stratification period at +4°C, and germination was conducted at alternating 30/20°C for 8 hours light and 16 hours dark, respectively, for 21 days). Each female parent was represented by four replicates of 100 seeds. Germination capacity was estimated following El-Kassaby et al. (2008) and was used in the optimization protocol along with several vectors simulated from a Poisson distribution to obtain a low mean germination capacity ($\mu = 0.1, 0.2, \text{ and } 0.3$).

8.2.3 Optimization

The impact of among-family variation in germination capacity on optimum solutions (i.e., optimum female contributions) obtained by the optimization protocol and genetic gain and diversity estimates was evaluated. Two different scenarios were tested: in Scenario 1, the germination capacity was not included in the optimization and the amount of viable seed (i.e., seed that actually germinated) as well as the genetic gain and diversity estimates were calculated *a posteriori* whereas in Scenario 2, the total amount of available (harvested) seed (vector **B**) was adjusted prior to optimization whereby the variation in germination capacity was accounted for during the optimization process. The protocol employed in this study was similar to that presented in Sections 7.2.2–7.2.4; however, in Scenario 2, Equation 7.8 had to be modified to accommodate the germination capacity variation such that

$$0 \leq f_i \leq \min \left(1, \frac{b_i g_i}{s \sum_{i=1}^N b_i g_i} \right) \quad [8.1]$$

where g_i represents the germination capacity of parent i and $0 \leq g_i \leq 1$. Variables used in this extension to the original model are listed in Table 8.1.

Table 8.1 List of variables used in the optimization protocol.

	Variable	Symbol
	Breeding values ¹	X
	Upper bounds set by parental seed production ¹	B
	Parental seed germination capacity ¹	G
Input	Male gametic contributions ¹	M
	Co-ancestry among parents ²	C
	Selected fraction of total seed crop	s
	Minimum desired effective population size*	$N_{e\ min}$
Output	Optimum female gametic contributions (forming custom seedlots) ¹	F

¹ vector $N \times 1$, ² matrix $N \times N$ (N = number of parents), * N_e expressed as status effective number (Lindgren et al. 1996).

The simplified reproductive investment assessment methods developed for predicting female and male reproductive success (Woods 2005) have been shown to provide an acceptable balance between their costs and information accuracy (Funda et al. 2011) (Chapter 6). In this study, the parental reproductive success was approximated by seed- and pollen-cone production. Although female reproductive success approached based on seed-cone volume may differ substantially from that based on the amount of seed (El-Kassaby and Cook 1994) or that derived from pedigree reconstruction, here I only intended to show the effect of including the germination capacity in the optimization model. In general, any female reproductive output data available can be used.

8.3 Results and Discussion

8.3.1 Seed germination

Results from the seed germination test show that differences in germination capacity among the tested 37 western larch families were significant. Family accounted for 64.8% of the total variance (i.e., broad sense heritability was 0.65), which also indicates that western larch germination capacity is under strong, maternally driven genetic control (Table 8.2 & Figure 8.1). This finding, along with those obtained by numerous studies on other species (Bramlett

et al. 1983; El-Kassaby et al. 1993b), justifies the inclusion of this parameter in the optimization.

Table 8.2 Germination capacity variation among 37 western larch parents.

Source of variance	Degrees of freedom	EMS	Germination capacity ¹
Parent	36	$\sigma_e^2 + 4\sigma_{fam}^2$	64.8*
Residual	111	σ_e^2	35.2

¹Percent variance component, *significant at $\alpha = 0.01$

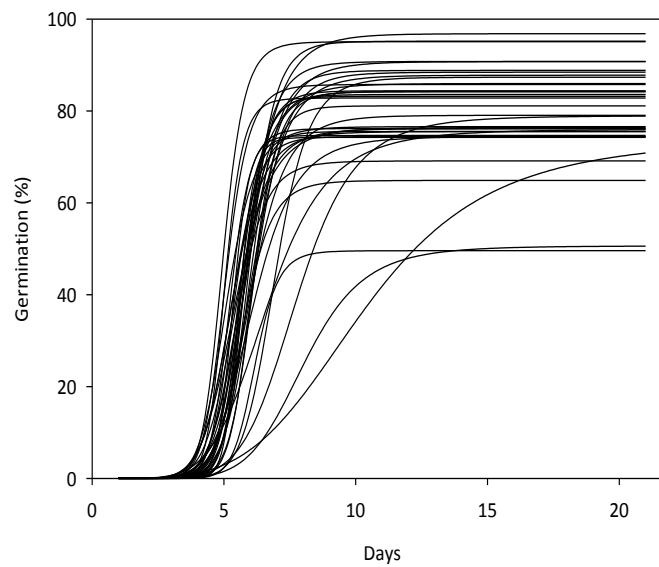


Figure 8.1 Germination curves for 37 western larch parents fitted using the four-parametric Hill function (El-Kassaby et al. 2008).

8.3.2 Genetic gain and diversity

The effect of among-family variation in germination capacity on the estimates of both genetic gain and diversity was found to be negligible. When the actual germination data was considered, the joint female and male (parental) effective status numbers (N_e) dropped slightly from 20.14 and 20.11 to 19.45 and 19.85, respectively, but the genetic gain remained almost unchanged (15.28 vs. 15.21). These small differences were probably mainly due to the small variation in the germination capacity (Figure 8.1) as well as the limited range of parental breeding values (10 to 26). Under a simulated germination (mean germination capacity of 10%), but keeping the actual breeding values and reproductive output data

unchanged, the N_e and genetic gain reached 19.59 ± 0.11 and 15.29 ± 0.02 , respectively, pointing out that the estimated genetic parameters are robust to the variation in the germination capacity and that the differences are small.

8.3.3 Optimization

As expected, the genetic gain estimates gradually increased as smaller fractions of the “most valuable” seed in terms of the objective function and the three constraints to optimization were retained. Using actual germination data (Figure 8.2), the genetic gain curves produced by the optimization completely overlapped regardless of the minimum N_e set prior to optimization (curves corresponding to $N_{e \min}$ of 15 and 20 are shown), indicating that the gain estimates were not at all affected by the variation in seed germination. This further implies that the original approach, which does not account for this additional parameter, is robust to producing reliable genetic gain estimates of the designed seedlots.

On the other hand, N_e estimates could be affected substantially, which would result in the production of seedlots with either under- or overestimated genetic diversity. For instance, when the variation in germination capacity was not included in the optimization, N_e of the best 50% of the total seed crop only reached 14.5 although a value of 15 had been set to the model as a constraint (Figure 8.2, point-dashed line, square markers). This means that the resultant seedlot's N_e , after the production of seedlings, could be lower than required ($N_{e \min} \geq 10$). Similarly, N_e of the best 40% reached 20.5 although a value of 20 had been constrained (Figure 8.2, point-dashed line, triangle markers). Considering the fact that any increase in diversity forces the inclusion of reproductive material from additional female parents with lower breeding values and thus results in a decrease in the genetic gain of the designed seedlot, then the minimum desired genetic diversity should be attained precisely whereby more space could be dedicated to improving the gain. In this situation, more gain was sacrificed to genetic diversity than was actually necessary.

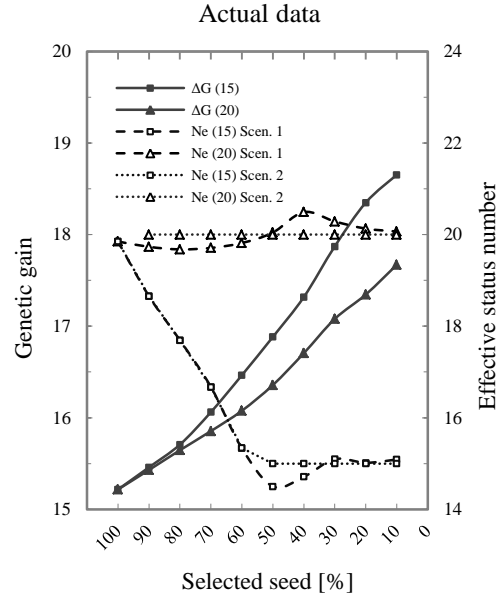


Figure 8.2 Genetic gain (ΔG ; left axes, solid markers) and effective status number (N_e ; right axes, open markers) estimates from a western larch seed orchard crop after a stepwise (10%) exclusion of seed. The two marker shapes, square and triangle, represent $N_{e \min}$ constraints of 15 and 20, respectively, and the point and line dashing represent the posterior adjustment of the genetic parameters due to variation in germination capacity following Scenario 1 and the inclusion of the germination capacity in the optimization model following Scenario 2.

In Figure 8.3, the effect of the variation in germination capacity was even more accentuated. While the genetic gain curves between the two scenarios were just negligibly diverse, N_e estimates of the new seedlots under Scenario 1 differed substantially across the whole range of the X axis (range 18.7–20.8; see Figure 8.3 – left), depending solely on the actual vector of the germination capacity. When the germination capacity was included in the optimization (Figure 8.3 – right), the final value of N_e always equaled to 20, no matter what germination capacity was used. This discrepancy was a result of the fact that, in the former case, the optimum solutions (i.e., optimum female contributions forming a designed seedlot) were produced regardless of how many seeds would actually develop into seedlings. As mentioned earlier, this situation will generally occur whenever variation in seed germination among orchard parents is present.

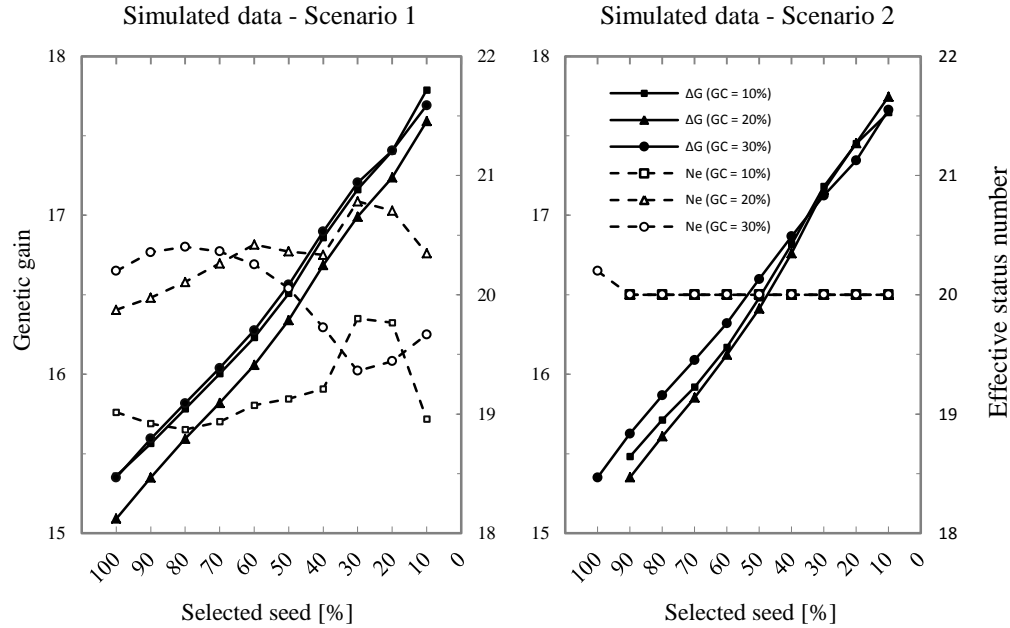


Figure 8.3 Genetic gain (ΔG ; left axes, solid markers) and effective status number (N_e ; right axes, open markers) estimates from a western larch seed orchard crop after a stepwise (10%) exclusion of seed. The three marker shapes, square, triangle, and circle, stand for simulated mean germination capacities of 10, 20, and 30%, respectively, and N_e is constrained by the value of 20 for illustration. The left and right parts of the figure are based on Scenarios 1 and 2, respectively.

8.4 Conclusion

Results from this study have shown that small among-family variation in seed germination capacity had a negligible effect on the optimum genetic gain–diversity solutions. However, when it was substantial, the genetic diversity was either under- or overestimated, indicating that it is worthwhile to include this additional parameter in the optimization protocol.

9. Conclusion

9.1 Thesis Summary

The goal of this thesis has been to deepen knowledge of the population genetics of artificial forest tree production populations, known as seed orchards. In theory, a seed orchard's offspring generation, represented by a seed crop, will reflect the genetic composition of the selected parental population; in other words, the population will be in Hardy-Weinberg equilibrium whereby desired genes will be effectively transmitted from the parents to the offspring. This would ensure that selection of superior individuals conducted in previous stages of tree improvement will be effective and that expected genetic gains will be realized. Considerable time and expenses have been dedicated to studying seed orchard populations in the past because seed orchards are responsible for production of genetically improved material used for re-establishment of forest stands. It was however not surprising that the ideal scenario was never fulfilled and that seed orchard populations were found to significantly deviate from Hardy-Weinberg equilibrium, causing bias in the crops' anticipated genetic parameters. Such deviations have been observed in seed orchards and natural populations of many species and can actually be considered as the norm rather than an exception. This phenomenon is a result of non-random union of gametes or variation in parental reproductive success, which is due to a variety of factors including simple differences in parental male and female fecundities, reproductive phenology asynchrony among parents, and gene flow.

All population genetic studies conducted in seed orchards so far have focused exclusively on the male component of the reproductive success, investigating pollination dynamics among parents within a population and among replicates within a particular parent, or on quantifying gene flow from additional pollen donors or background sources to determine pollen contamination. Regardless of genetic markers used, these studies were limited to partial pedigree reconstruction of seed crops only, because an equal or similar number of seeds collected from an array of known maternal parents (e.g., family array) was analyzed, allowing only the male component of parentage to be inferred without bias.

In this thesis, I studied pollination dynamics of four different conifer species' seed orchards (western larch, Douglas-fir, lodgepole pine, and western redcedar) and confirmed that

reproductive success is greatly imbalanced among parents, as the top 20% of males contributed 56, 54, 45, and 46% of successful within-orchard pollen.

Gene flow (pollen contamination), one of the greatest constraints to the proper functioning of seed orchards, varied markedly across the four studied seed orchards and while it was estimated to be moderate to low in Douglas-fir and lodgepole pine (10 and 14%, respectively), it was substantial in western larch and especially in western redcedar (22 and 42%, respectively). The most likely explanation for the latter high values is the presence of other conspecific orchards in the immediate vicinity of the two studied orchards, but reproductive phenology overlap between outside-orchard pollen sources' male flowering and within-orchard parents' female receptivity may also have played an important role.

Selfing, the highest possible level of inbreeding, known to negatively affect a population's fitness through the increase in frequency of rare deleterious alleles, was found to be relatively low in lodgepole pine (1.6%), moderate in western larch and western redcedar (8.3 and 7.3%, respectively), and high in Douglas-fir (15.2%). While selfing rates between 2 and 10% are common among conifer species, a value of 7.3% in western redcedar is surprisingly low, considering that values up to one order in magnitude higher have been reported in both artificial (El-Kassaby et al. 1994) and natural (O'Connell et al. 2008; O'Connell et al. 2001) populations. The observed low selfing can be attributed to high gene flow, as every pollen contamination is also an outcrossing event, and to top-grafting, which creates two different genotypes within a single tree and thus significantly reduces the chance for self-fertilization. The estimate of 15.2% selfing in Douglas-fir was unexpectedly high, especially considering that the seed orchard population had been treated with a fine water mist system to delay and synchronize parental reproductive phenology and that additional pollen from both within- and outside pollen donors (supplemental mass pollination; SMP) was applied. While synchronizing reproductive phenology is supposed to create an equal mating environment and ultimately promote panmixia within an orchard's population, SMP is expected to reduce self-fertilization by introducing additional pollen that can, depending on the number and timing of applications, avoid or outcompete self-pollen as well as pollen from background pollen sources. Top-grafting along with high gene flow is a plausible explanation for the overall high outcrossing rate of 92.7% estimated in the western redcedar seed orchard.

Male N_e was markedly reduced in all four seed orchards, reaching 45, 50, 58, and 62% of the census numbers for western larch, Douglas-fir, lodgepole pine, and western redcedar, respectively.

Partial pedigree reconstruction of maternal family array provides unbiased estimates of male-associated parameters such as pollen contamination, selfing, effectiveness of SMP, bloom delay, or top-grafting, effective number of fathers and male genetic worth. However, this method lacks the capability of quantifying the female component of reproductive success and in turn makes it impossible to infer any female-related parameters. On the other hand, the independent genetic analysis of diploid embryos and their corresponding maternal-haploid megagametophytes (Lai et al. 2010 and Chapter 4) and the combined utilization of nuclear and organelle (chloroplast) genetic markers (Chapter 5) on a sample of bulk seed enabled the simultaneous estimation of both male and female components of reproductive success because bulk seed sample represents the entire offspring population and thus reflects actual female gametic contributions. The bulk seed analyses also provided a more complex image of the populations' natural mating as compared with family array analyses (Chapters 2 and 3), as documented by a greater number of assembled full-sib families (Chapter 5). Results presented in Chapters 4 and 5 illustrate that genetic parameters of an offspring population (such as a seed orchard crop) can be determined with no prior information on the female parentage and that natural crosses among the populations members can be assembled *a posteriori*, forming a mating design matrix that would otherwise require extremely expensive as well as labor intensive controlled pollination. Posterior assemblage of crosses and subsequent determination of parental breeding values is an integral part of a novel concept called "breeding without breeding" (El-Kassaby and Lstibůrek 2009), an approach that significantly reduces breeding time and simplifies testing efforts, thus allowing for faster generational turnover.

Female reproductive success was enormously distorted in Douglas-fir, with the most successful single mother and the top 20% of mothers having produced over one third and 77% of the total seed crop, respectively. Although in lodgepole pine the distortion was less apparent (the top 20% of mothers only contributed 54% of the seed crop), this finding can be viewed as an experimental verification of the phenotype-derived approximation that 20% of mothers produce as much as 80% of seed in a crop (Anonymous 1976).

The joint male and female reproductive success was also markedly imbalanced, as the top 20% of parents contributed 60 and 43% of gametes in Douglas-fir and lodgepole pine seed crops, respectively.

Female N_e was even more reduced from the census number of parents than male N_e . While in lodgepole pine the relative N_e reached 50%, in Douglas-fir it was as little as 13%. This strikingly low value should draw orchard managers' attention, because had the N_e been calculated solely from parental representation, which is the common practice, the relative reduction would not be so substantial, leaving the estimate apparently satisfactorily high at 79% of the census number.

Combined male and female N_e was estimated to be 46.5 and 13.3 for lodgepole pine and Douglas-fir crops, respectively. The latter estimate is particularly noteworthy, as it exceeded the minimum N_e requirement mandated for planting on public land in British Columbia by only 33%.

Determining the actual parental reproductive success requires the availability of informative molecular genetic markers for a given species as well as statistical methods that can assign offspring to candidate parents (females or males or both) with an acceptable level of confidence. While this has become a relatively straightforward and reasonably inexpensive task recently, seed orchard managers had been looking for methods that could reliably predict the reproductive success without the need for any laboratory work.

I evaluated the reliability of several such methods by comparing parental representation and fecundity scores with actual estimates of the reproductive success obtained from microsatellite DNA-based pedigree reconstruction (Chapter 6). Parental representation was found to be the weakest predictor, explaining 19 and 20% of the variance in female and 42, 31, 27, and 8% of the variance in male reproductive success of Douglas-fir, lodgepole pine, western larch, and western redcedar, respectively. On the other hand, the inclusion of parental fecundity scores substantially improved the predictions' accuracy in all four species: Seed-cone volume explained 89 and 66% of the variance in female reproductive success of Douglas-fir and lodgepole pine, respectively, while full pollen survey explained 55, 58, 73, and 61% of the variance in male reproductive success in Douglas-fir, lodgepole pine, western larch, and western redcedar, respectively, demonstrating that fecundity variation is a very important factor to be considered when seed crop genetic parameters are estimated.

Furthermore, high correlation between fecundity data from full (all trees) and partial (50% of trees) pollen surveys in all four species indicated that scoring just a subset of ramets may suffice, as it does not compromise the estimates' accuracy despite reduced effort.

Seed-cone volume and partial pollen survey also provided sufficiently accurate proxies to estimates of both genetic gain and effective number of parents obtained from pedigree reconstruction. Effective number of parents was overestimated by most of the evaluated methods, however, gene flow from background pollen sources was not included in N_e calculations and thus the estimates shown in Table 7.1 may not necessarily reflect the actual reduction in genetic diversity. For instance, although male N_e was estimated to be just 62% of the census number in the western redcedar seed orchard, Table 6.1 clearly shows that genetic diversity actually improved, as the offspring population contained over six alleles per locus more than the parental population.

A multitude of methods have been developed and utilized by seed orchard managers to simultaneously balance crops' genetic gain and diversity both in newly established and existing orchards, such as linear (Lindgren and Matheson 1986) or optimum deployment of parents (Lindgren et al. 2009) or selective seed cone harvesting (Lindgren and El-Kassaby 1989). Following introduction of quadratic optimization, first used by animal breeders for balancing genetic gain and diversity in populations of related individuals (Meuwissen 1997; Villanueva and Woolliams 1997), I developed an optimization protocol that maximizes seed crops' genetic gain given a certain predefined value of N_e , while collectively considering parental female and male fecundities, co-ancestry among parents, inbreeding (Chapter 7), and variation in germination capacity (% of seed that germinate) (Chapter 8). This protocol is suitable for any generation's seed orchards and has proven to be effective in creating custom seedlots with multiple gain levels, being restricted only by predicted parental reproductive success and the distribution of breeding values. While this method cannot deal with the potential threat posed by inbreeding (neither selfed nor otherwise inbred seed can be filtered out from the computed optimum female gametic contributions and thence from the custom seedlots, as no information about their genetic status is available), it can be utilized for selecting "more valuable" portions of seed crops when seed supply exceeds demand, for mixing surpluses from multiple years, or if a given seedlot fails to meet minimum N_e requirements.

9.2 Research Limitations

The majority of work presented in this thesis is based on pedigree reconstruction of seed orchards' crops. Pedigree reconstruction is a complex process consisting of a number of steps, in each of which errors can be made that lead to reducing its accuracy.

- Mislabeling of ramets and the presence of alien genotypes represent the most common issues in seed orchards. While mislabeled ramets that perfectly match a parent's multilocus genotype are easily identified and back-assigned to the correct parent, alien genotypes – whether an overgrowing rootstock or mistakenly introduced parent – may negatively affect parentage analyses because a certain portion of pollen as well as seed may be a product of none of the sampled candidate parents. In family array analyses, an alien's offspring (i.e., seed contamination) is likely to be nearly always revealed through its multilocus genotype and can be subsequently excluded from any analyses. However, in bulk seed samples the female component of parentage could be confused with its male counterpart, which could artificially inflate pollen contamination. A solution can be provided by genotyping every single ramet of each parent, but due to their usually large number in seed orchards (~1000) this would require additional costs and effort. The number of mislabeled ramets detected in the studied seed orchards was generally low, as one, two, and three mislabeling incidents (0.7, 1.4, and 2.3%) and zero, three and three alien ramets (0.0, 2.1, and 2.3%) were detected in the seed orchards of Douglas-fir, lodgepole and western redcedar, respectively.
- Another factor that might have affected the accuracy of results presented in this thesis is the informativeness of the microsatellite markers used. Whereas in Douglas-fir, lodgepole pine, and western redcedar highly informative arrays of loci with low frequencies of null alleles were employed and thus both false assignment and false exclusion can be considered at minimum, in western larch several loci were used with relatively high null allele frequencies (for instance, 0.50 for locus UBCLXdi_16; see Table S1 in El-Kassaby et al. 2011 for more details), which might have artificially inflated the estimate of pollen contamination in the studied seed lot. This presumption can be well supported by the much lower estimate of pollen contamination obtained by El-Kassaby et al. (2011) (7.7 versus 22.0% reported in Chapter 2), who in their pedigree reconstruction analyses utilized only a subset of the best three out of nine loci, but

supplemented them with additional three loci developed by Khasa et al (2000) and four by Isoda and Watanabe (2006). In fact, the seedlot was originally genotyped for a total of 17 loci; however, the assignment rate gradually dropped as loci with null alleles were added to the best ten, indicating that the utilization of more loci does not necessarily improve the parentage analyses' accuracy unless the loci possess no or very low null allele frequencies; otherwise their effect may even be detrimental.

- Furthermore, although CERVUS (Kalinowski et al. 2007) is one of the most popular software programs for parentage analyses and I used it throughout this thesis, Koch et al. (2008) point out that unlike their newer algorithm, CERVUS produces a very high type I error rate (i.e., false assignment) when the size of the non-sampled population is assumed to be lower than it actually is. This can pose a serious problem to the correct assignment of offspring to candidate parents particularly in seed orchards, as the size of background pollen sources is mostly unknown and is usually guessed based on the best knowledge of the occurrence of conspecific stands in the orchard's vicinity.

9.3 Future Directions

- All of the results presented in this thesis are based on single-year seed crops only. Due to the commonly observed among-year variation in both male and female fecundities, parental reproductive success is also likely to vary correspondingly. It would therefore be worthwhile to analyze several more consecutive years' seed crops, at least for one of the four studied species, so that more general conclusions regarding temporal variation of crops' genetic quality could be drawn.
- The effectiveness of seed orchard designs could be evaluated by focusing on seed crops from every ramet of just one single parent separately, rather than analyzing mixed crops from all ramets by parents (i.e., maternal half-sib family array) as was the case of most studies presented in this thesis. The model parent should be selected with respect to the uniqueness of its genotypic profile, which should be as distinct from other parents (and potential background pollen sources) as possible, as well as with respect to the number of its replicates and distance between them so that the relative gametic contribution from surrounding ramets to their crops and thus within-orchard pollen flow could be determined. While this research has already been carried out in a number of seed

orchards, additional data would be helpful for a more general conclusion, as the previously reported results were not fully consistent (see Section 1.6.3 for review).

- As briefly mentioned in Section 1.6.4 on crown management in seed orchards, top-grafting, a method practiced with the aim to induce early flowering and reduce the time needed for one tree improvement cycle (Bramlett and Burris 1995; Goadling et al. 2000), has never been applied specifically for promoting outcrossing in seed orchards, although a positive effect of this technique can be anticipated. A simple experiment with a single factor (grafting) and two levels (regular grafting as a control and top-grafting) could be designed where seed from both types of grafting would be collected and analyzed separately and the effect of this management practice on outcrossing rate would be evaluated. The null hypothesis would be that the difference in outcrossing rates between the two types of ramets is not statistically significant and if the null hypothesis were rejected, then top-grafting could be applied to reduce selfing in species or seed orchards where selfing is (or is expected to be) a serious issue.
- Finally, all parentage analyses in this thesis have been conducted using microsatellite DNA markers. As new high-throughput technologies (e.g., Affymetrix, Illumina, or Perlegen) have been providing cost-effective ways to produce large amounts of SNP data for many species, including non-model species, it would be particularly attractive to compare microsatellite- and SNP-based results to determine whether and to what extent conclusions would differ based on these two approaches.

Bibliography

- Adams, G.W., and Kunze, H.A. 1996. Clonal variation in cone and seed production in black and white spruce seed orchards and management implications. *Forestry Chronicle* **72**(5): 475-480.
- Adams, R.I., Brown, K.M., and Hamilton, M.B. 2004. The impact of microsatellite electromorph size homoplasy on multilocus population structure estimates in a tropical tree (*Corythophora alta*) and an anadromous fish (*Morone saxatilis*). *Molecular Ecology* **13**(9): 2579-2588.
- Adams, W.T. 1982. Clonal variation in pollen-related characteristics of Douglas-fir. *Canadian Journal of Forest Research* **12**(2): 403-408.
- Adams, W.T. 1992. Gene dispersal within forest tree populations. *New Forests* **6**(1-4): 217-240.
- Adams, W.T., and Birkes, D.S. 1989. Mating patterns in seed orchards. *In* Proceedings of the 20th Southern Forest Tree Improvement Conference, Clemson University, Charleston, South Carolina. pp. 75-86.
- Allard, R.W. 1960. Principles of Plant Breeding. John Wiley and Sons, New York, NY, USA.
- Allen, G.S. 1961. Testing Douglas-fir seed for provenance. *In* Proc Int Seed Test Assoc 26. pp. 388-403.
- Almqvist, C., Bergsten, U., Bondesson, L., and Eriksson, U. 1998. Predicting germination capacity of *Pinus sylvestris* and *Picea abies* seeds using temperature data from weather stations. *Canadian Journal of Forest Research* **28**(10): 1530-1535.
- Almqvist, C., and Ekberg, I. 2001. Interstock and GA 4/7 effects on flowering after topgrafting in *Pinus sylvestris*. *Forest Genetics* **8**(4): 279-284.
- Amos, B., and Pemberton, J. 1992. DNA fingerprinting in non-human populations. *Current opinion in genetics & development* **2**(6): 857-860.

- Andersson, E., Jansson, R., and Lindgren, D. 1974. Some results from second generation crossings involving inbreeding in Norway spruce (*Picea abies*). *Silvae Genetica* **23**: 34–43.
- Anonymous. 1976. Twentieth annual report on cooperative tree improvement and hardwood research program, North Carolina State University, Raleigh, NC.
- Anonymous. 2002. The MOSEK.NET API manual. Version 4.0 (Revision 60). Mosek ApS, Denmark.
- Anonymous. 2009. Alberta Forest Genetic Resource Management and Conservation Standards: Second Revision of STIA. Alberta Sustainable Resource Development, Forestry Division. Edmonton, Alberta.
- Askew, G.R. 1986. Implications of non-synchronous flowering in clonal conifer seed orchards. *In* IUFRO Joint Meeting of Working Parties on Breeding Theory, Progeny Testing and Seed Orchards. *Edited by* A. V. Hatcher, and R. J. Weir. Tree Improvement Co-op., North Carolina State University, Raleigh, NC, Williamsburg, Virginia. pp. 182-191.
- Askew, G.R. 1988. Estimation of gamete pool compositions in clonal seed orchards. *Silvae Genetica* **37**(5-6): 227-232.
- Askew, G.R. 1992. Potential genetic improvement due to supplemental mass pollination management in conifer seed orchards. *Forest Ecology and Management* **47**(1-4): 135-147.
- Askew, G.R., and Blush, T.D. 1990. An index of phenological overlap in flowering for clonal conifer seed orchards. *Silvae Genetica* **39**(3-4): 168-171.
- Awise, J.C., and Hamrick, J.L. 1996. Conservation genetics: case histories from nature. Chapman and Hall, New York, N.Y.
- Baradat, P. 1987. A method for evaluating consanguinity among plants from 1st generation seed orchards. III. Application and optimization of an economic alternative to elite clonal seed orchards. *Silvae Genetica* **36**(3-4): 134-144.
- Beckmann, J.S., and Weber, J.L. 1992. Survey of human and rat microsatellites. *Genomics* **12**(4): 627-631.

- Begon, M., Mortimer, M., and Thompson, D.J. 1996. Population Ecology: A Unified Study of Animal and Plants. 3 ed. Oxford University Press, Cambridge, MA.
- Bell, G.D., and Fletcher, A.M. 1978. Computer organized orchard layouts (COOL) based on the permuted neighborhood design concept. *Silvae Genetica* **27**(6): 223-225.
- Bila, A.D., Lindgren, D., and Mullin, T.J. 1999. Fertility variation and its effect on diversity over generations in a teak plantation (*Tectona grandis* L.f.). *Silvae Genetica* **48**(3-4): 109-114.
- Bilir, N., Kang, K.S., and Ozturk, H. 2002. Fertility variation and gene diversity in clonal seed orchards of *Pinus brutia*, *Pinus nigra* and *Pinus sylvestris* in Turkey. *Silvae Genetica* **51**(2-3): 112-115.
- Bishir, J., and Namkoong, G. 1987. Unsound seeds in conifers - estimation of numbers of lethal alleles and of magnitudes of effects associated with the maternal parent. *Silvae Genetica* **36**(3-4): 180-184.
- Bishir, J., and Pepper, W.D. 1977. Estimation of number of embryonic lethal alleles in conifers. 1. self-pollinated seed. *Silvae Genetica* **26**(1): 50-54.
- Blouin, M.S. 2003. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends in Ecology and Evolution* **18**(10): 503-511.
- Blush, T.D., Bramlett, D.L., and El-Kassaby, Y.A. 1993. Reproductive phenology of conifer seed orchards. *In* Advances in Pollen Management. *Edited by* D. L. Bramlett, G. A. Askew, T. D. Blush, F. E. Bridgwater, and J. B. Jett. USDA Forest Service Agriculture Handbook 698. pp. 15-23.
- Bondesson, L., and Lindgren, D. 1993. Optimal utilization of clones and genetic thinning of seed orchards. *Silvae Genetica* **42**(4-5): 157-163.
- Botstein, D., White, R.L., Skolnick, M., and Davis, R.W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**(3): 314-331.
- Boyer, W.D. 1973. Air temperature, heat sums, and pollen shedding phenology of longleaf pine. *Ecology* **54**(2): 420-426.

- Bramlett, D.L. 1997. Genetic gain from mass controlled pollination and topworking. *Journal of Forestry* **95**(3): 15-19.
- Bramlett, D.L., and Burris, L.C. 1995. Topworking young scions into reproductively-mature loblolly pine. *In* Proceedings of the 23rd Southern Forest Tree Improvement Conference. *Edited by* J. R. Weir, and A. V. Hatcher, Asheville, NC. pp. 234-241.
- Bramlett, D.L., and Burris, L.C. 1998. Topworking genetic selections to reduce the breeding cycle in loblolly pine. *In* Proceedings of the 9th Biannual Southern Silviculture Research Conference. U.S.D.A. Forest Service, Southern Research Station, Asheville, North Carolina, Clemson, South Carolina. pp. 12-15.
- Bramlett, D.L., Dell, T.R., and Pepper, W.D. 1983. Genetic and maternal influences on virginia pine seed-germination. *Silvae Genetica* **32**(1-2): 1-4.
- Bramlett, D.L., and Popham, T.W. 1971. Model relating unsound seed and embryonic lethal alleles in self-pollinated pines. *Silvae Genetica* **20**(5-6): 192-193.
- Bridgwater, F.E., Blush, T.D., and Wheeler, N.C. 1993. Supplemental mass pollination. *In* Advances in pollen management. *Edited by* D. L., G. R. Askew, T. D. Blush, F. E. Bridgwater, and J. B. Jett. U.S.D.A. Handbook 698, U.S.D.A. Handbook 698. p. 69–77.
- Bridgwater, F.E., Bramlett, D.L., Byram, T.D., and Lowe, W.J. 1998. Controlled mass pollination in loblolly pine to increase genetic gains. *Forestry Chronicle* **74**(2): 185-189.
- Bridgwater, F.E., Bramlett, D.L., and Matthews, F.R. 1987. Supplemental mass pollination is feasible on an operational scale. *In* Proceedings from the 19th Southern Forest Tree Improvement Conference, College Sta, TX. pp. 216–222.
- Bridgwater, F.E., and Trew, I.F. 1981. Supplemental mass pollination. *In* Agriculture Handbook 587. *Edited by* E. C. Franklin. U.S.D.A. Forest Service, Washington D.C. pp. 52-56.
- Brisbane, J.R., and Gibson, J.P. 1995. Balancing selection response and rate of inbreeding by including genetic-relationships in selection decisions. *Theoretical and Applied Genetics* **91**(3): 421-431.

- Brondani, R.P.V., Brondani, C., Tarchini, R., and Grattapaglia, D. 1998. Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. *Theoretical and Applied Genetics* **97**(5-6): 816-827.
- Brown, A.H.D. 1989. Genetic characterization of plant mating system. *In* Plant Population Genetics, Breeding and Genetic Resources. *Edited by* A. H. D. Brown, M. T. Clegg, A. L. Kahler, and B. S. Weir. Sinauer, Sunderland, Massachusetts. pp. 145-162.
- Burczyk, J. 1996. Variance effective population size based on multilocus gamete frequencies in coniferous populations: An example of a Scots pine clonal seed orchard. *Heredity* **77**: 74-82.
- Burczyk, J., Adams, W.T., Moran, G.F., and Griffin, A.R. 2002. Complex patterns of mating revealed in a *Eucalyptus regnans* seed orchard using allozyme markers and the neighbourhood model. *Molecular Ecology* **11**(11): 2379-2391.
- Burczyk, J., Adams, W.T., and Shimizu, J.Y. 1996. Mating patterns and pollen dispersal in a natural knobcone pine (*Pinus attenuata* Lemmon) stand. *Heredity* **77**: 251-260.
- Burczyk, J., and Chalupka, W. 1997. Flowering and cone production variability and its effect on parental balance in a Scots pine clonal seed orchard. *Annales Des Sciences Forestieres* **54**(2): 129-144.
- Burczyk, J., Nikkanen, T., and Lewandowski, A. 1997. Evidence of an unbalanced mating pattern in a seed orchard composed of two larch species. *Silvae Genetica* **46**(2-3): 176-181.
- Burczyk, J., and Prat, D. 1997. Male reproductive success in *Pseudotsuga menziesii* (Mirb.) France: the effects of spatial structure and flowering characteristics. *Heredity* **79**: 638-647.
- Burdon, R.D. 2001. Genetic diversity and disease resistance: Some considerations for research, breeding, and deployment. *Canadian Journal of Forest Research* **31**(4): 596-606.
- Burdon, R.D., and Kumar, S. 2004. Forwards versus backwards selection: Trade-offs between expected genetic gain and risk avoidance. *New Zealand Journal of Forestry Science* **34**(1): 3-21.

- Burdon, R.D., and Shelbourne, C.J.A. 1971. Breeding populations for recurrent selection conflicts and possible solutions. *New Zealand Journal of Forestry Science* **1**(2): 174-193.
- Burrows, P.M. 1984. Inbreeding under selection from unrelated families. *Biometrics* **40**(2): 357-366.
- Butcher, P.A., Glaubitz, J.C., and Moran, G.F. 1999. Applications for microsatellite markers in the domestication and conservation of forest trees. *Forest Genetic Resources* **27**: 34-42.
- Butler, K., Field, C., Herbiner, C.M., and Smith, B.R. 2004. Accuracy, efficiency and robustness of four algorithms allowing full sibship reconstruction from DNA marker data. *Molecular Ecology* **13**(6): 1589-1600.
- Byram, T.D., Lowe, W.J., and McGriff, J.A. 1986. Clonal and annual variation in cone production in loblolly pine seed orchards. *Forest Science* **32**(4): 1067-1073.
- Caballero, A., Santiago, E., and Toro, M.A. 1996. Systems of mating to reduce inbreeding in selected populations. *Animal Science* **62**: 431-442.
- Callen, D.F., Thompson, A.D., Shen, Y., Phillips, H.A., Richards, R.I., Mulley, J.C., and Sutherland, G.R. 1993. Incidence and origin of null alleles in the (AC)_n microsatellite markers. *American Journal of Human Genetics* **52**(5): 922-927.
- Caron, G.E., and Powell, G.R. 1989. Patterns of seed-cone and pollen-cone production in young *Picea mariana* trees. *Canadian Journal of Forest Research* **19**(3): 359-364.
- Chaisurisri, K., Edwards, D.G.W., and El-Kassaby, Y.A. 1992. Genetic-control of seed size and germination in sitka spruce. *Silvae Genetica* **41**(6): 348-355.
- Chaisurisri, K., and El-Kassaby, Y.A. 1993. Estimation of clonal contribution to cone and seed crops in a Sitka spruce seed orchard. *Annales Des Sciences Forestieres* **50**(5): 461-467.
- Chaisurisri, K., and El-Kassaby, Y.A. 1994. Genetic diversity in a seed production population vs natural populations of Sitka spruce. *Biodiversity and Conservation* **3**(6): 512-523.

- Chaisurisri, K., Mitton, J.B., and El-Kassaby, Y.A. 1994. Variation in the mating system of sitka spruce (*Picea sitchensis*) - evidence for partial assortative mating. *American Journal of Botany* **81**(11): 1410-1415.
- Chaix, G., Gerber, S., Razafimaharo, V., Vigner, P., Verhaegen, D., and Hamon, S. 2003. Gene flow estimation with microsatellites in a Malagasy seed orchard of *Eucalyptus grandis*. *Theoretical and Applied Genetics* **107**(4): 705-712.
- Chaix, G., Vigner, P., Razafimaharo, V., and Hamon, S. 2007. Are phenological observations sufficient to estimate the quality of seed crops from a *Eucalyptus grandis* open-pollinated seed orchard? Consequences for seed collections. *New Forests* **33**(1): 41-52.
- Chakraborty, R., Deandrade, M., Daiger, S.P., and Budowle, B. 1992. Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. *Annals of Human Genetics* **56**: 45-57.
- Chakravarty, G.N., and Bagchi, S.K. 1993. A computer-program for permuted neighborhood seed orchard design. *Silvae Genetica* **42**(1): 1-5.
- Chakravarty, G.N., and Bagchi, S.K. 1994. Enhancement of the computer program of the permuted neighbourhood seed orchard design. *Silvae Genetica* **43**(2-3): 177-179.
- Charlesworth, D., and Charlesworth, B. 1981. Allocation of resources to male and female functions in hermaphrodites. *Biological Journal of the Linnean Society* **15**(1): 57-74.
- Charnov, E.L. 1982. *The theory of sex allocation*. Princeton University Press, Princeton, NJ.
- Chen, C.C., Liewlaksaneeyanawin, C., Funda, T., Kenawy, A.M.A., Newton, C.H., and El-Kassaby, Y.A. 2009. Development and characterization of microsatellite loci in western larch (*Larix occidentalis* Nutt.). *Molecular Ecology Resources* **9**(3): 843-845.
- Chen, Z., Kolb, T.E., Clancy, K.M., Hipkins, V.D., and DeWald, L.E. 2001. Allozyme variation in interior Douglas-fir: association with growth and resistance to western spruce budworm herbivory. *Canadian Journal of Forest Research* **31**(10): 1691-1700.
- Chistiakov, D.A., Hellemans, B., and Volckaert, F.A.M. 2006. Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. *Aquaculture* **255**(1-4): 1-29.

- Clutton-Brock, T.H. 1990. Reproductive Success: Studies of Individual Variation in Contrasting Breeding Systems. University Of Chicago Press.
- Cockerham, C. 1967. Group inbreeding and coancestry. *Genetics* **56**(1): 89-104.
- Colangeli, A.M., and Owens, J.N. 1990. The relationship between time of pollination, pollination efficiency, and cone size in western red cedar (*Thuja plicata*). *Canadian Journal of Botany* **68**(2): 439-443.
- Copes, D.L. 1974. Genetics of graft rejection in Douglas-fir. *Canadian Journal of Forest Research* **4**(2): 186-192.
- Copes, D.L. 1981. Isoenzyme uniformity in western red cedar seedlings from Oregon and Washington. *Canadian Journal of Forest Research* **11**(2): 451-453.
- Copes, D.L., and Snieszko, R.A. 1991. The influence of floral bud phenology on the potential mating system of a wind-pollinated Douglas-fir orchard. *Canadian Journal of Forest Research* **21**(6): 813-820.
- Critchfield, W.B. 1984. Impact of the pleistocene on the genetic structure of North American conifers. *In* Proceedings of the 8th North American Forest Biology Workshop. *Edited by* R. M. Lanner. Department of Forest Resources, Utah State University, Logan, Utah. pp. 70–117.
- Crow, J.F., and Denniston, C. 1988. Inbreeding and variance effective population numbers. *Evolution* **42**(3): 482-495.
- Crow, J.F., and Kimura, M. 1970. An Introduction to Population Genetics Theory. Harper and Row, New York.
- Crow, J.F., and Simmons, M.J. 1983. The mutation load in *Drosophila*. *In* The Genetics and Biology of *Drosophila*. Volume 3C. *Edited by* M. Ashburner, H. L. Carson, and J. N. Thompson. Academic Press, New York. pp. 1–35.
- Dakin, E.E., and Avise, J.C. 2004. Microsatellite null alleles in parentage analysis. *Heredity* **93**(5): 504-509.
- Daniels, J.D. 1978. Efficacy of supplemental mass-pollination in a Douglas-fir seed orchard. *Silvae Genetica* **27**(2): 52-58.

- David, A., Pike, C., and Stine, R. 2003. Comparison of selection methods for optimizing genetic gain and gene diversity in a red pine (*Pinus resinosa* Ait.) seedling seed orchard. *Theoretical and Applied Genetics* **107**(5): 843-849.
- Dawkins, R. 1982. *The Extended Phenotype*. Oxford University Press.
- de Nettancourt, D. 1997. Incompatibility in angiosperms. *Sexual Plant Reproduction* **10**(4): 185-199.
- Denti, D., and Schoen, D.J. 1988. Self-fertilization rates in white spruce - effect of pollen and seed production. *Journal of Heredity* **79**(4): 284-288.
- Devlin, B., and Ellstrand, N.C. 1990. The development and application of a refined method for estimating gene flow from angiosperm paternity analysis. *Evolution* **44**(2): 248-259.
- Devlin, B., Roeder, K., and Ellstrand, N.C. 1988. Fractional paternity assignment - theoretical development and comparison to other methods. *Theoretical and Applied Genetics* **76**(3): 369-380.
- Dow, B.D., and Ashley, M.V. 1998. High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. *Journal of Heredity* **89**(1): 62-70.
- Doyle, J.J., and Doyle, J.L. 1990. Isolation of plant DNA from fresh tissue. *Focus* **12**: 13-15.
- Dunlap, J.R., and Barnett, J.P. 1983. Influence of seed size on germination and early development of loblolly pine (*Pinus taeda* L.) germinants. *Canadian Journal of Forest Research* **13**(1): 40-44.
- Edwards, A.W.F. 1994. The fundamental theorem of natural-selection. *Biological Reviews of the Cambridge Philosophical Society* **69**(4): 443-474.
- Edwards, D.G.W. 1980. Maturity and quality of tree seeds - a state-of-the-art review. *Seed Science and Technology* **8**(4): 625-657.
- Edwards, D.G.W., and El-Kassaby, Y.A. 1988. Effect of flowering phenology, date of cone collection, cone-storage treatment and seed pretreatment on yield and germination of seeds from a Douglas-fir seed orchard. *Forest Ecology and Management* **25**(1): 17-29.

- Edwards, D.G.W., and El-Kassaby, Y.A. 1996. The biology and management of coniferous forest seeds: Genetic perspectives. *Forestry Chronicle* **72**(5): 481-484.
- El-Kassaby, Y., and Ritland, K. 1996. Impact of selection and breeding on the genetic diversity in Douglas-fir. *Biodiversity and Conservation* **5**(6): 795-813.
- El-Kassaby, Y.A. 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, Y.A. 1992. Domestication and genetic diversity - should we be concerned? *Forestry Chronicle* **68**: 687-700.
- El-Kassaby, Y.A. 1995. Evaluation of the tree-improvement delivery system - factors affecting genetic potential. *Tree Physiology* **15**(7-8): 545-550.
- El-Kassaby, Y.A. 1999. Phenotypic plasticity in western redcedar. *Forest genetics* **6**(4): 235-240.
- El-Kassaby, Y.A. 2000. Effect of forest tree domestication on gene pools. *In* Forest Conservation Genetics: Principles and Practice. *Edited by* A. Young, D. Boshier, and T. Boyle. Commonwealth Scientific and Industrial Research Organisation (CSIRO). CSIRO Publishing-CABI Publishing, Canberra, Australia. pp. 197-213.
- El-Kassaby, Y.A. 2003. Clonal-row vs. random seed orchard designs: Mating pattern and seed yield of western hemlock (*Tsuga heterophylla* (Raf.) Sarg.). *Forest Genetics* **10**: 121-127.
- El-Kassaby, Y.A., and Askew, G.R. 1991. The relation between reproductive phenology and reproductive output in determining the gametic pool profile in a Douglas-fir seed orchard. *Forest Science* **37**(3): 827-835.
- El-Kassaby, Y.A., and Askew, G.R. 1998. Seed orchards and their genetics. *In* Forest Genetics and Tree Breeding. *Edited by* A. K. Mandal, and G. L. Gibson. CBS Publishers and Distributors, Daryaganj, New Delhi.
- El-Kassaby, Y.A., Barnes, S., Cook, C., and Macleod, D.A. 1993a. Supplemental mass pollination success rate in a mature Douglas-fir seed orchard. *Canadian Journal of Forest Research* **23**(6): 1096-1099.

- El-Kassaby, Y.A., Cappa, E.P., Liewlaksaneeyanawin, C., Klápště, J., and Lstibůrek, M. 2011. Breeding without breeding: Is a complete pedigree necessary for efficient breeding? *In* PLoS ONE. p. e25737.
- El-Kassaby, Y.A., Chaisurisri, K., Edwards, D.G.W., and W., T.D. 1993b. Genetic control of germination parameters for Douglas-fir, Sitka spruce, western redcedar, and yellow-cedar and its impact on container nursery production. *In* Dormancy and barriers to germination. *Edited by* Edwards D. G. W., Forestry Canada, Pacific Forestry Centre, Victoria, BC. pp. 37-42.
- El-Kassaby, Y.A., and Cook, C. 1994. Female reproductive energy and reproductive success in a Douglas-fir seed orchard and its impact on genetic diversity. *Silvae Genetica* **43**(4): 243-246.
- El-Kassaby, Y.A., and Davidson, R. 1990. Impact of crop management-practices on the seed crop genetic quality in a Douglas-fir seed orchard. *Silvae Genetica* **39**(5-6): 230-237.
- El-Kassaby, Y.A., and 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, Y.A., and Edwards, D.G.W. 1998. Genetic control of germination and the effects of accelerated aging in mountain hemlock seeds and its relevance to gene conservation. *Forest Ecology and Management* **112**(3): 203-211.
- El-Kassaby, Y.A., Edwards, D.G.W., and Cook, C. 1990. Impact of crop management practices on seed yield in a Douglas-fir seed orchard. *Silvae Genetica* **39**(5-6): 226-230.
- El-Kassaby, Y.A., Edwards, D.G.W., and Taylor, D.W. 1992. Genetic control of germination parameters in Douglas-fir and its importance for domestication. *Silvae Genetica* **41**(1): 48-54.
- El-Kassaby, Y.A., Fashler, A.M.K., and Crown, M. 1989. Variation in fruitfulness in a Douglas-fir seed orchard and its effect on crop-management decisions. *Silvae Genetica* **38**(3-4): 113-121.

- El-Kassaby, Y.A., Fashler, A.M.K., and Sziklai, O. 1984. Reproductive phenology and its impact on genetically improved seed production in a Douglas-fir seed orchard. *Silvae Genetica* **33**(4-5): 120-125.
- El-Kassaby, Y.A., Funda, T., and Lai, B.S.K. 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**(2): 164-168.
- El-Kassaby, Y.A., and Jaquish, B. 1996. Population density and mating pattern in western larch. *Journal of Heredity* **87**(6): 438-443.
- El-Kassaby, Y.A., and Lstibůrek, M. 2009. Breeding without breeding. *Genetics Research* **91**(2): 111-120.
- El-Kassaby, Y.A., Moss, I., Kolotelo, D., and Stoehr, M. 2008. Seed germination: Mathematical representation and parameters extraction. *Forest Science* **54**(2): 220-227.
- El-Kassaby, Y.A., Parkinson, J., and Devitt, W.J.B. 1986. The effect of crown segment on the mating system in a Douglas-fir (*Pseudotsuga menziesii* (Mirb) Franco) seed orchard. *Silvae Genetica* **35**(4): 149-155.
- El-Kassaby, Y.A., and Reynolds, S. 1990. Reproductive phenology, parental balance, and supplemental mass pollination in a Sitka spruce seed-orchard. *Forest Ecology and Management* **31**(1-2): 45-54.
- El-Kassaby, Y.A., and Ritland, K. 1986a. Low levels of pollen contamination in a Douglas-fir seed orchard as detected by allozyme markers. *Silvae Genetica* **35**(5-6): 224-229.
- El-Kassaby, Y.A., and 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**(5-6): 240-244.
- El-Kassaby, Y.A., Ritland, K., Fashler, A.M.K., and Devitt, W.J.B. 1988. The role of reproductive phenology upon the mating system of a Douglas-fir seed orchard. *Silvae Genetica* **37**(2): 76-82.
- El-Kassaby, Y.A., Russell, J., and Ritland, K. 1994. Mixed mating in an experimental population of western red cedar, *Thuja-plicata*. *Journal of Heredity* **85**(3): 227-231.

- El-Kassaby, Y.A., Stoehr, M.U., Reid, D., Walsh, C.G., and Lee, T.E. 2007. Clonal-row versus random seed orchard designs: interior spruce mating system evaluation. *Canadian Journal of Forest Research* **37**(3): 690-696.
- El-Kassaby, Y.A., and Thomson, A.J. 1996. Parental rank changes associated with seed biology and nursery practices in Douglas-fir. *Forest Science* **42**(2): 228-235.
- Ellegren, H. 2000. Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics* **16**(12): 551-558.
- Emik, L.O., and Terrill, C.E. 1949. Systematic procedures for calculating inbreeding coefficients. *Journal of Heredity* **40**(2): 51-55.
- Epperson, B.K., and Allard, R.W. 1984. Allozyme analysis of the mating system in lodgepole pine populations. *Journal of Heredity* **75**: 212-214.
- Erickson, V.J., and Adams, W.T. 1989. Mating success in a coastal Douglas-fir seed orchard as affected by distance and floral phenology. *Canadian Journal of Forest Research* **19**(10): 1248-1255.
- Erickson, V.J., and Adams, W.T. 1990. Mating system variation among individual ramets in a Douglas-fir seed orchard. *Canadian Journal of Forest Research* **20**(10): 1672-1675.
- Eriksson, G., Ekberg, I., and Clapham, D. 2006. *An Introduction to Forest Genetics*. 2 ed. Swedish University of Agricultural Sciences, Uppsala, Sweden.
- Eriksson, G., Jonsson, A., and Lindgren, D. 1973. Flowering in a clone trial of *Picea abies*. *Studia Forestalia Suecica*(110): 1-45.
- Eriksson, U., Jansson, G., Yazdani, R., and Wilhelmsson, L. 1995. Effects of supplemental mass pollination (SMP) in a young and a mature seed orchard of *Pinus sylvestris*. *Tree Physiology* **15**(7-8): 519-526.
- Eriksson, U., Yazdani, R., Wilhelmsson, L., and Danell, O. 1994. Success rate of supplemental mass pollination in a seed orchard of *Pinus sylvestris* L. *Scandinavian Journal of Forest Research* **9**(1): 60-67.

- Estoup, A., Jarne, P., and Cornuet, J.M. 2002. Homoplasmy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* **11**(9): 1591-1604.
- Falconer, D.S., and Mackay, T.F.C. 1996. Introduction to quantitative genetics. 4 ed. Longmans Green, Harlow, Essex, UK.
- Fashler, A.M.K., and Devitt, W.J.B. 1980. A practical solution to Douglas-fir seed orchard pollen contamination. *Forestry Chronicle* **56**(5): 237-241.
- Fashler, A.M.K., and El-Kassaby, Y.A. 1987. The effect of water spray cooling treatment on reproductive phenology in a Douglas-fir seed orchard. *Silvae Genetica* **36**(5-6): 245-249.
- Fast, W., Dancik, B.P., and Bower, R.C. 1986. Mating system and pollen contamination in a Douglas-fir clone bank. *Canadian Journal of Forest Research* **16**(6): 1314-1319.
- Feilberg, L., and Sjøgaard, B. 1975. Historical Review of Seed Orchards. Forestry Committee Bulletin 54. Her Majesty's Stationary Office, London, UK.
- Fins, L., and Seeb, L.W. 1986. Genetic variation in allozymes of western larch. *Canadian Journal of Forest Research* **16**(5): 1013-1018.
- Fisher, P.J., Richardson, T.E., and Gardner, R.C. 1998. Characteristics of single- and multi-copy microsatellites from *Pinus radiata*. *Theoretical and Applied Genetics* **96**(6-7): 969-979.
- Fisher, R.A. 1999. The Genetical Theory of Natural Selection — A Complete Variorum Edition (edited by Henry Bennett). Oxford University Press, Oxford.
- Fowells, H.A. 1965. Silvics of Forest Trees of the United States. U.S.D.A. Agriculture Handbook No. 271.
- Franklin, E.C. 1974. Pollination in slash pine: first come, first served. *In* Seed Yield from Southern Pine Seed Orchards. Edited by J. Kraus. Georgia Forestry Center, Macon, GA. pp. 15–20.

- Funda, T., Chen, C.C., Liewlaksaneeyanawin, C., Kenawy, A.M.A., and El-Kassaby, Y.A. 2008. Pedigree and mating system analyses in a western larch (*Larix occidentalis* Nutt.) experimental population. *Annals of Forest Science* **65**(7).
- Funda, T., Liewlaksaneeyanawin, C., Fundova, I., Lai, B., Walsh, C., Van Niejenhuis, A., Cook, C., Graham, H., Woods, J., and El-Kassaby, Y. 2011. Congruence between parental reproductive investment and success determined by DNA-based pedigree reconstruction in conifer seed orchards. *Canadian Journal of Forest Research* **41**(2): 380-389.
- Funda, T., Liewlaksaneeyanawin, C., Fundova, I., Na Takuathung, C., and El-Kassaby, Y.A. Partial vs. full pedigree reconstruction in a lodgepole pine experimental population.
- Funda, T., Lstibůrek, M., Lachout, P., Klápště, J., and El-Kassaby, Y. 2009. Optimization of combined genetic gain and diversity for collection and deployment of seed orchard crops. *Tree Genetics & Genomes* **5**(4): 583-593.
- Geburek, T. 1986. Some results of inbreeding depression in Serbian spruce (*Picea omorika* (Panc) Purk). *Silvae Genetica* **35**(4): 169-172.
- Gietrych, M. 1975. Seed orchard designs. *In* Forestry Commission Bulletin 54. HMSO, London, UK. pp. 25–37.
- Glaubitz, J.C., El-Kassaby, Y.A., and Carlson, J.E. 2000. Nuclear restriction fragment length polymorphism analysis of genetic diversity in western redcedar. *Canadian Journal of Forest Research* **30**(3): 379-389.
- Goadin, G.D., Bridgwater, F.E., Bramlett, D.L., and Lowe, W.J. 2000. Top grafting loblolly pine in the Western Gulf region. *In* Proceedings of the 25th Biennial Southern Forest Tree Improvement Conference. Southern Forest Tree Improvement Committee, New Orleans, Louisiana, USA. pp. 60-66.
- Godt, M.J.W., Hamrick, J.L., Edwards-Burke, M.A., and Williams, J.H. 2001. Comparisons of genetic diversity in white spruce (*Picea glauca*) and jack pine (*Pinus banksiana*) seed orchards with natural populations. *Canadian Journal of Forest Research* **31**(6): 943-949.

- Goto, S., Miyahara, F., and Ide, Y. 2002. Monitoring male reproductive success in a Japanese black pine clonal seed orchard with RAPD markers. *Canadian Journal of Forest Research* **32**(6): 983-988.
- Goto, S., Watanabe, A., Miyahara, F., and Mori, Y. 2005. Reproductive success of pollen derived from selected and non-selected sources and its impact on the performance of crops in a nematode-resistant Japanese black pine seed orchard. *Silvae Genetica* **54**(2): 69-76.
- Greenwood, M.S. 1986. Gene exchange in loblolly pine - the relation between pollination mechanism, female receptivity and pollen availability. *American Journal of Botany* **73**(10): 1443-1451.
- Gregorius, H.R. 1989. Characterisation and Analysis of Mating System. Ekopan Verlag, Germany. p. 158.
- Griffin, A.R. 1982. Clonal variation in radiata pine seed orchards. 1. Some flowering, cone and seed production traits. *Australian Forest Research* **12**(4): 295-302.
- Griffin, A.R. 1984. Clonal variation in radiata pine seed orchards. 2. Flowering phenology. *Australian Forest Research* **14**(4): 271-281.
- Griffin, A.R., and Lindgren, D. 1985. Effect of inbreeding on production of filled seed in *Pinus radiata* - experimental results and a model of gene action. *Theoretical and Applied Genetics* **71**(2): 334-343.
- Grundy, B., Villanueva, B., and Woolliams, J.A. 1998. Dynamic selection procedures for constrained inbreeding and their consequences for pedigree development. *Genetical Research* **72**(2): 159-168.
- Gullberg, U., and Kang, H. 1985. Application of a model for tree breeding to conifers in southern Sweden. *Studia Forestalia Suecica*(170): 1-8.
- Gur-Arie, R., Cohen, C.J., Eitan, Y., Shelef, L., Hallerman, E.M., and Kashi, Y. 2000. Simple sequence repeats in *Escherichia coli*: Abundance, distribution, composition, and polymorphism. *Genome Research* **10**(1): 62-71.

- Gömöry, D., Bruchánik, R., and Longauer, R. 2003. Fertility variation and flowering asynchrony in *Pinus sylvestris*: consequences for the genetic structure of progeny in seed orchards. *Forest Ecology and Management* **174**(1-3): 117-126.
- Hahn, G.J., and Meeker, W.Q. 1991. *Statistical Intervals: A Guide for Practitioners*. John Wiley and Sons, Inc., Hoboken, NJ.
- Hamada, H., Petrino, M.G., and Kakunaga, T. 1982. A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **79**(21): 6465-6469.
- Hamrick, J.L., and Godt, M.J.W. 1989. Allozyme diversity in plant species. *In Plant Population Genetics, Breeding, and Genetic Resources. Edited by A. H. D. Brown, M. T. Clegg, A. L. Kahler, and B. S. Weir*. Sinauer Associates, Sunderland, Massachusetts. pp. 43-63.
- Handel, S.N. 1983. Pollination ecology, plant population structure, and gene flow. *In Pollination Biology. Edited by L. Real*. Academic Press, Inc., New York. pp. 163-211.
- Hansen, O.K., and Kjær, E.D. 2006. Paternity analysis with microsatellites in a Danish *Abies nordmanniana* clonal seed orchard reveals dysfunctions. *Canadian Journal of Forest Research* **36**(4): 1054-1058.
- Hansen, O.K., and Nielsen, U.B. 2010. Microsatellites used to establish full pedigree in a half-sib trial and correlation between number of male strobili and paternal success. *Annals of Forest Science* **67**(7).
- Hardy, G.H. 1908. Mendelian proportions in a mixed population. *Science* **28**: 49-50.
- Harju, A., and Muona, O. 1989. Background pollination in *Pinus sylvestris* seed orchards. *Scandinavian Journal of Forest Research* **4**(4): 513-520.
- Hartman, H.T., and Kestler, D.E. 1983. *Plant propagation: principles and practices*. Prentice Hall, Inc., Engelwood Cliffs, N.J.
- Hedrick, P.W. 2001. Conservation genetics: where are we now? *Trends in Ecology & Evolution* **16**: 629-636.

- Hedrick, P.W. 2005. Genetics of Populations. 3 ed. Jones and Bartlett, Boston, MA.
- Hirayama, D., Itoh, A., and Yamakura, T. 2004. Implications from seed traps for reproductive success, allocation and cost in a tall tree species *Lindera erythrocarpa*. *Plant Species Biology* **19**(3): 185-196.
- Hodge, G.R., and White, T.L. 1993. Advanced-generation wind-pollinated seed orchard design. *New Forests* **7**(3): 213-236.
- Hosius, B., Bergmann, F., Konnert, M., and Henkel, W. 2000. A concept for seed orchards based on isoenzyme gene markers. *Forest Ecology and Management* **131**(1-3): 143-152.
- Hunt, J., Bussiere, L.F., Jennions, M.D., and Brooks, R. 2004. What is genetic quality? *Trends in Ecology & Evolution* **19**(6): 329-333.
- Husband, B.C., and Schemske, D.W. 1996. Evolution of the magnitude and timing of inbreeding depression in plants. *Evolution* **50**(1): 54-70.
- Int Seed Test, A. 1985. International Rules for Seed Testing 1985. *Seed Science and Technology* **13**(2): 299-514.
- Isoda, K., and Watanabe, A. 2006. Isolation and characterization of microsatellite loci from *Larix kaempferi*. *Molecular Ecology Notes* **6**(3): 664-666.
- Jaquish, B., and El-Kassaby, Y.A. 1998. Genetic variation of western larch in British Columbia and its conservation. *Journal of Heredity* **89**(3): 248-253.
- Jaquish, B., Howe, G.T., Fins, L., and Rust, M. 1995. Western larch tree improvement programs in the Inland Empire and British Columbia. *In* Proceedings of an international symposium Ecology and Management of *Larix* forests: a look ahead. Whitefish, Montana, U.S.A. *Edited by* W. C. Schmidt, and K. J. McDonald. U.S.D.A. Forest Service, Intermountain Research Station, Ogden, UT. General Technical Report INT-319. pp. 452-460.
- Johnsen, K., Major, J.E., and Maier, C.A. 2003. Selfing results in inbreeding depression of growth but not of gas exchange of surviving adult black spruce trees. *Tree Physiology* **23**(14): 1005-1008.

- Johnson, R., and Lipow, S. 2002. Compatibility of breeding for increased wood production and longterm sustainability: the genetic variation of seed orchard seed and associated risks. *In* Wood compatibility initiative workshop 18. pp. 169-179.
- Jones, A.G. 2001. GERUD1.0: a computer program for the reconstruction of parental genotypes from progeny arrays using multilocus DNA data. *Molecular Ecology Notes* **1**(3): 215-218.
- Jones, A.G., and Ardren, W.R. 2003. Methods of parentage analysis in natural populations. *Molecular Ecology* **12**(10): 2511-2523.
- Jones, A.G., and Avise, J.C. 1997. Polygynandry in the dusky pipefish *Syngnathus floridae* revealed by microsatellite DNA markers. *Evolution* **51**(5): 1611-1622.
- Jones, O.R., and Wang, J. 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* **10**(3): 551-555.
- Jonsson, A., Ekberg, I., and Eriksson, G. 1976. Flowering in a seed orchard of *Pinus sylvestris*. *Studia Forestalia Suecica*(135): 1-38.
- Kalinowski, S.T., Taper, M.L., and Marshall, T.C. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* **16**(5): 1099-1106.
- Kang, H. 1982. Components of a tree breeding plan. *In* Proceedings of the IUFRO Joint Meeting of Working Parties on Genetics about Breeding Strategies including Multiclinal Varieties. Lower Saxony Forest Research Institute, Department of Forest Tree Breeding, Staufenberg-Escherode, Germany, Sensenstein, Germany. pp. 119-135.
- Kang, H., and Namkoong, G. 1988. Inbreeding effective population size under some artificial selection schemes. 1. Linear distribution of breeding values. *Theoretical and Applied Genetics* **75**(2): 333-339.
- Kang, H., and Nienstaedt, H. 1987. Managing long-term tree breeding stock. *Silvae Genetica* **36**(1): 30-39.
- Kang, K.-S. 2001. Genetic gain and gene diversity of seed orchard crops. PhD Thesis. *Acta Universitatis Agriculturae Sueciae. Silvestria* 187. p. 75.

- Kang, K.-S., and Lindgren, D. 1999. Fertility variation among clones of Korean pine (*Pinus koraiensis* S. et Z.) and its implications on seed orchard management. *Forest Genetics* **6**: 191-200.
- Kang, K.S. 2000. Clonal and annual variation of flower production and composition of gamete gene pool in a clonal seed orchard of *Pinus densiflora*. *Canadian Journal of Forest Research* **30**(8): 1275-1280.
- Kang, K.S., and El-Kassaby, Y.A. 2002. Considerations of correlated fertility between genders on genetic diversity: the *Pinus densiflora* seed orchard as a model. *Theoretical and Applied Genetics* **105**(8): 1183-1189.
- Kang, K.S., Harju, A.M., Lindgren, D., Nikkanen, T., Almqvist, C., and Suh, G.U. 2001a. Variation in effective number of clones in seed orchards. *New Forests* **21**(1): 17-33.
- Kang, K.S., Kjaer, E.D., and Lindgren, D. 2003a. Balancing gene diversity and nut production in *Corylus avellana* collections. *Scandinavian Journal of Forest Research* **18**(2): 118-126.
- Kang, K.S., Kjær, E.D., and Lindgren, D. 2003b. Balancing gene diversity and nut production in *Corylus avellana* collections. *Scandinavian Journal of Forest Research* **18**(2): 118-126.
- Kang, K.S., and Lindgren, D. 1998. Fertility variation and its effect on the relatedness of seeds in *Pinus densiflora*, *Pinus thunbergii* and *Pinus koraiensis* clonal seed orchards. *Silvae Genetica* **47**(4): 196-201.
- Kang, K.S., Lindgren, D., and Mullin, T.J. 2001b. Prediction of genetic gain and gene diversity in seed orchard crops under alternative management strategies. *Theoretical and Applied Genetics* **103**(6-7): 1099-1107.
- Kang, K.S., Lindgren, D., and Mullin, T.J. 2004. Fertility variation, genetic relatedness, and their impacts on gene diversity of seeds from a seed orchard of *Pinus thunbergii*. *Silvae Genetica* **53**(5-6): 202-206.
- Kang, K.S., Lindgren, D., Mullin, T.J., Choi, W.Y., and Han, S.U. 2005. Genetic gain and diversity of orchard crops under alternative management options in a clonal seed orchard of *Pinus thunbergii*. *Silvae Genetica* **54**(3): 93-96.

- Kawauchi, H., and Goto, S. 1999. Monitoring clonal management in nematode-resistant Japanese black pine seed orchard in Kagoshima Prefecture. *Journal of the Japanese Forestry Society* **81**: 338-340.
- Kaya, N., Isik, K., and Adams, W.T. 2006. Mating system and pollen contamination in a *Pinus brutia* seed orchard. *New Forests* **31**(3): 409-416.
- Keller, L.F., and Waller, D.M. 2002. Inbreeding effects in wild populations. *Trends in Ecology & Evolution* **17**(5): 230-241.
- Kent, J., and Richardson, T.E. 1997. Fluorescently labelled, multiplexed chloroplast microsatellites for high-throughput paternity analysis in *Pinus radiata*. *New Zealand Journal of Forestry Science* **27**(3): 305-312.
- Khasa, P.D., Newton, C.H., Rahman, M.H., Jaquish, B., and Dancik, B.P. 2000. Isolation, characterization, and inheritance of microsatellite loci in alpine larch and western larch. *Genome* **43**(3): 439-448.
- Kimura, M., and Crow, J.F. 1963. Measurement of effective population number. *Evolution* **17**(3): 279-&.
- Kjær, E.D. 1996. Estimation of effective population number in a *Picea abies* (Karst) seed orchard based on flower assessment. *Scandinavian Journal of Forest Research* **11**(2): 111-121.
- Kjær, E.D., and Wellendorf, H. 1998. Studies on the effect of unequal flowering on the effective population number in Danish seed orchard crops. *Forest Tree Improvement* **26**: 1-9.
- Koch, M., Hadfield, J.D., Sefc, K.M., and Sturmbauer, C. 2008. Pedigree reconstruction in wild cichlid fish populations. *Molecular Ecology* **17**(20): 4500-4511.
- Koelewijn, H.P., Koski, V., and Savolainen, O. 1999. Magnitude and timing of inbreeding depression in Scots pine (*Pinus sylvestris* L.). *Evolution* **53**(3): 758-768.
- Koller, D. 1962. Preconditioning of germination in lettuce at time of fruit ripening. *American Journal of Botany* **49**(8): 841-&.

- Koski, V. 1971. Embryonic lethals of *Picea abies* and *Pinus sylvestris*. Metsantutkimuslaitoksen Julkaisuja **75**(3): 9-30.
- Koski, V. 1973. On self pollination genetic load and subsequent inbreeding in some conifers. Metsantutkimuslaitoksen Julkaisuja **78**(10): 1973.
- Kozłowski, T.T., and Gentle, A.C. 1959. Influence of the seed coat on germination, water absorption, and oxygen uptake of Eastern white pine seed. *Forest Science* **5**: 389-395.
- Krakowski, J., and El-Kassaby, Y.A. 2005. Lodgepole pine and white spruce germination: Effects of stratification and simulated aging. *Silvae Genetica* **54**(3): 138-144.
- Krebs, C. 2002. *Ecology: The Experimental Analysis of Distribution and Abundance*. 5 ed. Prentice Hall, New York.
- Kroon, J., Wennström, U., Prescher, F., Lindgren, D., and Mullin, T.J. 2008. Estimation of clonal variation in seed cone production over time in a Scots pine (*Pinus sylvestris* L.) seed orchard. *Silvae Genetica* **58**(1-2): 53-62.
- Kärkkäinen, K., Koski, V., and Savolainen, O. 1996. Geographical variation in the inbreeding depression of Scots pine. *Evolution* **50**(1): 111-119.
- Lai, B.S., Funda, T., Liewlaksaneeyanawin, C., Klapste, J., Van Niejenhuis, A., Cook, C., Stoehr, M.U., Woods, J., and El-Kassaby, Y.A. 2010. Pollination dynamics in a Douglas-fir seed orchard as revealed by pedigree reconstruction. *Annals of Forest Science* **67**(8).
- Lamy, S., Bouchard, A., and Simon, J.P. 1999. Genetic structure, variability, and mating system in eastern white cedar (*Thuja occidentalis*) populations of recent origin in an agricultural landscape in southern Quebec. *Canadian Journal of Forest Research* **29**(9): 1383-1392.
- Ledig, F.T. 1998. Genetic variation in *Pinus*. In *Ecology and Biogeography of Pinus*. Edited by D. M. Richard. Cambridge University Press, Cambridge, United Kingdom. pp. 251-280.
- Levin, D.A., and Kerster, H.W. 1969. Dependence of bee-mediated pollen and gene dispersal upon plant density. *Evolution* **23**(4): 560-571.

- Levin, D.A., and Kerster, H.W. 1974. Gene flow in seed plants. Dobzhansky, T., Hecht, M. K., Steere, W. C. (Eds.). *Evolutionary Biology*, Vol. 7. X+314p. Illus. Plenum Press: New York, NY, USA; London, England.: 139-220.
- Li, W.-H. 2006. *Molecular Evolution*. Sinauer Associates Inc. p. 432.
- Libby, W.J., McCutchan, B.G., and Millar, C.I. 1981. Inbreeding depression in selfs of redwood. *Silvae Genetica* **30**(1): 15-25.
- Liewlaksaneeyanawin, C., Ritland, C.E., and El-Kassaby, Y.A. 2002. Inheritance of null alleles for microsatellites in the white pine weevil (*Pissodes strobi* Peck Coleoptera : Curculionidae). *Journal of Heredity* **93**(1): 67-70.
- Liewlaksaneeyanawin, C., Ritland, C.E., El-Kassaby, Y.A., and Ritland, K. 2004. Single-copy, species-transferable microsatellite markers developed from loblolly pine ESTs. *Theoretical and Applied Genetics* **109**(2): 361-369.
- Lindgren, D. 1974. Aspects on suitable number of clones in a seed orchard. *In* Proceedings from joint IUFRO meeting S2.04.01-3, Stockholm. pp. 293-306.
- Lindgren, D. 1989. Genetisk diversitet, klonantal och fröplantager. Sveriges Skogsvårdsförbunds tidskrift 4/87: 3-13.
- Lindgren, D., Cui, J.G., Son, S.G., and Sonesson, J. 2004. Balancing seed yield and breeding value in clonal seed orchards. *New Forests* **28**(1): 11-22.
- Lindgren, D., Danusevicius, D., and Rosvall, O. 2009. Unequal deployment of clones to seed orchards by considering genetic gain, relatedness and gene diversity. *Forestry* **82**(1): 17-28.
- Lindgren, D., and El-Kassaby, Y.A. 1989. Genetic consequences of combining selective cone harvesting and genetic thinning in clonal seed orchards. *Silvae Genetica* **38**(2): 65-70.
- Lindgren, D., Gea, L., and Jefferson, P. 1996. Loss of genetic diversity monitored by status number. *Silvae Genetica* **45**(1): 52-59.
- Lindgren, D., Gea, L.D., and Jefferson, P.A. 1997. Status number for measuring genetic diversity. *Forest Genetics* **4**(2): 69-76.

- Lindgren, D., and Gregorius, H.R. 1976. Inbreeding and coancestry. *In* Proceedings of the IUFRO Joint Meeting on Advanced Generation Breeding, Bordeaux, France. pp. 49-69.
- Lindgren, D., Jefferson, P.A., and Gea, L.D. 1995. Status number - a measure of genetic diversity. *In* Evolution of Breeding Strategies for Conifers from the Pacific North West. Joint Meeting of the IUFRO Working Parties S2.02.05, 06, 12 and 14. *Edited by* C.-J. Bastien, Limoges, France.
- Lindgren, D., and Kang, K.-S. 1997. Status number – a useful tool for tree breeding. Research Report of the Forest Genetic Research Institute of Korea 33.
- Lindgren, D., and Matheson, A.C. 1986. An algorithm for increasing the genetic quality of seed from seed orchards by using the better clones in higher proportions. *Silvae Genetica* **35**(5-6): 173-177.
- Lindgren, D., and Mullin, T.J. 1997. Balancing gain and relatedness in selection. *Silvae Genetica* **46**(2-3): 124-129.
- Lindgren, D., and Mullin, T.J. 1998. Relatedness and status number in seed orchard crops. *Canadian Journal of Forest Research* **28**(2): 276-283.
- Lindgren, D., and Prescher, F. 2005. Optimal clone number for seed orchards with tested clones. *Silvae Genetica* **54**(2): 80-92.
- Lindgren, D., Tellalov, Y., and Prescher, F. 2007. Seed set for Scots pine grafts is difficult to predict. *In* Proceedings of the IUFRO Division 2 Joint Conference: Low input breeding and conservation of forest genetic resources. *Edited by* F. Isik, Antalya, Turkey. pp. 139-141.
- Linhart, Y.B., Busby, W.H., Beach, J.H., and Feinsinger, P. 1987. Forager behavior, pollen dispersal, and inbreeding in 2 species of hummingbird-pollinated plants. *Evolution* **41**(3): 679-682.
- Litt, M., and Luty, J.A. 1989. A hypervariable microsatellite revealed by invitro amplification of a dinucleotide repeat within the cardiac-muscle actin gene. *American Journal of Human Genetics* **44**(3): 397-401.

- Lstibůrek, M., and El-Kassaby, Y.A. 2010. Minimum-inbreeding seed orchard design. *Forest Science* **56**(6): 603-608.
- Lynch, M., and Walsh, B. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer, Sunderland, MA.
- Magnussen, S., and Yeatman, C.W. 1990. Predictions of genetic gain from various selection methods in open pollinated *Pinus banksiana* progeny trials. *Silvae Genetica* **39**(3-4): 140-153.
- Malécot, G. 1948. *Les mathématiques de l'hérédité*. Masson, Paris.
- Mariette, S., Balsemin, E., Stoeckel, S., Tavaud, M., Le Boulter, H., Santi, F., and Verger, M. 2007. Parental participation in progeny and effective population sizes in experimental seed orchards of wild cherry *Prunus avium* L. (Batsch). *Annals of Forest Science* **64**(5): 533-539.
- Marshall, H.D., Newton, C., and Ritland, K. 2002. Chloroplast phylogeography and evolution of highly polymorphic microsatellites in lodgepole pine (*Pinus contorta*). *Theoretical and Applied Genetics* **104**(2-3): 367-378.
- Marshall, T.C., Slate, J., Kruuk, L.E.B., and Pemberton, J.M. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* **7**(5): 639-655.
- Massah, N., Wang, J., Russell, J.H., Van Niejenhuis, A., and El-Kassaby, Y.A. 2010. Genealogical relationship among members of selection and production populations of yellow cedar (*Callitropsis nootkatensis* D. Don Oerst.) in the absence of parental information. *Journal of Heredity* **101**(2): 154-163.
- Matziris, D.I. 1994. Genetic variation in the phenology of flowering in black pine. *Silvae Genetica* **43**(5-6): 321-328.
- McKeand, S., Mullin, T., Byram, T., and White, T. 2003. Deployment of genetically improved loblolly and slash pines in the south. *Journal of Forestry* **101**(3): 32-37.
- Meagher, T.R., and Thompson, E. 1986. The relationship between single parent and parent pair genetic likelihoods in genealogy reconstruction. *Theoretical Population Biology* **29**(1): 87-106.

- Metzgar, D., Bytof, J., and Wills, C. 2000. Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Research* **10**(1): 72-80.
- Meuwissen, T.H.E. 1997. Maximizing the response of selection with a predefined rate of inbreeding. *Journal of Animal Science* **75**(4): 934-940.
- Minore, D. 1990. *Thuja plicata* Donn ex D. Don. In *Silvics of North America*. Vol. 1. Conifers. Edited by R. M. Burns, and B. H. Honkala. U.S.D.A. Forest Service, Agriculture Handbook No. 654, Washington, DC. pp. 590-600.
- Mitton, J.B. 1995. Genetics and the physiological ecology of conifers. *Physiological Ecology; Ecophysiology of coniferous forests*: 1-36.
- Morgante, M., Vendramin, G.G., Rossi, P., and Olivieri, A.M. 1993. Selection against inbreds in early life-cycle phases in *Pinus leucodermis* Ant. *Heredity* **70**: 622-627.
- Moriguchi, Y., Iwata, H., Ujino-Ihara, T., Yoshimura, K., Taira, H., and Tsumura, Y. 2003. Development and characterization of microsatellite markers for *Cryptomeria japonica* D.Don. *Theoretical and Applied Genetics* **106**(4): 751-758.
- Moriguchi, Y., Taira, H., Tani, N., and Tsumura, Y. 2004. Variation of paternal contribution in a seed orchard of *Cryptomeria japonica* determined using microsatellite markers. *Canadian Journal of Forest Research* **34**(8): 1683-1690.
- Moriguchi, Y., Tani, N., Ito, S., Kanehira, F., Tanaka, K., Yomogida, H., Taira, H., and Tsumura, Y. 2005. Gene flow and mating system in five *Cryptomeria japonica* D. Don seed orchards as revealed by analysis of microsatellite markers. *Tree Genetics & Genomes* **1**(4): 174-183.
- Moriguchi, Y., Tsuchiya, S., Iwata, H., Ito, S., Tani, N., Taira, H., and Tsumura, Y. 2007. Factors influencing male reproductive success in a *Cryptomeria japonica* seed orchard revealed by microsatellite marker analysis. *Silvae Genetica* **56**(5): 207-214.
- Muona, O., and Harju, A. 1989. Effective population sizes, genetic variability, and mating system in natural stands and seed orchards of *Pinus sylvestris*. *Silvae Genetica* **38**(5-6): 221-228.
- Müller-Starck, G. 1977. Cross-fertilization in a conifer stand inferred from enzyme gene-markers in seeds. *Silvae Genetica* **26**(5-6): 223-226.

- Namkoong, G., and Bishir, J. 1987. The frequency of lethal alleles in forest tree populations. *Evolution* **41**(5): 1123-1127.
- Namkoong, G., Kang, H.C., and Brouard, J.S. 1988. Tree breeding: principles and strategies. Monographs on Theoretical and Applied Genetics 11. Springer-Verlag, New York.
- Nanson, A. 1972. The provenance seedling seed orchard. *Silvae Genetica* **21**(6): 243-248.
- Navascues, M., and Emerson, B.C. 2005. Chloroplast microsatellites: measures of genetic diversity and the effect of homoplasy. *Molecular Ecology* **14**(5): 1333-1341.
- Neale, D. 1984. Population genetic structure of the Douglas-fir shelterwood regeneration system in southwest Oregon. PhD thesis. Oregon State University, Corvallis, Oregon, USA.
- Neale, D.B., and Adams, W.T. 1985. The mating system in natural and shelterwood stands of Douglas-fir. *Theoretical and Applied Genetics* **71**(2): 201-207.
- Neale, D.B., and Sederoff, R.R. 1989. Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. *Theoretical and Applied Genetics* **77**(2): 212-216.
- Neale, D.B., Wheeler, N.C., and Allard, R.W. 1986. Paternal inheritance of chloroplast DNA in Douglas-fir. *Canadian Journal of Forest Research* **16**(5): 1152-1154.
- Neff, B.D., and Pitcher, T.E. 2005. Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. *Molecular Ecology* **14**(1): 19-38.
- Nei, M. 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America* **70**(12): 3321-3323.
- Nikkanen, T. 2001. Reproductive phenology in a Norway spruce seed orchard. *Silva Fennica* **35**(1): 39-53.
- Nikkanen, T., and Velling, P. 1987. Correlations between flowering and some vegetative characteristics of grafts of *Pinus sylvestris*. *Forest Ecology and Management* **19**(1-4): 35-40.
- O'Connell, L.M. 2003. The evolution of inbreeding in western redcedar (*Thuja plicata*: Cupressaceae). PhD Thesis. University of British Columbia.

- O'Connell, L.M., and Ritland, C.E. 2000. Characterization of microsatellite loci in western redcedar (*Thuja plicata*). *Molecular Ecology* **9**(11): 1920-1922.
- O'Connell, L.M., and Ritland, K. 2005. Post-pollination mechanisms promoting outcrossing in a self-fertile conifer, *Thuja plicata* (Cupressaceae). *Canadian Journal of Botany* **83**(3): 335-342.
- O'Connell, L.M., Ritland, K., and Thompson, S.L. 2008. Patterns of post-glacial colonization by western redcedar (*Thuja plicata*, Cupressaceae) as revealed by microsatellite markers. *Botany* **86**(2): 194-203.
- O'Connell, L.M., Russell, J., and Ritland, K. 2004. Fine-scale estimation of outcrossing in western redcedar with microsatellite assay of bulked DNA. *Heredity* **93**(5): 443-449.
- O'Connell, L.M., Viard, F., Russell, J., and Ritland, K. 2001. The mating system in natural populations of western redcedar (*Thuja plicata*). *Canadian Journal of Botany* **79**(6): 753-756.
- O'Reilly, C., Parker, W.H., and Barker, J.E. 1982. Effect of pollination period and strobili number on random mating in a clonal seed orchard of *Picea mariana*. *Silvae Genetica* **31**(2-3): 90-94.
- Oddou-Muratorio, S., Houot, M.L., Demesure-Musch, B., and Austerlitz, F. 2003. Pollen flow in the wildservice tree, *Sorbus torminalis* (L.) Crantz. I. Evaluating the paternity analysis procedure in continuous populations. *Molecular Ecology* **12**(12): 3427-3439.
- Olsson, T., Lindgren, D., and Li, B. 2001. Balancing genetic gain and relatedness in seed orchards. *Silvae Genetica* **50**(5-6): 222-227.
- Orr-Ewing, A.L. 1974. The incidence of dwarfing in inbred Douglas-fir. B.C. Forestry Service Research Note 64.
- Owens, J.N., Colangeli, A.M., and Morris, S.J. 1990. The effect of self-pollination, cross-pollination and no pollination on ovule, embryo, seed, and cone development in western red cedar (*Thuja plicata*). *Canadian Journal of Forest Research* **20**(1): 66-75.
- Owens, J.N., and Molder, M. 1980. Sexual reproduction in western red cedar (*Thuja plicata*). *Canadian Journal of Botany* **58**(12): 1376-1393.

- Owens, J.N., and Simpson, S.J. 1982. Further observations on the pollination mechanism and seed production of Douglas-fir. *Canadian Journal of Forest Research* **12**(2): 431-434.
- Owens, J.N., Simpson, S.J., and Caron, G.E. 1987. The pollination mechanism of Engelmann spruce (*Picea engelmannii*). *Canadian Journal of Botany* **65**(7): 1439-1450.
- Owens, J.N., Simpson, S.J., and Molder, M. 1981a. Sexual reproduction of *Pinus contorta*. 1. Pollen development, the pollination mechanism, and early ovule development. *Canadian Journal of Botany* **59**(10): 1828-1843.
- Owens, J.N., Simpson, S.J., and Molder, M. 1981b. The pollination mechanism and the optimal time of pollination in Douglas-fir (*Pseudotsuga menziesii*). *Canadian Journal of Forest Research* **11**(1): 36-50.
- Paetkau, D., and Strobeck, C. 1995. The molecular basis and evolutionary history of a microsatellite null allele in bears. *Molecular Ecology* **4**(4): 519-520.
- Pakkanen, A., Nikkanen, T., and Pulkkinen, P. 2000. Annual variation in pollen contamination and outcrossing in a *Picea abies* seed orchard. *Scandinavian Journal of Forest Research* **15**(4): 399-404.
- Park, Y.S., and Fowler, D.P. 1984. Inbreeding in black spruce (*Picea mariana* (Mill) BSP) - self-fertility, genetic load, and performance. *Canadian Journal of Forest Research* **14**(1): 17-21.
- Pemberton, J.M. 2008. Wild pedigrees: the way forward. *Proceedings of the Royal Society B-Biological Sciences* **275**(1635): 613-621.
- Pemberton, J.M., Slate, J., Bancroft, D.R., and Barrett, J.A. 1995. Nonamplifying alleles at microsatellite loci - a caution for parentage and population studies. *Molecular Ecology* **4**(2): 249-252.
- Perez, A.M.M., White, T.L., Huber, D.A., and Martin, T.A. 2007. Graft survival and promotion of female and male strobili by topgrafting in a third-cycle slash pine (*Pinus elliottii* var. *elliottii*) breeding program. *Canadian Journal of Forest Research* **37**(7): 1244-1252.
- Perry, D.J., and Dancik, B.P. 1986. Mating system dynamics of lodgepole pine in Alberta, Canada. *Silvae Genetica* **35**(5/6): 190-195.

- Perry, D.J., and Knowles, P. 1990. Evidence of high self-fertilization in natural populations of eastern white cedar (*Thuja occidentalis*). *Canadian Journal of Botany* **68**(3): 663-668.
- Philipson, J.J. 1997. Predicting cone crop potential in conifers by assessment of developing cone buds and cones. *Forestry* **70**(1): 87-96.
- Plomion, C., LeProvost, G., Pot, D., Vendramin, G., Gerber, S., Decroocq, S., Brach, J., Raffin, A., and Pastuszka, P. 2001. Pollen contamination in a maritime pine polycross seed orchard and certification of improved seeds using chloroplast microsatellites. *Canadian Journal of Forest Research* **31**(10): 1816-1825.
- Pompanon, F., Bonin, A., Bellemain, E., and Taberlet, P. 2005. Genotyping errors: Causes, consequences and solutions. *Nature Reviews Genetics* **6**(11): 847-859.
- Powell, W., Machray, G.C., and Provan, J. 1996. Polymorphism revealed by simple sequence repeats. *Trends in Plant Science* **1**(7): 215-222.
- Prescher, F. 2007. Seed orchards - genetic considerations on function, management and seed procurement. Doctoral Dissertation No. 2007:75, Swedish University of Agricultural Sciences, Umeå, Sweden.
- Prescher, F., Lindgren, D., and El-Kassaby, Y.A. 2006. Is linear deployment of clones optimal under different clonal outcrossing contributions in seed orchards? *Tree Genetics & Genomes* **2**(1): 25-29.
- Prescher, F., Lindgren, D., and Karlsson, B. 2008. Genetic thinning of clonal seed orchards using linear deployment may improve both gain and diversity. *Forest Ecology and Management* **254**(2): 188-192.
- Quinton, M., and Smith, C. 1995. Comparison of evaluation-selection systems for maximizing genetic response at the same level of inbreeding. *Journal of Animal Science* **73**: 2208-2212.
- Reed, D.H., and Frankham, R. 2003. Correlation between fitness and genetic diversity. *Conservation Biology* **17**(1): 230-237.
- Reynolds, S., and El-Kassaby, Y.A. 1990. Parental balance in Douglas-fir seed orchards - Cone crop vs. seed crop. *Silvae Genetica* **39**(1): 40-42.

- Ritland, K. 2002. Extensions of models for the estimation of mating systems using *n* independent loci. *Heredity* **88**: 221-228.
- Ritland, K., and El-Kassaby, Y.A. 1985. The nature of inbreeding in a seed orchard of Douglas-fir as shown by an efficient multilocus model. *Theoretical and Applied Genetics* **71**(3): 375-384.
- Roberds, J.H., Friedman, S.T., and El-Kassaby, Y.A. 1991. Effective number of pollen parents in clonal seed orchards. *Theoretical and Applied Genetics* **82**(3): 313-320.
- Robertson, A. 1961. Inbreeding in artificial selection programmes. *Genetical Research* **2**: 189-194.
- Robinson, L.W., and Wareing, P.F. 1969. Experiments on juvenile-adult phase change in some woody species. *New Phytologist* **68**(1): 67-78.
- Ross, S.D. 1983. Enhancement of shoot elongation in Douglas-fir by gibberellin A4/7 and its relation to the hormonal promotion of flowering. *Canadian Journal of Forest Research* **13**(5): 986-994.
- Ross, S.D., and Bower, R.C. 1989. Cost-effective promotion of flowering in a Douglas-fir seed orchard by girdling and pulsed stem injection of gibberellin A4/7. *Silvae Genetica* **38**(5-6): 189-195.
- Ross, S.D., and Bower, R.C. 1991. Promotion of seed production in Douglas-fir grafts by girdling plus gibberellin A4/7 stem injection and effect of retreatment. *New Forests* **5**(1): 23-34.
- Ross, S.D., Piesch, R.F., and Portlock, F.T. 1981. Promotion of cone and seed production in rooted ramets and seedlings of western hemlock by gibberellins and adjunct cultural treatments. *Canadian Journal of Forest Research* **11**(1): 90-98.
- Rosvall, O. 1999. Enhancing gain from long-term forest tree breeding while conserving genetic diversity. PhD thesis. *Acta Universitatis Agriculturae Sueciae. Silvestria* 109. p. 65.
- Ruotsalainen, S. 2002. Managing breeding stock in the initiation of a long-term tree breeding program. The Finnish Forest Research Institute. p. 95+61.

- Ruotsalainen, S., and Lindgren, D. 1998. Predicting genetic gain of backward and forward selection in forest tree breeding. *Silvae Genetica* **47**(1): 42-50.
- Ruotsalainen, S., and Lindgren, D. 2000. Stratified sublining: a new option for structuring breeding populations. *Canadian Journal of Forest Research* **30**(4): 596-604.
- Russell, J.H., Burdon, R.D., and Yanchuk, A.D. 2003. Inbreeding depression and variance structures for height and adaptation in self- and outcross *Thuja plicata* families in varying environments. *Forest Genetics* **10**: 171–184.
- Samuel, C.J.A., B., R.C., and Fletcher, A.M. 1972. A diallel cross in sitka spruce: assessment of first year characters in an early glasshouse test. *Theoretical and Applied Genetics* **42**: 53–61.
- Sancristobal, M., and Chevalet, C. 1997. Error tolerant parent identification from a finite set of individuals. *Genetical Research* **70**(1): 53-62.
- Savolainen, O., Karkkainen, K., and Kuittinen, H. 1992. Estimating numbers of embryonic lethals in conifers. *Heredity* **69**: 308-314.
- Savolainen, O., Kärkkäinen, K., Harju, A., Nikkanen, T., and Rusanen, M. 1993. Fertility variation in *Pinus sylvestris* - a test of sexual allocation theory. *American Journal of Botany* **80**(9): 1016-1020.
- Sawhney, R., and Naylor, J.M. 1979. Dormancy studies in seed of *Avena fatua*. 9. Demonstration of genetic variability affecting the response to temperature during seed development. *Canadian Journal of Botany* **57**(1): 59-63.
- Schmidt, W.C. 1995. Around the world with *Larix*: an introduction. *In* Proceedings of an international symposium Ecology and Management of *Larix* forests: a look ahead. Whitefish, Montana, U.S.A. *Edited by* W. C. Schmidt, and K. J. McDonald. U.S.D.A. Forest Service, Intermountain Research Station, Ogden, UT. General Technical Report INT-319. pp. 6-18.
- Schmidt, W.C., and McDonald, K.J. 1995. Proceedings of an international symposium Ecology and Management of *Larix* forests: a look ahead. Whitefish, Montana, U.S.A. U.S.D.A. Forest Service, Intermountain Research Station, Ogden, UT. General Technical Report INT-319.

- Schmidt, W.C., and Shearer, R.C. 1995. *Larix occidentalis*: a pioneer of the North American west. In Proceedings of an international symposium Ecology and Management of *Larix* forests: a look ahead. Whitefish, Montana, U.S.A. Edited by W. C. Schmidt, and K. J. McDonald. U.S.D.A. Forest Service, Intermountain Research Station, Ogden, UT. General Technical Report INT-319. pp. 33-37.
- Schmidt, W.C., Shearer, R.C., and Roe, A.L. 1976. Ecology and silviculture of western larch forests. U.S.D.A. Forest Service, Technical Bulletin No. 1520.
- Schmidtling, R.C. 1983. Genetic variation in fruitfulness in a loblolly pine (*Pinus taeda* L.) seed orchard. *Silvae Genetica* **32**(3-4): 76-80.
- Schmidtling, R.C., and Hipkins, V. 1998. Genetic diversity in longleaf pine (*Pinus palustris*): influence of historical and prehistorical events. *Canadian Journal of Forest Research* **28**(8): 1135-1145.
- Schoen, D.J., Denti, D., and Stewart, S.C. 1986. Strobilus production in a clonal white spruce seed orchard - evidence for unbalanced mating. *Silvae Genetica* **35**(5-6): 201-205.
- Schoen, D.J., and Stewart, S.C. 1986. Variation in male reproductive investment and male reproductive success in white spruce. *Evolution* **40**(6): 1109-1120.
- Schoen, D.J., and Stewart, S.C. 1987. Variation in male fertilities and pairwise mating probabilities in *Picea glauca*. *Genetics* **116**(1): 141-152.
- Schreiner, E.J. 1950. Genetics in relation to forestry. *Journal of Forestry* **48**: 33-38.
- Schreiner, E.J. 1961. Clonal or seedling seed orchards? In 9th Northeastern Forest Tree Improvement Conference, Syracuse, New York, USA. pp. 53-57.
- Selkoe, K.A., and Toonen, R.J. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**(5): 615-629.
- Semerikov, V.L., and Lascoux, M. 1999. Genetic relationship among Eurasian and American *Larix* species based on allozymes. *Heredity* **83**: 62-70.
- Semerikov, V.L., and Matveev, A.V. 1995. Investigation of genetic variation of isozyme loci in Siberian larch *Larix sibirica* Ldb. *Genetika* **31**(8): 1107-1113.

- Semerikov, V.L., Semerikov, L.F., and Lascoux, M. 1999. Intra- and interspecific allozyme variability in Eurasian *Larix* Mill. species. *Heredity* **82**: 193-204.
- Shaw, D.V., and Allard, R.W. 1982. Estimation of outcrossing rates in Douglas-fir using isoenzyme markers. *Theoretical and Applied Genetics* **62**(2): 113-120.
- Shea, K.L. 1987. Effects of population structure and cone production on outcrossing rates in Engelmann spruce and subalpine fir. *Evolution* **41**(1): 124-136.
- Shoulder, E. 1961. Effect of seed size on germination, growth, and survival of slash pine. *Journal of Forestry* **59**: 363-365.
- Silen, R.R. 1970. Further trials of cold water spray to reduce seed orchard contamination. Industrial Forestry Association 13, Olympia, Washington.
- Silen, R.R., and Keane, G. 1969. Cooling a Douglas-fir seed orchard to avoid pollen contamination. U.S.D.A. Forest Service Research Note PNW-101.
- Slavov, G.T., Howe, G.T., and Adams, W.T. 2005a. Pollen contamination and mating patterns in a Douglas-fir seed orchard as measured by simple sequence repeat markers. *Canadian Journal of Forest Research* **35**(7): 1592-1603.
- Slavov, G.T., Howe, G.T., Gyaourova, A.V., Birkes, D.S., and Adams, W.T. 2005b. Estimating pollen flow using SSR markers and paternity exclusion: accounting for mistyping. *Molecular Ecology* **14**(10): 3109-3121.
- Slavov, G.T., Howe, G.T., Yakovlev, I., Edwards, K.J., Krutovskii, K.V., Tuskan, G.A., Carlson, J.E., Strauss, S.H., and Adams, W.T. 2004. Highly variable SSR markers in Douglas-fir: Mendelian inheritance and map locations. *Theoretical and Applied Genetics* **108**(5): 873-880.
- Smith, D.N., and Devey, M.E. 1994. Occurrence and inheritance of microsatellites in *Pinus radiata*. *Genome* **37**(6): 977-983.
- Smith, J.S.C., Chin, E.C.L., Shu, H., Smith, O.S., Wall, S.J., Senior, M.L., Mitchell, S.E., Kresovich, S., and Ziegler, J. 1997. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparisons with data from RFLPS and pedigree. *Theoretical and Applied Genetics* **95**(1-2): 163-173.

- Smith, R.F., and Adam, C.I.G. 2003. A manual for managing conifer seed orchards in eastern Canada. Forest Genetics Ontario.
- Snyder, E.B. 1972. Five-year performance of self-pollinated slash pines. *Forest Science* **18**: 246–257.
- Sokal, R.R., and Rohlf, F.J. 1981. Biometry. 2nd ed. W.H. Freeman and Co., New York.
- Son, S.-G., Kang, K.-S., Lindgren, D., and Hyun, J.-O. 2003. Qualification for the value of seed orchard considering breeding value and seed productivity. *Journal of Korean Forestry Association* **91**: 601–608.
- Soranzo, N., Provan, J., and Powell, W. 1999. An example of microsatellite length variation in the mitochondrial genome of conifers. *Genome* **42**(1): 158-161.
- Sorensen, F. 1969. Embryonic genetic load in coastal Douglas-fir, *Pseudotsuga menziesii* var. *menziesii*. *American Naturalist* **103**(932): 389-398.
- Sorensen, F.C. 1982. The roles of polyembryony and embryo viability in the genetic system of conifers. *Evolution* **36**(4): 725-733.
- Sorensen, F.C. 1999. Relationship between self-fertility, allocation of growth, and inbreeding depression in three coniferous species. *Evolution* **53**(2): 417-425.
- Sorensen, F.C., and Cress, D.W. 1994. Effects of sib mating on cone and seed traits in coastal Douglas-fir. *Silvae Genetica* **43**(5-6): 338-345.
- Sorensen, F.C., Franklin, J.F., and Woollard, R. 1976. Self-pollination effects on seed and seedling traits in noble fir. *Forest Science* **22**(2): 155-159.
- Squillace, A.E., and Kraus, J.F. 1963. The degree of natural selfing in slash pine as estimated from albino frequencies. *Silvae Genetica* **12**: 46-50.
- St. Clair, J.B. 1993. Evaluating realized genetic gains from tree improvement. *In* Proceedings from the IUFRO S4.01 Conference on Modelling Stand Response to Silvicultural Practices. Edited by H. E. Burkhart, T. G. Gregoire, and J. L. Smith. Publication FWS-1-93 of the Virginia Polytechnic Institute and State University. Blacksburg, VA. pp. 145-157.

- Stoehr, M., Mehl, H., Nicholson, G., Pieper, G., and Newton, C. 2006. Evaluating supplemental mass pollination efficacy in a lodgepole pine orchard in British Columbia using chloroplast DNA markers. *New Forests* **31**(1): 83-90.
- Stoehr, M., Webber, J., and Woods, J. 2004. Protocol for rating seed orchard seedlots in British Columbia: quantifying genetic gain and diversity. *Forestry* **77**(4): 297-303.
- Stoehr, M.U. 2000. Seed production of western larch in seed-tree systems in the southern interior of British Columbia. *Forest Ecology and Management* **130**(1-3): 7-15.
- Stoehr, M.U., and El-Kassaby, Y.A. 1997. Levels of genetic diversity at different stages of the domestication cycle of interior spruce in British Columbia. *Theoretical and Applied Genetics* **94**(1): 83-90.
- Stoehr, M.U., and Newton, C.H. 2002. Evaluation of mating dynamics in a lodgepole pine seed orchard using chloroplast DNA markers. *Canadian Journal of Forest Research* **32**(3): 469-476.
- Stoehr, M.U., Orvar, B.L., Vo, T.M., Gawley, J.R., Webber, J.E., and Newton, C.H. 1998. Application of a chloroplast DNA marker in seed orchard management evaluations of Douglas-fir. *Canadian Journal of Forest Research* **28**(2): 187-195.
- Stonecypher, R.W. 1969. Recurrent selection in forest tree breeding. *Proceedings of Southern Conference on Forest Tree Improvement*. Houston, Texas. pp. 7-16.
- Szmidt, A.E., Alden, T., and Hallgren, J.E. 1987. Paternal inheritance of chloroplast DNA in *Larix*. *Plant Molecular Biology* **9**(1): 59-64.
- Szmidt, A.E., El-Kassaby, Y.A., Sigurgeirsson, A., Alden, T., Lindgren, D., and Hallgren, J.E. 1988. Classifying seedlots of *Picea sitchensis* and *Picea glauca* in zones of introgression using restriction analysis of chloroplast DNA. *Theoretical and Applied Genetics* **76**(6): 841-845.
- Tani, N., Takahashi, T., Ujino-Ihara, T., Iwata, H., Yishimura, K., and Tsumura, Y. 2004. Development and characteristics of microsatellite markers for sugi (*Cryptomeria japonica* D. Don) derived from microsatellite-enriched libraries. *Annals of Forest Science* **61**(6): 569-575.

- Tautz, D. 1989. Hypervariability of simple sequence as a general source for polymorphic DNA markers. *Nucleic Acids Research* **17**: 6463-6471.
- Toda, R. 1964. A brief review and conclusion of the discussion on seed orchards. *Silvae Genetica* **13**: 1-4.
- Toro, M., and Pérez-Enciso, M. 1990. Optimization of selection response under restricted inbreeding. *Genetics Selection Evolution* **22**(1): 93-107.
- Toth, G., Gaspari, Z., and Jurka, J. 2000. Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Research* **10**(7): 967-981.
- van Buijtenen, J.P. 1984. Strategies for maintaining the genetic base. How will advanced generation tests differ from first? Charleston, USDA Forest Service, Washington, D.C., USA. pp. 566-572.
- Vandeputte, M., Mauger, S., and Dupont-Nivet, M. 2006. An evaluation of allowing for mismatches as a way to manage genotyping errors in parentage assignment by exclusion. *Molecular Ecology Notes* **6**(1): 265-267.
- Varnell, R.S., Squillace, A.E., and Bengtson, W. 1967. Variation and heritability of fruitfulness in slash pine. *Silvae Genetica* **16**: 125-128.
- Varshney, R.K., Graner, A., and Sorrells, M.E. 2005. Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology* **23**(1): 48-55.
- Varshney, R.K., Thiel, T., Stein, N., Langridge, P., and Graner, A. 2002. In silico analysis on frequency and distribution of microsatellites in ESTs of some cereal species. *Cellular & Molecular Biology Letters* **7**(2A): 537-546.
- Vendramin, G.G., and Hansen, O.K. 2005. Molecular markers for characterizing diversity in forest trees. *In Conservation and Management of Forest Genetic Resources in Europe. Edited by T. Geburek, and J. Turok. Arbora Publishers, Zvolen, Slovakia. pp. 337–368.*
- Vendramin, G.G., Lelli, L., Rossi, P., and Morgante, M. 1996. A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae. *Molecular Ecology* **5**(4): 595-598.

- Vendramin, G.G., and Ziegenhagen, B. 1997. Characterisation and inheritance of polymorphic plastid microsatellites in *Abies*. *Genome* **40**(6): 857-864.
- Vignal, A., Milan, D., SanCristobal, M., and Eggen, A. 2002. A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution* **34**(3): 275-305.
- Villanueva, B., Avenano, S., and Woolliams, J.A. 2006. Prediction of genetic gain from quadratic optimisation with constrained rates of inbreeding. *Genetics Selection Evolution* **38**(2): 127-146.
- Villanueva, B., and Woolliams, J.A. 1997. Optimization of breeding programmes under index selection and constrained inbreeding. *Genetical Research* **69**(2): 145-158.
- Wagner, D.B., Furnier, G.R., Saghaiaroof, M.A., Williams, S.M., Dancik, B.P., and Allard, R.W. 1987. Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. *Proceedings of the National Academy of Sciences of the United States of America* **84**(7): 2097-2100.
- Wakeley, P.C., Wells, O.O., and Campbell, T.E. 1966. Mass production of shortleaf \times 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. pp. 78–79.
- Wallace, B. 1981. *Basic Population Genetics*. Columbia University Press.
- Wang, T., Aitken, S.N., Woods, J.H., Polsson, K., and Magnussen, S. 2004. Effects of inbreeding on coastal Douglas fir growth and yield in operational plantations: a model-based approach. *Theoretical and Applied Genetics* **108**(6): 1162-1171.
- Wang, T.L., Hagqvist, R., and Tigerstedt, P.M.A. 1999. Inbreeding depression in three generations of selfed families of silver birch (*Betula pendula*). *Canadian Journal of Forest Research* **29**(6): 662-668.
- Wang, Z., Weber, J.L., Zhong, G., and Tanksley, S.D. 1994. Survey of plant short tandem DNA repeats. *Theoretical and Applied Genetics* **88**(1): 1-6.

- Webber, J.E. 2000. Western hemlock: a manual for tree improvement seed production. B.C. Ministry of Forests, Research Branch Work Paper 44.
- Webber, J.E., and Yeh, F.C.H. 1987. Test of the first-on, first-in pollination hypothesis in coastal Douglas-fir. *Canadian Journal of Forest Research* **17**(1): 63-68.
- Weber, J.L., and May, P.E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* **44**(3): 388-396.
- Weinberg, W. 1908. Über den Nachweis der Vererbung beim Menschen. *Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg* **64**: 368–382.
- Weir, B.S. 1990. *Genetic Data Analysis*. Sinauer Associates, Inc., Sunderland, Massachusetts.
- Wheeler, N., and 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, Virginia. pp. 160-171.
- Wheeler, N.C., Adams, W.T., and Hamrick, J.L. 1993. Pollen distribution in wind-pollinated seed orchards. *In Advances in Pollen Management. Edited by D. L. Bramlett, G. R. Askew, T. D. Blush, Bridgwater F. E., and J. B. Jett*. U.S.D.A. Forest Service Handbook 698, Washington, DC. pp. 25-31.
- Wheeler, N.C., and Jech, K.S. 1992. The use of electrophoretic markers in seed orchard research. *New Forests* **6**(1-4): 311-328.
- White, T.L. 1987. A conceptual framework for tree improvement programs. *New Forests* **1**(4): 325-342.
- Williams, C.G., Barnes, R.D., and Nyoka, I. 1999. Embryonic genetic load for a neotropical conifer, *Pinus patula* Schiede et Deppe. *Journal of Heredity* **90**(3): 394-398.
- Williams, C.G., Hamrick, J.L., and Lewis, P.O. 1995. Multiple-population versus hierarchical conifer breeding programs - a comparison of genetic diversity levels. *Theoretical and Applied Genetics* **90**(3-4): 584-594.

- Williams, C.G., and Savolainen, O. 1996. Inbreeding depression in conifers: Implications for breeding strategy. *Forest Science* **42**(1): 102-117.
- Wittwer, R.F., Tauer, C.G., Huebschmann, M.M., and Huang, Y. 1997. Estimating seed quantity and quality in shortleaf pine (*Pinus echinata* Mill) cones from natural stands. *New Forests* **14**(1): 45-53.
- Wodehouse, R.P. 1935. Pollen grains. McGraw-Hill, New York.
- Woods, J.H. 2005. Methods for estimating gamete contributions to orchard seed crops and vegetative lots in British Columbia. B.C. Ministry of Forests and Range, Research Branch, Victoria, B.C. Tech. Rep. 025. p. 17.
- Woods, J.H. 2008. Business Plan. Forest Genetics Council of B.C.
- Woods, J.H., and Heaman, J.C. 1989. Effect of different inbreeding levels on filled seed production in Douglas-fir. *Canadian Journal of Forest Research* **19**(1): 54-59.
- Woods, J.H., Stoeck, M.U., and Webber, J.E. 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**(5-6): 203-209.
- Wray, N.R., and Goddard, M.E. 1994. Increasing long-term response to selection. *Genetics Selection Evolution* **26**(5): 431-451.
- Wright, S. 1931. Evolution in Mendelian populations. *Genetics* **16**(2): 0097-0159.
- Wright, S. 1969. Evolution and the genetics of populations. Vol. 2. The theory of gene frequencies. University of Chicago Press, Chicago.
- Wu, H.X., Matheson, A.C., and Spencer, D. 1998. Inbreeding in *Pinus radiata*. I. The effect of inbreeding on growth, survival and variance. *Theoretical and Applied Genetics* **97**(8): 1256-1268.
- Xiang, B., Li, B.L., and McKeand, S. 2003. Genetic gain and selection efficiency of loblolly pine in three geographic regions. *Forest Science* **49**(2): 196-208.

- Xie, C.Y., Dancik, B.P., and Yeh, F.C. 1991. The mating system in natural populations of *Thuja orientalis*. Canadian Journal of Forest Research **21**(3): 333-339.
- Xie, C.Y., and Knowles, P. 1992. Male fertility variation in an open-pollinated plantation of Norway spruce (*Picea abies*). Canadian Journal of Forest Research **22**(10): 1463-1468.
- Xie, C.Y., Woods, J., and Stoehr, M. 1994. Effects of seed orchard inputs on estimating effective population size of seedlots - a computer simulation. Silvae Genetica **43**(2-3): 145-154.
- Yanchuk, A.D. 2001. A quantitative framework for breeding and conservation of forest tree genetic resources in British Columbia. Canadian Journal of Forest Research **31**(4): 566-576.
- Yazdani, R., Hadders, G., and Szmidt, A.E. 1986. Supplemental mass pollination in a seed orchard of *Pinus sylvestris* L. investigated by isozyme analyses. Scandinavian Journal of Forest Research **1**(3): 309-316.
- Yazdani, R., Lindgren, D., and Stewart, S. 1989. Gene dispersion within a population of *Pinus sylvestris*. Scandinavian Journal of Forest Research **4**(3): 295-306.
- Yazdani, R., Scotti, I., Jansson, G., Plomion, C., and Mathur, G. 2003. Inheritance and diversity of simple sequence repeat (SSR) microsatellite markers in various families of *Picea abies*. Hereditas **138**(3): 219-227.
- Yeh, F.C. 1988. Isozyme variation of *Thuja plicata* (Cupressaceae) in British Columbia. Biochemical Systematics and Ecology **16**(4): 373-377.
- Ying, C.C., Murphy, J.C., and Andersen, S. 1985. Cone production and seed yield of lodgepole pine grafts. Forestry Chronicle **61**(3): 223-228.
- Zobel, B.J., Barber, J., Brown, C.L., and Perry, T.O. 1958. Seed orchards – their concept and management. Journal of Forestry **56**: 815-825.
- Zobel, B.J., and Talbert, J.T. 1984. Applied forest tree improvement. John Wiley & Sons, New York.

Appendix A Characterization of microsatellite loci for western larch (*Larix occidentalis* Nutt.) (after Chen et al. 2009).

Locus	Primer sequences (5' to 3')	t_a (°C)	Repeat motif	Size (bp)	A	H_o	H_e^*	G.A.N.
UBCLXtet_1-22	F: TTCCAAATTCCAGACTTCTGTTCA R: ACAATCTATGCCTGTGAACCCA	58	(TATC) ₉ (TA) ₁₂	175-250	8	0.19	0.57*	EU306567
UBCLXtet_21	F: TCCATGCATGAGCTAGCAATC R: CTAGATTGATTTTATGGCCATC	58	(TA) ₄ (GATA) ₁₄	100-200	7	0.58	0.80*	EU306568
UBCLXtet_32	F: CAAATAAGATTAAAAGGATATCC R: CAAACTCCGCACTTGTTTACC	58	(TATC) ₇ (TA) ₃	145-200	9	0.42	0.48	EU306569
UBCLXtet_2-12	F: AGGCGACTAGACTAGCTTGAC R: GAGGAGACAATGAGACCATGAC	60	(GA) ₂ A(TA) ₂ (TCTA) ₈	240-300	5	0.11	0.60*	EU306570
UBCLXtet_2-11	F: ACACGAGAACAATAGATGTGTC R: CTTGCAGGTGATCTCTTAAC	58	(TATC) ₈ (TA) ₈	135-200	9	0.51	0.54*	EU306571
UBCLX_1-10	F: TCGGCGCTTGATAGAGCCTAGG R: GTGATCTCATATTTAAGTGATTC	58	(TATC) ₃ (TA) ₂ TCTA(TATC) ₉	250-310	9	0.83	0.74	EU306572
UBCLXA4_1	F: GATAAATTTGAAGCCTGAACAC R: GCAATTCCTCTCAAGCCATAC	62	(ACAT) ₁₇ (AT) ₁₁	190-230	11	0.63	0.72	EU306573
UBCLXdi_21	F: CTAGAGTGTGCTCTTTTAAAACC R: CTACCTTTATTCTCCTGCACC	55	(TC) ₁₆	300-350	7	0.27	0.65*	EU306574
UBCLXdi_16	F: CCTTCGGTTTCCTTACTATATG R: TGTACCCAAGGATGCACAAAGA	58	(GT) ₁₂ CA(GT) ₅ (GA) ₂₀	204-280	19	0.20	0.80*	EU306575

T_a annealing temperature; bp base pairs; A number of alleles; H_o observed heterozygosity; H_e expected heterozygosity; G.A.N. gene accession number

* denotes significant departure from Hardy-Weinberg equilibrium