THE GENETICS OF SELECTIVE BREEDING IN WESTERN HEMLOCK *TSUGA HETEROPHYLLA*

by

HUGH FRANCIS WELLMAN

B.Sc., The University of British Columbia, 2001

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

MASTERS OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

GENETICS PROGRAM

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

August 2004

Library Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

HUGH WELLMAN
Name of Author *(please print)*

09/09/2004
Date (dd/mm/yyyy)

Title of Thesis: THE GENETICS OF SELECTIVE BREEDING IN WESTERN HEMLOCK TSUGA HETEROPHYLLA

Degree: MSc Year: 2004

Department of GENETICS
The University of British Columbia
Vancouver, BC Canada

ABSTRACT

Rates of genetic diversity were inferred from allozyme allelic variation and used to compare differing seed crops that had originated from seed orchards under various seed production conditions with natural stands of western hemlock (*Tsuga heterophylla* (Raf.) Sarg). The supplemental mass pollination (SMP) technique was found to maximize both H_t and H_o when compared with unimproved wind pollinated seed crops and controlled full-sib crosses. This was identified as being the result of minimized self-fertilization and an improved parental balance. It also can be noted that 8 private alleles were lost in the seed orchard populations when compared with natural populations. Results of an inbreeding analysis have suggested that when compared with natural populations, inbreeding decreased in the wind pollinated unimproved and controlled full-sib seed crops of first generation seed orchard. The SMP treatment also showed lower values of inbreeding and when compared with controlled crosses had the highest genetic diversity. These finding suggest that the SMP technique best maintains genetic diversity and minimizes inbreeding while retaining the required selection intensity for genetic gains.

Fifteen microsatellite markers were isolated from western hemlock genomic DNA and six of these markers were optimized for use in mountain hemlock [*Tsuga mertensiana* (Bong) Carr.]. The developed SSR's were then implemented in a test of the Heterozygosity-Fitness Correlation (HFC) in western hemlock elite families. In general no significant relationship was identified between the genetic distance/relatedness of top western hemlock families and the fitness of their progeny based on phenotypic indicators. One exception were the MKL and STV progeny trials which were found to be significantly correlated at 95%. They both showed negative slope indicating that as genetic distance increase, height decreases which is indicative of outbreeding depression. Reasons that may have lent to the null result include limited parental lines, concluding in too little distance variation, a low degree of relatedness resulting in very little inbreeding depression and very little evidence for linkage disequilibrium, a cause of associative overdominance. The lack of structure identified in Western hemlock populations may not have allowed for strong heterosis therefore resulting in the poor predictive power of genetic distance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I am indebted to many people who helped me throughout this thesis. I would like to thank both, my supervisor, Dr. Kermit Ritland and committee members Dr. Carol Ritland and Dr. Yousry El-Kassaby. Without their encouragement and useful commentary I would have never finished this project. Carol in particular was always there when I had a dire lab issue. Our lab technician, Allyson MisCampbell was the other shinning beacon when you ran into lab problems. I also must thank all my lab mates, Jaclyn Beland, Dilara Ally, Charles Chen, Cherdsak (Liew) Liewlaksaneeyanawin, Jennifer Wilkin and Yanik Berube who were always there to put things in perspective. Yanik was especially helpful in all aspect of this rollercoaster ride.

I must also thank Charlie Cartwright and the B.C. Ministry of Forests, with which there would have never been a project without their provided data. Their countless man hours are greatly appreciated. I also have to thank one of my best friends and school buds, Graeme Alexander who managed to pull me away for lunch at just the right times. I would have never been able to maintain my cool perspective on this whole project without him.

I acknowledge funding through the Centre for Forest Gene Conservation, initially from Forest Renewal British Columbia, and subsequently from the Forestry Investment Account of British Columbia, obtained with support from the Forest Genetics Council of British Columbia.

I have to thank all of friends, new and old, as they all in their own little way have helped me along this path. Last but not least I have to thank my beautiful girlfriend, Caryn Lubiner. Her "kind" encouragement was the real clincher in all of this. I am completely owing to her tactful way of keeping me focused.

CHAPTER 1: General Introduction

The development and application of molecular biological techniques has become a viable way of making significant genetic gains, in major agricultural crops. This has been shown specifically in crop yield (Stuber et al. 1999). Molecular techniques, specifically molecular markers, have more recently been implemented into a limited number of conifer breeding programs (Walter 1998). Molecular markers can be used to both, monitor allelic diversity, therefore conserving genetic diversity and identify heterozygosity and fitness correlations that may allow the detection of individuals with superior performance. Phenotypic gains have been significant in conventional conifer breeding programs and further potential gains may be achieved through the use of molecular makers (King et al. 1997). The objective of this study is to test molecular marker techniques in the current Western hemlock [*Tsuga heterophylla* (Raf.) Sarg.] breeding program.

Conventional conifer improvement programs have existed for a century but intensive breeding programs were initiated in the 1940's. Selections were initially made through bulk collection of the best ("plus") trees as identified visually in natural systems (Shelbourne et al. 1989). The intensive programs consisted of selections by region, followed by development of both breeding and seed production populations. Reselection, based on progeny performance, then defined "elite" conifer families (Shelbourne et al. 1989). Large amounts of genetic improvement have been obtained in many conventional conifer programs (King et al. 1997).

Molecular markers, through their ability to reveal molecular differences between individual trees, are of potential use for breeding at two levels. At the individual-locus level, they can be associated with quantitative trait loci, and used as the basis for "marker aided selection" however this has yet to realize any industrial application in forest tree breeding programs. At the genome-wide level, they can provide a fine-scale picture of relatedness between individual trees. This application of genetic markers has the potential to minimize the effects of either inbreeding or outbreeding depression, or to predict heterosis.

Inbreeding depression occurs normally as a consequence of recessive deleterious alleles that are expressed in the progeny of matings between closely related individuals (Wright 1921), or by the loss of overdominance effects in homozygous inbred progeny (Crow 1948). In contrast outbreeding depression occurs when locally adapted genes are lost due to the mating of distantly related individuals (Templeton 1986). In either case, the extent of homozygosity at marker loci can be used as a predictor of the level of inbreeding or outbreeding depression in mature trees.

Alternatively, if heterosis occurs, markers may also be used as a predictive tool for the identification of elite genotypes. Increased heterozygosity will occur when distantly related individuals mate. If overdominance is a general phenomenon across the genome, heterozygosity at marker loci will correlate with fitness. This trend is known as hybrid vigor or heterosis and can be quantified using the heterozygosity-fitness correlation (HFC) (Shull 1952). However, this effect may be balanced against the occurrence of outbreeding depression.

In this chapter I will first review both the ecology and genetics of Western hemlock. I will then review the breeding characteristics, the history of the current breeding program across the Pacific Northwest, and the impact of selective breeding in Western hemlock. Finally I will review the heterozygosity-fitness correlation (HFC). This correlation may be used as a predictive technique of more fit individuals in selective breeding programs.

1.1 Biology of Western hemlock

Western hemlock, *Tsuga heterophylla* (Raf.) Sarg. also known as Pacific hemlock and west coast hemlock, is a member of the largest conifer family, Pinaceae. Of the 14 species of *Tsuga*, Western hemlock is one of only 3 hemlock species found in Canada. It is found mainly on the west coast of the country, although, it does maintain some interior habitat in British Columbia. Western hemlock grows to an average age of approximately 400 years but an individual was recorded at an excess of 700 years in the Queen Charlotte Islands (Hepting 1971). Western hemlock, on average, reaches diameters greater than 100 cm and heights of 60 meters.

Western hemlock exhibits some of the most productive growth in the world (Packee 1990). It is a very shade tolerant species but is sensitive to dry soils and frost (Owens and Molder 1984). Given Western hemlock's shade tolerance it is found to regenerate efficiently on shaded, moist sites (Williamson 1976). Stands investigated in the Cascade Head Experimental Forest in Oregon have recorded foliar biomass averaging 22,724 kg/ha, significantly higher than estimates of highly productive Douglas-fir (*Pseudotsuga menziesii* (Mirb) Franco.) stands, 12,107 kg/ha (Fujimori 1971). Western hemlock performs well in a mild, humid climate with constant precipitation. This species is found in areas with mean annual precipitation ranging from 380 mm in Alaska to 6650 mm in British Columbia (Packee 1976).

1.1.1 Species Range

Western hemlock, found primarily in coastal areas, is distributed in a narrow band along the Pacific Coast from Kenai Peninsula in Alaska to Northern California (Figure 1.1). This range extends a distance of 3200 km (Packee 1990). It is also scattered amongst the humid interior of British Columbia, northeastern Washington, northern Idaho, and northwestern Montana (Owens and Molder 1984). This species is found mainly at low elevations, below 600 m, but it has been found up to 2000 m in the Rocky Mountains (Hosie 1975). Western hemlock is associated with many species in the coastal regions. Some major associations in the area studied include Sitka spruce (*Picea sitchensis* (Bong.) Carr.), Alaska-cedar (*Chamaecyparis nootkatensis* (D. Don) Spach.) and Pacific silver fir (*Abies amambilis* (Doug.) ex. Forbes) (Packee 1990). With higher elevations Western hemlock is usually associated with sub-alpine fir (*Abies lasiocarpa* (Hook.) Nutt.) (Owens and Molder 1984). This species grows on many soil types; however, height growth has been shown to decrease with increasing clay content or soil bulk density. This has been attributed to poor soil aeration or the inability of roots to penetrate compact soils (Wooldridge 1961).

1.1.2 Economic and Biological Importance

Western hemlock is recognized as an all-purpose raw material in the forest industry (Packee 1990). Its strength and nailing characteristics have led to its distinction as a preferred species for construction timber. Hemlock ranks third in British Columbia for volume cut annually (Owens and Molder 1984). In 1997 sowing requests were 7.9 million (Webber 2000). This species has been defined as good to excellent in its pulping characteristics. Its fiber is a major source for groundwood, thermomechanical, kraft, and sulfite pulps (Packee 1990). Also, its fine grain and resin free characteristics, have defined it as a suitable finishing lumber. Given its diverse economic attributes, this species also plays a major ecological role. Western hemlock is an important browse species for deer and elk and makes up a significant portion of the canopy for many national parks in both Canada and the United States.

1.1.3 Reproductive Characteristics

Western hemlock, a monoecious tree, is a good cone and seed producer. Good cone crops occur by about 25 years of age (Owens and Molder 1984). Cones are primarily found in the upper regions of the crown, but this is dependent upon density. With a less dense stand, seed cones can be found throughout the crown. A good crop occurs approximately every 3-4 years but some seeds are produced every year (Fowells 1965). Estimates of filled seeds in cones have been recorded at an average of 10 to 20 per cone (Meagher 1976). Release of seed, that is, cone scale opening, is dependent on moisture content in the atmosphere (Packee 1990). Seed travel, following release, has been recorded at distances of more than 1.6 km in a strong wind, but average distance is estimated at about 600 m in a 20 km wind (Isaac 1930). Its prolific seed production and high shade tolerance make Western hemlock a perfect candidate for management.

1.1.4 The Genetics of Western hemlock

Very little molecular genetic work has been performed to date on Western hemlock. Some extrapolation can be made by investigating Western hemlock's morphological characteristics. Western hemlock has a large and continuous geographic distribution, it is wind pollinated and has small seed; therefore it should show high gene flow (Mosseler et al. 1994). This would lead to the inference that this species should show very little population differentiation. (Hamrick and Godt 1990) note that most molecular-based studies, particularly those using isoenzyme variation, suggest low levels of interpopulation differentiation in trees, with less than 10% of the total genetic variation occurring among different populations. Some other characteristics of Western hemlock that would promote high gene flow and therefore low genetic structure would include its breeding structure. This species favors cross-pollination, and extensive pollen range and high population density leading to high gene dispersal.

Some physiological studies have been performed using Western hemlock. Malavasi and Perry (1993) found, in a shade tolerance study, variability within populations. They suggest that Western hemlock's competitive variability is a function of high genetic variation within populations. This study used the competitive ability of the shade intolerant conifer, *P. menziesii*, for comparison. This study was correlated with findings made in an allozyme study, using different conifers that showed similar results (Hamrick et al. 1981).

One common garden study showed latitudinal differentiation in Western hemlock, when comparing seedlings in California through to Alaska (Kuser and Ching 1980, 1981). The 21 Western hemlock provenances showed trends in cold hardiness, survival and seedling growth factors. It is also of interest, that this trend was no longer significant in any of the three factors, when the distribution was minimized to a range of 46° to 51°. Foster and Lester (1983) found similar results using height as their indicator. They found no differentiation in a 3° latitude distribution of Washington families of Western hemlock.

Given little differentiation in terms of latitude, (King 1991) did find significant differentiation when studying elevation. Western hemlock families that grow well above

600 m were found to be significantly different then families that grow below 600 m. Given the significant findings, for the most part this is the exception, as very little Western hemlock is found above 600 m. The high potential for gene flow in Western hemlock makes differentiation unlikely (King 1991). King (1991) notes that gene flow both by gamete and zygote are so extensive that there would be very little chance for genetically distinct demes to develop, resulting in very little geographic differentiation. This can be extended to a population context. With high gene flow you would also expect very little among population differentiation, therefore leading to all the variation being harbored within populations.

Realized gain trials of Western hemlock, carried out by the B.C. Ministry of Forests, found orchard mean seedlots out-performing wild stands by 5 to 8% (Cartwright 2001). Some elites have been recorded with gains greater than 20% (Cartwright 2001). This is much higher than the theoretical 2% assigned to unimproved (Un) seed orchards. Possible explanations include a higher level of inbreeding in wild collection or efficiency of first generation selection (Ritland 2000). No molecular work to date has looked at this anomalous genetic gain.

1.2 Western hemlock Breeding Programs and Design

The goal of all tree breeding programs is to produce premium seedlings that show good performance when out-planted in natural systems. The expression of seedlings performance is related to that seedlings growth potential and the degree to which the field conditions allow this potential to manifest (Grossnickle and Sutton 1999). Tree improvement in western hemlock began in the 1970's; 2000 families are represented at present in seed orchards and provenances today.

The Western Hemlock International Cooperative was created in 1991 (the Western Hemlock Tree Improvement Cooperative, HEMTIC). Most Western hemlock orchards and provenances are managed within this cooperative to date. The objective of HEMTIC is to produce parents for seed production that can provide genetic gain for wood quality and quantity with well adapted genotypes at appropriate cost levels. Their objective is achieved by capitalizing on extensive investment already made in progeny

tests and seed orchard development. HEMTIC selection is, for the most part based on height data collected in 75 open-pollinated progeny test installations in 5 different programs (BC, Forks (Rayonier), Grays Harbour (Rayonier, QIN), CZ/ Willamette (formerly Cavenham), and Tillamook). The primary goals of HEMTIC are as follows:

1. the construction of a filial (F-1) generation from diallel mating among such first-generation orchard clones.
2. the reselection of these parents for use as high-gain seed orchard stock ('A' parents); (King and Cress 1991)

The B.C. program will now be described in detail as it is the oldest and most developed of the programs. Comprehensive long-term trials have never been planted in Oregon or Washington for Western hemlock (Jayawickrama 2003).

1.2.1 British Columbia Program

Western hemlock comprises 40% of the coastal harvest and 60% of the provinces export market. Natural regeneration is prolific in western hemlock but there is currently 8 million hemlock seedling planted annually (King et al. 1997). Given this regeneration capacity combined with the biological diversity observed in western hemlock stands, this species is a prime candidate for tree improvement (Webber 2000). The cost of planting stock is compensated by the faster green up and the more even spaced stocking. The utilization of improved seed can therefore justify hemlock plantations. Production from earlier established seed orchards has contributed to 41, 48, and 44% of the annual sowing requests in British Columbia for 1996, 1997 and 1998, respectively (SPAR nursery extract files, B.C. Ministry of Forests).

The Background of B.C.'s Western hemlock Tree Improvement

Parent tree selection began in 1959 but most of the plus tree selection was made between 1975 and 1981. As of September 1997 a total of 1339 plus tree selections had been made. Some original work was carried out by the Canadian Forestry Service (CFS) including provenance trials on 4 sites on Vancouver Island and a half-sib test and clone bank at Cobble Hill, Vancouver Island (Piesch 1974).

Some family-level genetic information could be discerned from the Thasis trial. Two year old seedlings were planted at Gold River in 1970 and 1972 (site Tahsis 15). This site included 58 open-pollinated families, tested by 4 replications of 50 tree row plots. Height and diameter data are now available for ages 5, 10, 15, and 20 years. This row plot design was not effective for discerning family-level differences so it has not been used as a stand alone progeny test. However, this information has been quite useful as a model to compare to later testing (MM series) carried out by Mike Meager.

Three series of open-pollinated progeny tests, (MM-79, MM-80, MM-81) comprising a total of 144 OP families were established on 23 test sites throughout Vancouver Island and the lower mainland in 1979, 80 and 81 respectively (King et al 1998). Continued progeny testing was carried out to enlarge the pool of tested B.C. western hemlock parents. Most of the first generation seed orchard parents were placed under test, using a polycross. This series included PX-91, PX-92, PX-93, and PX-95.

Genetic selection for wood and fiber properties has also been deemed as quite profitable given relatively high heritability and population variability (King et al. 1998). Strong family effects have been identified for both the pulp and paper characteristics in western hemlock. Family effects have been noted in the internal wood properties of relative wood density, average fiber length and fiber coarseness with heritability estimates ranging from $h^2_i = 0.5$ to 0.9 (King et al. 1998).

MM-Series (MM79, MM-80 and MM-81)

Each progeny test site covered a specific seed zone as delineated in the old seed orchard planning zones (Crown 1981). The MM-79 series tested 29 parents from the South-coast Mainland (SCM) seed planning zone. The MM-80 series tested 39 parents, 35 from West Vancouver Island and 4 from coastal areas of Washington's Olympic Peninsula. The MM-81 series tested 76 parents from North Vancouver Island and Johnstone Straits planning zones. Each series were tested over approximately 8 sites, 6 of these sites typically lying within the seed planning zone and 2 placed outside the seed planning zones. Five, 10, 15 and 20 year heights and diameters have been measured and analyzed. Fifteen year wood density, pulp and fiber traits, and stem form data have also been collected on a limited number of sites. Selection involved backward selection

utilizing 15 year height data. It should be noted that a limited number of forward selections (the 3000 series) were also made.

The MM series was set up as Randomized non-contiguous complete block design (RCB) with every family randomly distributed 16 times across each of 4 replicated blocks (King et al. 1997). Selection was carried out using the site information as separate traits in a selection index based on progeny test information. The parents now represented in seed orchards and clone banks are the original plus trees parents with selection being based on their progeny means. Individual-tree narrow-sense heritability was measured at 0.14(0.05) for the MM-79 series. In the MM-80 series only two sites were deemed adequate for measurement at age 15. The individual narrow sense heritability for this series was measured at 0.06 (King et al. 1997). In the MM-81 series age 15 assessment indicated heritability for height growth at 0.09.

PX Series (PX-91, PX-92, PX-93 and PX-95)

This series encompassed all seed orchard clones that were not included in the open pollinated testing of the MM series. The PX Design was based on a sets in replications design with a RCB layout. The sets are a random partition of the polycross families split into 3 sets of 25 to 30 families. The incomplete block design was used to reduce the environmental heterogeneity within the test plots (King et al. 1997). Each test plot is a rectangular square of about 4x7 trees that will include a single tree representing each family of the set plus one or two control lots. Heritability estimates for height at age 5 were 0.12 for PX-91 and 0.007 for PX-92 (these estimates are higher than the corresponding heritability from the MM series (King et al. 1997).

Seed Orchard Programs and Materials

The seed orchard programs are based on the philosophy of having small zones where locally adapted seed sources can be implemented. Initial development called for 13 seed orchard planning zones established in the coastal B.C. areas (King et al. 1997). Ten different seed orchards for western hemlock were established to cover high and low elevations in at least 6 of the 13 seed orchard planning zones. The Forest Genetics Council plan lists eight orchards for the maritime low seed zone and five for the maritime

High seed zone (FGC 1998). The current level for genetic worth is between 5 and 10% for the maritime low with anticipated levels rising to 20% by 2007 (FGC 1998).

A 6 year seed surplus has been accumulated for the 10 low-elevation coastal hemlock seed orchards (King and Brown 1993). Forthcoming HEMTIC materials are far superior in terms of total gain potential and have therefore since replaced these orchards in terms of full management. Most BC cooperators have established new low elevation hemlock seed orchards, based on HEMTIC materials which are expected to have gains of 18% at minimum in final rotation age volumes. Orchards representing high elevation and the Queen Charlotte Islands are still being maintained and managed by Timber West, Western Forest Products and the B.C. Forest Service. Limited potential is expected from these orchards. A total of 325 parents (MM and PX series) have now been tested for the low-elevation south coast (south of 51 degrees and below 300m) and a total of 130 high elevation parents have been tested (MM-79, PX-93 and PX-95) (King et al. 1997).

1.2.2 Washington and Oregon Programs

Western hemlock occupies a significant area along the coast of both Washington and Oregon (Figure 1.1). Historically very little nursery grown seedling were produced given poor seedling quality and a low stumpage value. But with improvements in nursery technology and adaptational concerns with alternative species such as Swiss Needle Cast in Douglas-Fir, western hemlock has become a significant part of nursery stock.

The Industrial Forestry Association started selective breeding centered around Douglas-Fir as early as the 50's (King et al. 1997). Breeding co-op's were first organized in 1966 in western Oregon and expanded to include western Washington in 1986 when it was reorganization as the Northwest Tree Improvement Cooperative (NWTIC). The NWTIC association currently serves 33 member organizations engaged in over 75 local co-op programs across roughly 7 million acres (3.2 million hectares) of commercial forest land. Originally, Co-op's were primarily set up for Douglas-Fir but 2 co-ops, Tilamook co-op of Oregon and the Bray's Harbor co-op of Washington were specifically designed for western hemlock. Three programs independent of the NWTIC also exist for western hemlock, including the Forks program, Rayonier Inc., the Crown Zellerbach /

Willamette program (formerly known as Cavenham; now part of Willamette Indistries Inc.) and a Weyerhaeuser program (King et al. 1997). All of these programs have since been incorporated into the HEMTIC cooperative.

The programs included on the NWTIC co-op were generally laid out as replicates-in-sets. The design typically included the testing of 2 to 10 sets of approximately 30 families across each of 4 to 10 test sites. Generally each family was represented by 3 to 5 replications of 4 to 6 open-pollinated progeny per site.

Forks Program

This program was set by Rayonier Inc. with plus selection beginning in 1973. A total of 278 open-pollinated Western hemlock families were planted in 1979 and 1980 (King et al. 1997). To date 5, 10 and 15 year measurements have been made. The families were planted in an interlocking block design. Approximately 2 meter separated each tree and the design incorporated 6 reps on 4 sites. A seed producing orchard for the original parents has been established in proximity to Sequim Washington.

Gray's Harbor Program

In 1978 the Quinault Indian Nation (QIN) began plus tree selections of western hemlock. Open pollination resulted in 151 families which were planted on 4 testing sites the following year (King et al. 1997). Selections were also made by Rayonier Inc. resulting in 137 families on 9 sites which were planted in 1982. QIN, Rayonier, and Boise Cascade combined all programs to form a coop in 1981 and continued to make selections resulting in one more test site in 1983 and 8 more in 1983. The coop then represented a total of 447 families in 10 sets across 8 test sites. Spacing was generally 2m x 2m and measurements for 5 and 10 years have been made to date (King et al. 1997). This material has since been forgone due to the introduction of higher-gain HEMTIC material.

Willamette Program

Crown Zellerbach established some test sites based on selections made in the mid 1970's. Sites were set up to utilize a thinning design and some selection was made for

both volume and diameter (King et al. 1997). To date 5, 10 and 15 year measurements have been made.

Tillamook Cooperative Program

Four cooperators, BLM, USFS, the Oregon Department of Forestry and Publishers Papers initiated a selective breeding program in 1975. This coop is represented in 10 open pollinated progeny trials based on 270 selections which were planted in 1976. To date 18 parents have been identified and have been measured through year 15.

1.3 Selective Breeding in Western hemlock

In outcrossed species, erosion of population genetic diversity can lead to loss in viability and reproductive success (Mosseler et al. 1994). With changes in the environment, selection for reduced fitness can ensue. This can have profound consequences on the stability of the ecosystem. Therefore, maintaining genetic diversity is an important factor in maintaining the stability of managed areas. With the use of seed orchards to restock natural systems, genetic consequences of a limited genetic pool must be realized.

When a subsample of a population is used to represent a large area of a species range, genetic diversity can be compromised. Mating of closely related individuals may lead to a higher mutation rate as well as the expression of recessive traits that could conclude in reduced fitness. A special concern must be taken when dealing with small populations as the opportunity of mating with relatives increases with decreasing population size (Franklin 1980). This is particularly true in coniferous seed orchards as genetic diversity must be maintained to handle the heterogeneity of environmental pressures that trees endure (Muller-Stark 1995). Given these concerns extensive research has focused on controlled selection for desirable traits in managed forest systems. Selective breeding can lead to a marked increase in fitness. With this increase in fitness a balance between selection and genotypic diversity must be maintained.

The genetic quality of orchard seed lots is comprised of three components: genetic worth, adaptation to the planting zone, and genetic diversity (Webber 2000). In incremental value, the British Columbian Western hemlock selective tree breeding program is ranked third behind coastal Douglas-Fir and Prince George Lodgepole pine (FGC 1998). The Western hemlock realized gain trials were established to allow the prediction of unit-area volume gains, produce growth yield information for improved stock to calibrate growth models, and to demonstrate the gain from investment in tree improvement (King et al. 1997).

Three trials were established on Vancouver Island in 1992, 1994, and 1996. The 1996 trial included 3 genetic entries at 4 spacings on 6 sites. This trial was made up of a full-sib cross "elite" or "A cross" equivalent to a 1 in 30 selection proportion (comparable in gain to the HEMTIC advanced generation orchards). The 1994 trial was composed of 4 levels of improvement represented at 4 spacings on only 2 sites. This trial used a full-sib cross representing a selection proportion of 1 in 6 or "B cross" (equivalent to a rouged first generation orchard). The 1992 trial was composed of 3 levels of improvement on 3 sites of varying site indexes. In this trial the seedlings were all "wood run" composed of 14 wild stand seedlots. All trails have a 12 by 12 tree block design (0.1 ha at 3m spacing). Gains of approximately 30% in height growth are evidenced for the advanced generation material across sites and over 3 measurements at age 3, 4 and 5 years (King et al. 1997).

1.4 The Heterozygosity-Fitness Correlation (HFC)

A positive correlation between heterozygosity and fitness-related traits has been recognized across a diverse spectrum of organisms (David et al. 1997). This correlation termed the heterozygosity-fitness correlation (HFC) is the result of hybrid vigor or heterosis. HFC can be used as predictive tool in marker-assisted breeding programs. Conventional selective tree breeding programs are extremely laborious and require lengthy testing (King et al. 1997). This is the result of the long interval before assessment of phenotype. HFC could greatly reduce the number of individuals that

would require testing, as it would identify favorable crosses. Given the basis for this examination, I now describe how HFC arises.

1.4.1 Defining Heterosis

Heterosis can be defined as the increased vigor, size, fruitfulness, speed of development, resistance to disease and to insect pests, or to climactic rigors of any kind, manifested by crossbred organisms as compared with corresponding inbreds (Shull 1952). Heterosis is best understood as an empirical observation. It can be attributed to either overdominance, partial dominance, or total dominance. In the case of overdominance, a higher fitness is conferred by a heterozygous genotype at a given locus compared to the corresponding homozygous genotype (David 1997).

As microsatellites are obtained from non-coding DNA, it seems safe to assume there is no direct relationship between marker and selected loci. Any correlation between markers and fitness occurs because of associative dominance or overdominance. Associative dominance or overdominance can be cause by either linkage disequilibrium (nonrandom associations of alleles in gametes) or identity disequilibrium (nonrandom association of diploid genotypes in zygotes) (Ohta 1971).

1.4.2 Theoretical Context

The major goal is to be able to predict heterosis before mating, in order to reduce the number of parents that would be tested. The genetic distance between two individuals at their marker loci may be indicative of the closeness of their relationship. As such, this is predictive of the number of deleterious alleles shared by the two parents, and also the expected heterozygosity of progeny. In some perennial and herbaceous plants, heterosis is thought to increase with increasing genetic divergence, then decrease at higher levels of divergence (Waser and Price 1989). Fitness is believed to be the product of a balance between inbreeding depression at one extreme of relatedness and outbreeding depression in widely divergent genotypes at the other (Lynch 1991).

Reasons that one might expect a positive correlation between individual heterozygosity and components of fitness are (Houle 1989, Charlesworth 1991):

1. Heterozygotes may have intrinsically higher fitness than homozygotes, i.e. there is a heterozygote advantage (overdominance) at the marker loci.
2. Statistical association (linkage disequilibrium) between the marker loci and other selected loci (usually linked) can result in apparent heterozygote advantage at marker loci. Secondary loci may exhibit heterozygote advantage or themselves may be dominant.
3. There may be an association of genotypes and components of fitness due to the presence of different levels of inbreeding in the population, which generates genotypic association between unlinked loci and quantified using the concept of identity equilibrium (Weir and Cockerham 1973, Charlesworth 1991).

The power to predict heterosis is low if it is determined by few quantitative trait loci (QTL) or QTL with multiple alleles randomly distributed among trees (Leonardi et al. 1991). Two reasons for a lack of correlation are:

1. Lack of association between quantitative traits measured and markers involved in measures of genetic distance (Moser and Lee 1994).
2. Sample size of parental lines may not be large enough to give accurate enough genetic distance to correctly predict heterosis (particularly interspecific).

When using microsatellites, no relationship between marker loci and the loci that are directly contributing to the observed phenotypic variation exists, so one assumes associative overdominance (Ohta 1971). This genetic correlation occurs as linkage disequilibrium or identity disequilibrium (David 1998). Linkage disequlibriuim is mainly restricted to tightly linked loci. Strong linkage disequilibrium occurs in finite populations as a result of genetic drift (Hill and Robertson 1968). In comparison, identity disequilibrium is mainly generated by partial inbreeding (Weir and Cockerham 1973). When inbreeding equilibrium exists, the value for two completely linked loci is only twice that of independent loci. In this case inbreeding generates correlations among all loci of the genome and this is referred to as "general effect". When effects are restricted to a narrow chromosomal segment around the target locus and vanish with distance, it is

known to have a "local effect" (David et al. 1995). This is what is expected in finite populations. It should be noted that with linkage disequilibrium, two inbreds may have similar phenotypes, share common alleles at QTL's and display relatively high marker distances. I now examine how this relationship can be used in an applied context.

1.4.3 Applied Context

The heterozygosity fitness correlation has a long history of study in agricultural species. Species such as *Zea mays* (Dudley et al. 1991) (Lanza et al. 1997) (Melchinger et al. 1998) and rice (Liu and Wu 1998) have been subjected to countless studies in which the GDH correlation has been tested with varying degrees of success. To date there have been several studies that have tested for this correlation in conifers (Savolainen and Hedrick 1995, Vaillancourt et al. 1995, Arcade et al. 1996, Harfouche et al. 2000).

Arcade et al. (1996) looked at heterozygosity and hybrid performance in larch. Random Amplified Polymorphic DNAs (RAPD) was employed for the study of 12 European larches and 12 Japanese larches. One hundred and twelve fragments were used to measure genetic distance using Jacard's coefficient. They found a significant positive correlation between genetic distance of parents and performance of the hybrids. This was identified in total height, 6^{th} year stem increment, breast height girth and stem volume. No correlation was found for taper, number or diameter of branches, stem straightness or wood specific gravity. An interesting result noted in this study was that the correlation between genetic distance and height reached a maximum at age 6. Given the increase, Arcade et al. (1996) notes that the correlations were not significant beyond the age of 4 years and height may be controlled by additive effects after this point.

Vaillancourt et al. (1995) tested the ability to predict heterosis in Eucalyptus. Eight females and nineteen males were used from two provenances in five geographical areas. Sixty-six polymorphic RAPD loci were used to estimate Nei's genetic distance. Conic volume based on height and diameter was used as a measure of fitness. Genetic distance significantly predicted heterosis in Eucalyptus, but accounted for less than 5% of the variation and more accurately reflected total genetic differentiation between provenance crosses (Vaillancourt et al. 1995). Vaillancourt et al. (1995) pointed to the

low degree of relatedness that existed between parents as a possible reason for the poor predictive capabilities of volume.

Harfouche et al. (2000) examined the performance of provenance hybridization in a diallel mating scheme of maritime pine (*Pinus pinaster*). This research project included ten provenances belonging to three biosystematic groups. Parental genetic distance was estimated with various genetic markers and phenotypic traits, as well as terpenes and denatured proteins. A weak but significant positive correlation was noted between genetic distance and height. No significant relation ship was noted for stem crookedness and insect resistance. Harfouche et al. (2000) also found that the variance of heterosis for height growth decreased with increasing age.

Savolainen and Hedrick (1995) looked at the association of heterozygosity and fitness in Scots pine. In this study, 3 populations of Scots pine were used to associate six quantitative traits related to fitness with heterozygosity using 12 allozyme loci. No association was evident in this study. Between only 7 and 8% of their regression tests were significant at the 5% level. The relatively small amount of information about heterozygosity given by isozymes, due to their limited numbers of loci and few (2-4) alleles, may be a possible reason for such low correlation. With the background now outlined the objectives can be presented.

1.5 Objectives for the Thesis

1.5.1 Evaluation of the Genetic Effects of Domestication under Differing Orchard Conditions using Isozymes

When seed is selected from natural systems, some allelic diversity is lost. Therefore, various seed orchard management techniques have been developed to maintain diversity in seed orchards. Originally, seed orchard operators had very little control over pollen parentage, as trees were passively open-pollination. Some control of pollination was achieved by creating buffer areas in which other species were planted around orchards. The buffer zone system is not totally effective, resulting in diluted genetic gain.

Two techniques were then developed to minimize pollen contamination, supplemental mass pollination and controlled pollination. Controlled pollination is when pollen buds are removed from part of a branch, and female strobili are isolated with a paper/plastic bag and pollen from only one other parent is then applied. The supplemental mass pollination orchards have pollen from best trees applied to receptive clones.

In my study, isozymes will be used to compare heterozygosity and genetic diversity between differing seed orchard systems and natural populations. This will be done firstly, to identify losses of genetic diversity, and secondly, to identify which seed orchard technique best maintains genetic diversity.

1.5.2 Testing the Role of Inbreeding in Relation to Domestication

Breeders have both long term and short term objectives. In the short term, one seeks substantial gains in fitness, and in the long term, one wants to maintain allelic diversity and control inbreeding (Johnson 2001). A conflict arises, as a large breeding population is required to maintain allelic diversity and substantial gains require high selection intensities, therefore limiting the breeding population (Johnson 2001). Consequently, a breeder must try to balance between objectives. Different pollination techniques may have influences over rates of inbreeding and genetic structure. By identifying which technique best maintains diversity and minimizes inbreeding, the breeder can make more informed decisions.

Realized gain trials of Western hemlock, carried out by the B.C. Ministry of Forests, found orchard mean seedlots out-performing wild stands by 5 to 8% (Charlie Cartwright, pers. comm.). This is much higher than the theoretical 2% assigned to unimproved (Un) seed orchards. Two alternative explanations for this are (1) inbreeding effects expressed in wild collections are removed, or (2) the efficiency of first generation selection is greater than expected (Ritland 2000).

In my study, isozymes will be used to compare genetic structure and rates of inbreeding amongst orchards and natural populations. The will be done to identify the rates of inbreeding under differing pollination techniques and to determine if the anomalous genetic gain is the result of the removal of inbreeding in natural populations.

1.5.3 Development of Western hemlock Microsatellites

Simple sequence repeats (SSR) or microsatellites have become the most abundantly used marker in population based studies to date (Zane et al. 2002). This is primarily owing to their abundance throughout the genome and their hyper variability while compared with other molecular markers (Zane et al. 2002). Further advantages include the very small amount of DNA required for amplification and their repeatability. In addition microsatellites are selectively "neutral" and are detected in a co-dominant fashion. Therefore microsatellites are likely the best choice for measuring variation in heterozygosity, and are of greatest value in marker-assisted tree breeding programs.

Microsatellites require lengthy development as primers are generally specific to the genus. This is the case as microsatellites are generally found in non-coding regions where the substitution rates are high (Zane et al. 2002). The development cost is offset by the relatively low cost of use, once primers have been designed. To date, only 5 primer pairs have been developed for Western hemlock (Amerasinghe et al. 2002), and in house testing of these primers has indicated poor repeatability of these markers (Dilara Ally pers. comm.). Further development is required for full implementation of microsatellites into a marker-assisted Western hemlock breeding program, and in this chapter, I present the results of this development.

1.5.4 Testing the Heterozygosity-fitness Correlation in Western hemlock and Evaluate its Ability in Predicting Specific Combining Ability

One strategy that may be implemented into the current Western hemlock breeding program is the use of the heterozygosity-fitness correlation (HFC). HFC can be used to predict the fitness of offspring before mating, based on the genetic distance between their parents. The use of HFC could ultimately reduce the number of parents that would have to be tested by identifying superior crosses. Fitness has been shown to amplify with increasing genetic distance in many species including, *Zea mays* (Dudley et al. 1991; Melchinger et al. 1998), marine bivalves (Zouros 1987), *Oryza sativa* (Zhang et

al. 1994) and *Phaseolus vulgaris* (Beattie et al. 2003). The prediction of F1 performance is of particular value to conifer breeders as field testing is expensive and labor-consuming. Moreover the delays before assessment of individual performance are long (Neale and Williams 1991). If a strong relationship can be defined for HFC, heterozygosity can be maintained through selection for greater genetic distance while increasing the overall fitness of the next generation population.

Western hemlock is prime candidate for the application of HFC as it has a long history of selective breeding, with significant genetic gains (King et al. 1997). Progeny trials have been developed for crosses between top elite families (Charlie Cartwright per. Comm. 2003). Five year height, presence of frost damage and forks (damaged leader) are now available for F1 seedlings.

In this chapter, I genotype the parents of these families for microsatellites, and then correlate this F1 phenotypic data with the genetic distance of seedling parents. Any significant correlation will be indicative of HFC and may allow for the prediction of F1 performance in Western hemlock seedlings in the future.

1.6 References

Arcade, A., P. Faivre-Rampant, B. Le Guerroue, L. E. Paques, and D. Prat. 1996. Heterozygosity and hybrid performance in Larch. Theoretical Applied Genetics **93**:1271-1281.

Amerasinghe V, Brown GR, Mank JE, Carlson JE (2002) Microsatellite DNA loci for Western Hemlock [*Tsuga heterophylla* (Raf.) Sarg]. *Molecular Ecology Notes* **2**, 236-238.

Beattie, A. D., T. E. Michaels, and K. P. Pauls. 2003. Predicting progeny performance in common bean (*Phaseolus vulgaris* L.) using molecular marker-based cluster analysis. Genome **46**:259 - 267.

Cartwright, C. 2001. Western Hemlock. *in* H. Wellman, editor., Vancouver B.C.

Charlesworth, D. 1991. The apparent selection on neutral marker loci in partially inbreeding populations. Genetical Research **57**:159-175.

Crow, J. F.. 1948. Alternative hypotheses of hybrid vigor. Genetics. **33**:447 - 487.

Crown, M. 1981. B.C. coastal tree improvement council 1st progress report 1979 to 1981. B.C. Ministry of forests, Victoria B.C.

David, P. 1997. Modeling the genetic basis of heterosis: tests of alternative hypotheses. Evolution **51**:1049 - 1057.

David, P. 1998. Heterozygosity-fitness correlation: new perspectives on old problems. Heredity **80**:531-537.

David, P., B. Delay, P. Berthou, and P. Jarne. 1995. Alternative models for allozyme-associated heterosis in the marine bivalve *Spisula ovalis*. Genetics **139**:1719-1726.

Dudley, J. W., M. A. Saghai-Maroof, and G. K. Rufener. 1991. Molecular markers and grouping of parents in maize breeding programs. Crop Science **31**:718 - 723.

FGC. 1998. FGC Bussiness Plan 1998-1999. Business Plan Forest Genetics Council of British Columbia.

Foster, G. S., and D. T. Lester. 1983. Fifth-year height variation in western hemlock open-pollinated families growing on four test sites. Canadian Journal Forest Research **13**:251 - 256.

Fowells, H. A. 1965. Silvics of forest trees of the United States. Agricultural handbook 271, U.S.D.A., Washington, D.C.

Franklin, I. R. 1980. Evolutionary change in small populations. Pages 135-149 *in* M. E. Soule and B. A. Wilcox, editors. Conservation biology: an evolutionary-ecological perspective. Sinauer Associates, Sunderland, MA.

Fujimori, T. 1971. Primary productivity of a young Tsuga heterophylla stand and some speculations about biomass of coast communities on the Oregon coast. Research paper PNW - 123 USDA Forest service, Pacific Northwest Forest and Range Experiment Station, Portland, Oregon.

Grossnickle, S. C., and B. C. S. Sutton. 1999. Applications of biotechnology for forest regeneration. New Forests **17**:213-226.

Hamrick, J. L., and M. J. W. Godt. 1990. Allozyme diversity in plant species. Pages 43 - 63 *in* A.H.D. Brown, M. T. Clegg, A. L. Kahler, and B. S. Weir, editors. Plant population genetics, breeding, and genetic resources. Sinauer, Sunderland, Mass.

Hamrick, J. L., M. J. W. Godt, and S. L. Sherman-Broyles. 1992. Factors influencing levels of genetic diversity in woody plant species. New Forests **6**:95-124.

Hamrick, J. L., J. B. Mitton, and Y. B. Linhart. 1981. Levels of genetic variation in trees: influence of life history characteristics. Pages 35 - 41 *in* M. T. Conkle, editor. Isozymes of North America forest trees and forest insects. USDA Forest Service General Technical Representatives.

Harfouche, A., N. Bahrman, P. Baradat, J. Guyon, R. J. Petit, and A. Kremer. 2000. Provenance hybridization in a diallel mating scheme of maritime pine (*Pinus pinaster*). II. Heterosis. Canadian Journal Forest Research **30**:10-16.

Hepting, G. H. 1971. Disease of forest and shade trees of the United States. U.S. Department of Agriculture, Washington, DC.

Hill, W. G., and A. Robertson. 1968. Linkage disequilibrium in finite populations. Theoretical Applied Genetics **38**:226-231.

Hosie, R. C. 1975. Native trees of Canada. Canadian Forest Service, Ottawa, Ontario.

Houle, D. 1989. Allozyme-associated heterosis in *Drosophila melanogaster*. Genetics **123**:789-801.

Isaac, L. A. 1930. Seed flight in the Douglas-fir Region. Journal of Forestry **28**:492 - 499.

King, J. N. 1991. The significance of geographic variation patterns for western hemlock genetic improvement. Technical Report B.C. Ministry of Forests Research Branch, Victoria, B.C.

King, J. N., and P. Brown. 1993. Western hemlock TAC subcommittee seed orchard report 1993. internal report B.C. tree improvment council technical advisory commitee, Victoria B.C.

King, J. N., C. Cartwright, and D. Cress. 1997. Western hemlock tree improvment: selection of P-1 parents. B.C. Ministry of Forests, Victoria B.C.

King, J. N., C. Cartwright, J. V. Hatton, and A. D. Yanchuk. 1998. The potential of improving western hemlock pulp and paper quality. I. Genetic control and interrelationships of wood and fibre traits. Canadian Journal Forest Research **28**:863 - 870.

King, J. N., and D. Cress. 1991. Breeding plan proposal for western hemlock cooperative tree improvement. internal report HEMTIC cooperative.

Kuser, J. E., and K. K. Ching. 1980. Provenance variation in phenology and cold hardiness of western hemlock seedlings. Forest Science **26**:463 - 470.

Kuser, J. E., and K. K. Ching. 1981. Provenance variation in seed weight, cotyledon number, and growth rate of western hemlock seedlings. Forest Science **26**:463 - 470.

Lanza, L. L. B., C. L. Souza Jr., L. M. M. Ottoboni, M. L. C. Vieira, and S. A.P. 1997. Genetic distance of inbred lines and prediction of maize single-cross performance using RAPD markers. Theoretical Applied Genetics **94**:1023-1030.

Leonardi, A., C. Damerval, Y. Hebert, A. Gallais, and D. Devienne. 1991. Association of protein amount polymorphism (PAP) among maize lines with performance of their hybrids. Theoretical Applied Genetics **82**.

Liu, X. C., and J. L. Wu. 1998. SSR heterogenic patterns of parents for marking and predicting heterosis in rice breeding. Molecular Breeding **4**:263-268.

Lynch, M. 1991. The genetic interpretation of inbreeding depression and outbreeding depression. Evolution **45**:622-629.

Malavasi, U. C., and D. A. Perry. 1993. Genetic variation in competitive ability of some shade-tolerant and shade-intolerant pacific Coast (USA) conifers. Forest Ecology and Management **56**:69 - 81.

Meagher, M. D. 1976. Studies of variation in hemlock (Tsuga) populations and individuals from southern British Columbia. University of British Columbia, Vancouver, B.C.

Melchinger, A. E., R. B. Gumber, R. B. Leipert, M. Vulysteke, and M. Kuiper. 1998. Prediction of testcross means and variances among F3 progenies and F1 crosses fro testcross means and genetic distances of their parents in maize. Theoretical Applied Genetics **96**:503-512.

Moser, H., and M. Lee. 1994. RFLP variation and geneological distance, multivariate distance, heterosis, and genetic variances in oats. Theoretical Applied Genetics **87**:947-956.

Mosseler, A., K. N. Egger, and S. M. Carr. 1994. Molecular biology and genetic diversity in tree populations. Pages 114 - 124 *in* Recent progress in Forest Biotechnology in Canada. Charest and Duchense.

Muller-Stark, G. 1995. Protection of genetic variability in forest trees. Forest Genetics. **2**: 121-124.

Neale, D. B., and C. G. Williams. 1991. Restriction fragment length polymorphism mapping in conifers and applications to forest genetics and tree improvment. Canadian Journal Forest Research **21**:545 - 554.

Ohta, T. 1971. Associative overdominance caused by linked detrimental mutations. Genetical Research **18**:277-286.

Owens, J. N., and M. Molder. 1984. The reproductive cycles of western and mountain hemlock. Ministry of Forests, Victoria B.C.

Packee, E. C. 1976. The ecology of western hemlock. *in* W. A. Atkinson and R. J. Zasoski, editors. Western Hemlock Management Conference, University of Washington College of Forest Resources, Seattle.

Packee, E. C. 1990. Tsuga heterophylla (Raf.) Sarg: Western hemlock. *in* R. M. Burns and B. H. Honkala, editors. Silvics of North America: Conifers. U.S. Department of Agriculture, Forest Service, Washington, DC.

Piesch, R. F. 1974. Establishment of western hemlock tree improvment program in coastal British Columbia. PFRC information report BC-X-89, Canadian Forest Service, Victoria B.C.

Ritland, K. 2000. Forest renewal B.C. operational improvement proposal. Forest Genetic Council of British Columbia, Vancouver, B.C.

Savolainen, O., and P. Hedrick. 1995. Heterozygosity and Fitness: No Association in Scots Pine. Genetics **140**:755-766.

Shelbourne, C. J. A., M. J. Carson and M. D. Wilcox. 1989. New techniques in the genetic improvement of radiata pine. Commonwealth Forest Review **68**: 3.

Shull, G. H. 1952. Beginnings of the heterosis concept. Pages 14-48 *in* J. W. Gowen, editor. Heterosis. Iowa State College Press, New York.

Stuber, C. W., M. Polacco and M. L. Senior. 1999. Synergy of empirical breeding, marker-assisted selection, and genomics to increase crop yield potential. Crop Science. **39**: 1571-1583.

Templeton, A. R. 1986. Coadaptation and outbreeding depression. In Conservation biology. The science of scarcity and diversity (ed. M.E. Soule), pp. 105-116. Sunderland, MA: Sinauer Associates.

Vaillancourt, R. E., B. M. Potts, M. Watson, P. W. Volker, G. R. Hodge, J. B. Reid, and A. K. West. 1995. Detection and prediction of heterosis in *eucalyptus globulus*. Forest Genetics **2**:11-19.

Walter, C., S. D. Carson, M. I. Menzies, T. Richardson and M. Carson. 1998. Review: Application of biotechnology to forestry - molecular biology of conifers. World Journal of Microbiology and Biotechnology **14**:321 - 330.

Waser, N. M., and M. V. Price. 1989. Optimal outcrossing *Ipomopsis aggregata*: seed set and offspring fitness. Evolution **43**:1097-1109.

Webber, J. E. 2000. Western hemlock: a manual for tree improvment seed production. Working paper 44, B.C. Ministry of Forests, Victoria B.C.

Weir, B. S., and C. C. Cockerham. 1973. Mixed self and random mating at two loci. Genetical Research **21**:247-262.

Wellman, H., C. Ritland, and K. Ritland. 2003. Genetic effects of domestication in western hemlock [Tsuga heterophylla (Raf.) Sarg]. Forest Genetics 10 **3**:229-240

Williamson, R. L. 1976. Natural regeneration of western hemlock. Pages 166-169 *in* I. o. F. Products, editor. Western Hemlock Management. University of Washington, Seattle Wash.

Woods, J. 1984. Coastal western hemlock breeding program review and recommendation. Internal report B.C. Ministry of Forests, Victoria B.C.

Wooldridge, D. D. 1961. Factors related to growth and management of western hemlock Tsuga heterophylla (Raf.) Sarg. University of Washington, Seattle, Washington.

Wright, S. 1921. Systems of mating II. The effects of inbreeding on the genetic composition of a population. Genetics **6**: 124 - 143.

Zane, L., L. Bargelloni and T. Patarnello. 2002. Strategies for microsateellite isolation: a review. Molecular Ecology **11**:1 - 16.

Zhang, Q., Y. J. Gao, S. H. Yang, R. A. Ragab, M. A. Saghai-Maroof, and Z. B. Li. 1994. A diallel analysis of heterosis in elite hybrid rice based on RFLPs and microsatellites. Theoretical Applied Genetics **89**:185 - 192.

Zouros, E. 1987. On the relatioship between heterozygosity and heterosis: an evaluation of the evidence from marine mollusks. Isozymes **15**:255 - 270.

1.7 Tables and Figures

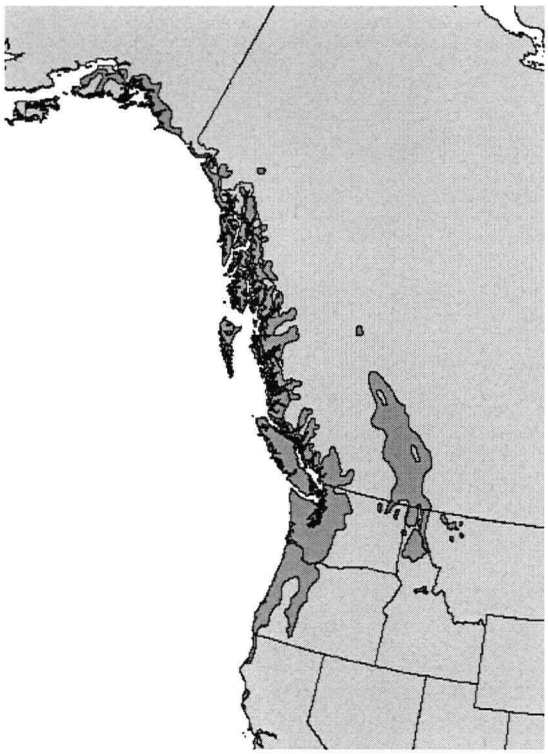


Figure 1.1. Range of Western hemlock *Tsuga heterophylla* (Raf.) Sarg..

CHAPTER 2: Genetic Effects of Domestication in Western hemlock, *Tsuga heterophylla*

2.1 Introduction

Western hemlock (*Tsuga heterophylla* (Raf.) Sarg) is distributed mainly in a narrow zone along the Pacific Coast from the Kenai Peninsula in Alaska to Northern California (Figure 2.1). Stands of western hemlock are among the most productive in the world (Packee 1990), and rank third in British Columbia for annual volume cut (Owens and Molder 1984). This conifer has good to excellent pulping characteristics, and its fiber is a major source for groundwood, thermomechanical, kraft, and sulfite pulps (Packee 1990). In addition, its fine grain and resin free characteristics make it a suitable finishing lumber. This species also plays a major ecological role, being an important browse species for deer and elk as well as making up a significant portion of forest canopies in both western Canada and the northwestern United States.

In British Columbia, seed orchards are used to provide seedlings for reforestation. The growth cycle of vegetative buds and reproductive cones is well understood and this information has allowed the efficient production of nursery grown seedlings. Advancement in breeding practices has led to genetic improvement projects. Trees are selected for superior fitness and are incorporated into breeding plans. Such selection can lead to loss of genetic variation and an elevated percentage of inbreeding. Without genetic variance there can be no adaptive response by the individual (Hedrick 1985). As a consequence, the higher inbreeding may lead to depression in fitness. Therefore it is important to understand the effect that phenotypic selection will have on genetic diversity in breeding programs.

In outcrossed species such as western hemlock, erosion of population genetic diversity can lead to a loss in viability and reproductive success (Mosseler et al. 1994). This can therefore hinder the ability of the species to respond to selection pressures. As a consequence changes in our environment can conclude in selection for this reduced fitness. This can have profound consequences on ecosystem stability. Therefore,

sustaining genetic diversity is an important factor in the maintenance of stability in managed areas. With the use of seed orchard based seed to restock natural systems, genetic consequences must be realized.

When a sub-sample of a population is used to represent a large area of a species range, genetic diversity can be compromised. With this, inbreeding and inbreeding depression arise as an issue. Inbreeding occurs when mates are more closely related than they would be if they had been chosen at random from the population (Crow and Kimura 1970). Mating of closely related individuals may increase the probability of homozygosity in recessive lethal alleles. The opportunity of mating with relatives increases with decreasing population size, therefore some concern must be taken when dealing with a decreaseing gene pool (Muller-Stark 1995). To date little molecular work has looked at the above issues in western hemlock. The maintenance of genetic diversity in breeding and production populations of commercially valuable species is a priority in all breeding programs, but has seldom been verified by direct genetic assays of levels of variation. However, the problems with identifying rare alleles and the general limitations of isozymes for characterizing genetic variation will limit our inferences.

The goal of this study is to interpret levels of genetic differentiation under different seed orchard treatments when compared with natural populations. The orchard treatments tested in this study include full-sib, supplemental mass pollination and unimproved (open-pollinated). Our purpose is then to (1) identify rates of genetic diversity under differing orchard conditions compared to natural stands of western hemlock and (2) identify the effectiveness of seed orchard and seed collection protocol in maintaining natural genetic structure.

2.2 Materials and Methods

The geographic origin of the natural and seed orchard source populations are given in Table 2.1 and Figure 2.1. Approximately 50 individuals were sampled from each of the 22 populations. The natural populations were collected throughout the 70's and 80's by the B.C. Ministry of Forests. Parents used in orchards were selected from

wild collections carried out by the B.C. Ministry of Forests. All seed orchard material used in this study was collected from F1 offspring. The unimproved seed orchards under went open-pollination. The full-sib orchard, collections are controlled crosses. The pollen buds were removed from part of a branch, and female strobili were isolated with a paper/plastic bag; next pollen from only one other parent was applied. This eliminates the possibility of contamination. In general 20 isolation bags are used on one ramet of one clone and each approximately cover 20 plus cones. To meet B.C.'s diversity requirements for public lands a minimum of 5 crosses are used for each full-sib pollination treatment. The SMP (Supplementary Mass Pollinations) orchard, collections had pollen from best trees applied to receptive seed clones. Stored best pollen can be applied with contamination from outside sources minimized. This is done by hand or with a blower. Breeding value for SMP treatments are based on applied pollen and concluding seed production given estimated contamination.

Seed was removed from the cone and was stored at 4° C until germination. The germination process included a 72 hour soaking in distilled water at 4° C followed by a 3 week period in which the seeds were left on wet filter paper at room temperature. Seedlings were removed once they had reached 2 cm in length and placed in cold storage (4° C).

Starch gel electrophoresis was chosen as the molecular genetic technique for this study. Fourteen loci were identified using nine isozyme stains. Two electrophoresis systems were used to optimize the resolution of each locus. These included one continuous system, Morpholine-citrate pH 8.0 (Clayton and Tretiak 1972) and one discontinuous system, Sodium-borate pH 8.0 (Poulik 1957). The stains used with the morpholine system included: Shikimate dehydrogenase (SKDH), 6-Phosphogluconate dehydrogenase (PGD), Malate dehydrogenase (MDH), Fructose-1,6-diphosphatase (FDP) and Isocitric dehydrogenase (IDH). The stains used with the Sodium-borate system included: Aspartate aminotransferase (AAT), Glutamate dehydrogenase (GDH), Phosphoglucoisomerase (PGI) and Phosphoglucomutase (PGM).

The gel was produced with a 12% starch (Starch-art corp.) and 5% sucrose (Sigma) concentration. This volume allowed 5 useable slices per gel for the same

isozyme extract. Different populations were placed on the same gel to allow comparison between alleles of different populations. The running time for the Poulik running systems was 4 hours at 230V and the running time for the Morpholine system was 4 hours at 160V. The histochemical staining solutions were obtained from Murphy et al. (1996) and Acquaah (1992). The banding pattern was recorded visually and each allele was denoted by an integer depending upon their mobility. Alleles were defined consecutively with 1 representing the fastest band. One tetramer, (MDH) was defined and it was scored as a monomer for two positions within the banding pattern, hence two loci for this stain.

Observed and expected mean heterozygosity, percentage of polymorphic loci, and mean number of alleles per locus were calculated from allelic frequencies using the BIOSYS-2 computer program (Swofford and Selander 1981, Black IV 1997). The diversity index (H_t) was calculated using POPGENE, version 1.32 (Yeh et al. 1997). The dendogram was produced using the computer program "gd" (http://genetics.forestry.ubc.ca/ritland/programs.html) whose algorithm is described in Ritland (1989).

2.3 Results

Mean heterozygosity ranged from 0.112 to 0.160 for the 11 natural populations collected throughout British Columbia, while the seed orchards ranged from 0.120 to 0.167 (Table 2.2). Mean number of alleles per locus ranged from 1.4 for the Camper population to 2.0 for the Holberg population. The percentage of polymorphic loci ranged from 28.6 for populations, Fleet, U.B.C., Camper and S.O 143UN to 50.0 for populations Holerg, Toba and Ucona. Mean heterozygosity, mean number of alleles, H_t and the percentage of polymorphic loci have also been expressed as an average for each of the grouped seed orchard samples, SMP, FS, UN as well as the natural population (Table 2.3).

Both the unimproved and the SMP seed orchard collections show higher mean number of alleles per locus compared with the natural populations, but the full-sib orchard collections resulted in lower estimates. This can be contrasted to the percentage

of polymorphic loci which showed a higher value in the natural populations. The SMP population showed higher heterozygosity (0.154) overall compared with all other groups. The lowest average heterozygosity was identified in the natural populations. The measure of gene diversity (H_t) portrayed a similar story with the SMP seed orchard showing the highest genetic diversity (0.164). But in contrast to the basic heterozygosity measures, both the FS and unimproved orchards fell lower than the natural population shown in Table 2.3.

A dendogram mapping thet collective allelic differences between each population is presented as Figure 2.2. Three distinct groups can be inferred. These are: (1) Sombrio, Holberg, Fleet, UBC and S.O. 133UN, (2) Toba, S.O. 133smp and S.O. 133fs and (3) the remaining populations. It can be noted that the standard error of branch length, depicted as the thicker shaded branches, is quite large, making most inferences fairly non-significant. The lack of significant structure is not surprising given the low value of maximal genetic divergence (0.008).

Allele frequencies are presented in Appendix A. Some alleles are present in the natural populations that are not present in the seed orchards. These include population 1, PGD-1 A, PGD-2 A, population 2, MDH C, IDH-1 D, population 3, IDH-1 C, population 8, IDH-1 B and population 9, PGM-1 D and PGM-2 D. Some alleles are also present in the seed orchards that are not present in the natural populations. These include population 12, PGM-1 C, population 14, SKDH-2 C, population 17 and 18, PGM-2 B and population 21, PGI C. All of the rare alleles presented have frequencies between .015 and .025.

2.4 Discussion

Our finding of large amounts of genetic variation in western hemlock, and little genetic differentiation of populations, accords with the life history and patterns of morphological and physiological variability exhibited by this species. Western hemlock has a large, continuous, geographic distribution, and is wind pollinated; therefore it should show little population differentiation (Mosseler et al. 1994), with perhaps less than

10% of the variation occurring among population, as gauged by studies with similar species (Hamrick and Godt 1990). Malavasi and Perry (1993) found, in a shade tolerance study of Western hemlock, physiological variability within populations. A common garden study of 21 western hemlock provenances from California to Alaska showed latitudinal differentiation for cold hardiness, survival and seedling growth factors (Kuser and Ching 1980, 1981). However, in the range of 46° to 51° no trend was found. Foster and Lester (1983) found similar results for height, with no differentiation in the 3° latitude distribution of hemlock in Washington State. Interestingly, King (1991) did find significant differentiation between trees below vs. above 600m elevation.

2.4.1 Seed Orchards vs. Natural Populations

Tsuga heterophylla, an outcrossing wind pollinated species, showed a relatively high gene diversity (H_t) in its natural populations, 0.146 (Table 2.3). This finding is similar to other outcrossing wind pollinated species (Hamrick et al., 1992). *T. heterophylla* had a lower H_t value than that determined as an average value in gymnosperms (0.281) but was similar to an average value determined for long-lived perennials, 0.148 (Hamrick et al., 1992). The unimproved and full-sib seed orchard collections showed a slight decrease in H_t (0.142) but this was not significant (Table 2.3). This trend was also seen in mean observed heterozygosity (H_o). The natural populations H_o (0.136) did not differ significantly from the unimproved and full-sib seed orchards, 0.136 and 0.137 respectively (Table 2.3). It is important to note that the observed heterozygosities are not significantly different than the expected heterozygosities (H_e), for the natural populations indicating that they are in Hardy-Weinberg equilibrium. These results indicate that gene diversity was retained in the unimproved and full-sib seed orchards. This result was also seen in the percentage of polymorphic loci which showed a slight decrease for orchard populations compared with natural populations (Table 2.3).

In contrast the SMP seed orchards showed different results. The SMP orchards had a higher H_t (0.164) than the unimproved orchards (0.142) (Table 2.3). This trend was also seen in H_o (0.154) for the SMP compared to the unimproved orchards (0.136) (Table

2.3). These results indicate that gene diversity may be higher for this treatment. This is expected as SMP has been shown to: 1) minimize self-fertilization (El-Kassaby and Ritland, 1986) 2) improve reproductive syncrocy (Reynolds and El-Kassaby, 1990) and 3) improve parental balance (Reynolds and El-Kassaby, 1990). All of these factors would result in higher levels of genetic diversity. The percentage of polymorphic loci worked in parallel with the H_t results. The SMP seed orchards were measured at P=39.3% where as the unimproved seed orchards were measured at P=36.9% (Table 2.3). Although slight, there is a decrease in the polymorphic estimate when comparing the SMP treatment to the natural populations (40.3%) (Table 2.3). This is most likely the result of a limited genetic base in the orchard population. It should also be noted that the mean number of alleles did not differ significantly between the natural and seed orchard populations.

The dendogram showed limited pattern (Figure 2.2), but does indicate three groups. These groups show little correlation between the seed orchards and the source areas of the orchards that the natural populations are located in. This result suggests that gene flow in the orchards has sufficiently mixed the existing gene pool creating more diverse populations. Genetic drift within the orchards has not seemed to homogenize the populations therefore leading to less genetic diversity. This was shown earlier in the mean heterozygosity (H_o) and genetic diversity (H_t) levels. One issue that may arise from high rates of gene flow is the loss of rare alleles as rare alleles may contribute to genetic differentiation that allow a population to adapt to a changing ecosystem.

The allele frequencies are listed for all loci in the 22 populations (Appendix 2.A). Eight alleles were recorded in the natural populations that were not present in any of the seed orchard populations. All alleles were very rare being found in only one of the natural populations studied. Within the populations their allele frequencies ranged from 0.020 to 0.025 (Appendix 2.A). In comparison, 4 alleles were present in the seed orchards that were not present in the natural populations. These alleles were also only found in one seed orchard for each. Their frequencies were also very low, ranging from 0.010 to 0.020. Chaisurisri and El-Kassaby (1994) noted in a similar study on Sitka spruce (*Picea sitchensis* (Bong.) Carr.) that sampling breadth may be the reason for such findings. The source ranges differed between the seed orchards and natural populations.

All of the orchard parents were collected throughout the range of western hemlock, where as the natural populations were centered around Vancouver Island and the Lower Mainland.

2.4.2 Comparison to Other Hemlock Species

Very little molecular work has been performed on hemlock. Zabinski (1992) looked at isozyme variation in eastern hemlock (*Tsuga canadensis* (L.) Carr.) and found very low rates of H_t, (0.04). The proportion of polymorphic loci was also very low for this species, 0.10. This is unexpected as much higher values have been observed for outcrossing, wind pollinated, long lived conifers with a very wide range. The hypothesis presented as an explanation are that of Fowler and Morris (1977). They point to a population bottleneck or a series of bottlenecks during the Pleistocene.

Ally et al. (2000) also found lower than expected levels of H_t in Mountain hemlock (*Tsuga mertensiana*). They found an H_t of 0.093 and a percentage of polymorphic alleles of 0.33. They point to two reasons for this relatively low estimate. First being the bottleneck hypothesis, presented earlier, and second, the genetic depauperacy of southern refugial populations hypothesized by Cwnyar and MacDonald (1987). They also note that flight of mountain hemlock seed is short and much more likely to promote family structure and the accumulation of local genetic differences. This is in contrast to western hemlock, which has been recorded, under windy conditions, with 1.6 km of seed travel (Isaac 1930).

2.4.3 Comparison to Similar Studies with Other Conifers

Chaisurisri and El-Kassaby (1994) performed a similar study with Sitka spruce (*Picea sitchensis* (Bong.) Carr.). In this study 10 range-wide natural populations were compared to 1 seed orchard population. Major finding include a non-significant higher value for mean heterozygosity in the seed orchard population when compared to the natural populations. It is also noted, by looking at genetic distance, that the seed orchard

is similar to three of the natural populations located within the area where the parent seed orchard seed was collected. Increased levels of heterozygosity in the seed orchard are hypothesized as being the result of sampling breadth in parent tree sampling (Chaisurisri and El-Kassaby 1994). As a conclusion Chaisurisri and El-Kassaby (1994) note that the seed orchard, composed of a production population of 139 clones, was sufficient to prevent loss of genetic variability.

El-Kassaby and Ritland (1996) performed a seed orchard study to identify the impact of selection and breeding on genetic diversity in Douglas-fir (*Pseudotsuga menziesii*). In this study two generations of seed orchards were compared against their 49 wild progenitor populations. Measures of heterozygosity, polymorphic loci and divergence were found to be similar or higher in the domesticated populations. El-Kassaby and Ritland (1996) therefore concluded that selection and breeding had not lead to a significant decrease in genetic variation. This was thought to be linked to the parents used to stock the orchards as it was pooled from a widely distributed natural population. They also noted that although the first generation seed orchard was not significantly different from the natural populations, the second generation was. This is thought to be the result of the reduced set up parents that formed the advanced seed orchard generation.

The dynamics of rare alleles in these domesticated populations needs study, using different types of genetic markers. Isozymes are well known to encompass a very restricted portion of the total genome of an organism. In a Douglas-fir study isozymes were shown to not fully measure the losses of variation that occur in the initial stages of domestication (El-Kassaby and Ritland 1996). The ideal markers would lie adjacent to loci controlling physiologically and ecologically important characters, and lacking knowledge of the locations of these genes, would at least be numerous and highly polymorphic. Further analysis should therefore be initiated.

2.5 References

Acquaah, G. 1992. *Practical Protein Electrophoresis for Genetic Research.* Theodore R. Dudley, general editor. Dioscorides Press. Portland Oregon.

Ally, D., Y.A. El-Kassaby and K. Ritland. 2000. Genetic diversity, differentiation and mating system in mountain hemlock (*Tsuga mertensiana*) across British Columbia. Forest Genetics 7, **2**: 97 – 108.

Chaisurisri, K. and Y.A. El-Kassaby. 1994. Genetic diversity in seed production population vs. natural populations of Sitka Spruce. Biodiversity and Conservation **3**: 512 – 523.

Clayton, J.W. and D.N. Tretiak. 1972. Amine-citrate buffers for pH control in starch gel electrophoresis. Journal of Fish Research. Board Canada. **29**: 1169 – 1172.

Crow, J.K. and M. Kimura. 1970. An introduction to population genetics theory. Harper and Row, New York.

Cwynar, L.C. and G.M. MacDonald. 1987. Geographical variation of lodgepole pine in relation to population history. American Naturalist **129**: 463 – 469.

El-Kassaby, Y.A. and K. Ritland. 1986. Low level of pollen contamination in a douglas-fir seed orchard as detected by allozyme markers. Silvae Genetica **35**: 224 – 229.

El-Kassaby, Y.A. and K. Ritland. 1996. Impact of selection and breeding on the genetic diversity in Douglas-fir. Biodiversity and Conservation **5**: 795 – 813.

Foster, G.S. and D.T. Lester. 1983. Fifth-year height variation in western hemlock open-pollinated families growing on four test sites. Canadian Journal Forest Research **13**: 251 – 256.

Fowler, D.P. and R.W. Morris. 1977. Genetic diversity in red pine: evidence for low heterozygosity. Canadian Journal of Forest Research **7**: 341 – 347.

Hamrick, J.L. and M.J.W. Godt. 1990. 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, Sunderland, Mass. 43 – 63.

Hamrick, J.L., M.J.W. Godt and S.L. Sherman-Broyles. 1992. Factors influencing levels of genetic diversity in woody plant species. New Forest **6**: 95 – 124.

Hedrick, P.W. 1985. Genetics of populations. Jones and Bartlett Publishers, Inc., Boston, MA.

Isaac, L.A. 1930. Seed flight in the Douglas-fir Region. Journal of Forestry **28**: 492 – 499.

King J.N. 1991. The significance of geographic variation patterns for western hemlock genetic improvement. Technical Report, B.C. Ministry of Forests Research Branch.

Kuser, J.E. and K.K. Ching. 1980. Provenance variation in phenology and cold hardiness of western hemlock seedlings. Forest Science **26**: 463 – 470.

Kuser, J.E. and K.K. Ching. 1981. Provenance variation in seed weight, cotyledon number, and growth rate of western hemlock seedlings. Forest Science **26**: 463 – 470.

Malavasi, U.C. and D.A. Perry. 1993. Genetic variation in competitive ability of some shade-tolerant and shade-intolerant pacific Coast (USA) conifers. Forest Ecology and Management **56**: 69 – 81.

Mosseler, A., K.N. Egger and S.M. Carr. 1994. Molecular biology and genetic diversity in tree populations. *In*. Recent progress in Forest Biotechnology in Canada. Ed. Charest and Duchense. 114 – 124.

Muller-Stark, G. 1995. Protection of genetic variability in forest trees. Forest Genetics. **2**: 121-124.

Murphy, R.W., J.W. Sites Jr., D.G. Buth, and C.H. Haufler. 1996. Proteins: Isozyme Electrophoresis. In: *Molecular Systematics*. 2nd Ed. Hillis, David M., Craig Moritz, Barbara K. Mable (eds.). Sinauer Associates, Inc. Sunderland, Massachusetts. 51 – 120.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89**: 583 – 590.

Owens, J.N. and M. Molder. 1984. The reproductive cycles of western and mountain hemlock. Biology Dept, UVIC, Victoria B.C.. Ministry of Forests. Pp 30.

Packee, E.C. 1990. *Tsuga heterophylla* (Raf.) Sarg: Western hemlock. In *Silvics of North America: Volume 1, conifers.* Burns, R. M.; Honkala, B. H. (tech. coords). Agriculture Handbook 654. Washington, DC: U.S. Department of Agriculture, Forest Service: 260 – 267.

Poulik, M.D. 1957. Starch gel electrophoresis in a discontinuous system buffers. Nature 180: 1477 – 1479.

Reynolds, S. and Y.A. El-Kassaby. 1990. Parental balance in douglas-fir seed orchards – cone crop vs. seed crop. Silvae Genetica **39**: 40 – 42.

Ritland, K. 1989. Genetic differentiation, diversity and inbreeding in the mountain monkeyflower (*Mimulus caespitosus*) of the Washington Cascades. Canadian Journal of Botany **67**: 2017-2024.

Swofford, D.L. and R.B. Selander. 1981. BIOSYS-1:A FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. Journal of Heredity **72**: 281 –283.

Yeh, F., Yang, C., Boyle, R.C., Timothy, B.J., Ye, Z.H., and J.X. Mao. 1997. POPGENE, the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada.

Zabinski, C. 1992. Isozyme variation in eastern hemlock. Canadian Journal of Forest Research. **22**: 1838 – 1842.

2.6 Tables and Figures

Table 2.1. Population codes, year of collection and geographic origin of *T. heterophylla* seed.

#	Seedlot	Location	Year	Latitude	Longitude	Elevation (m)
1	2685	Holberg	1975	50 38	128 05	92
2	2753	Toba River	1975	50 30	124 10	366
3	3471	Ucona River	1978	49 40	126 00	475
4	4088	Fleet River	1979	48 38	124 04	370
5	4538	Sombrio Creek	1978	48 32	124 18	365
6	4692	Camper Creek	1982	48 34	124 30	300
7	4787	Sechelt	1976			
8	7832	Nanaimo River	1987	49 00	124 10	550
9	9789	UBC Res. Forest	1985	49 17	122 33	275
10	18784	Kaouk Jnc.	1982	50 05	126 59	60
11	46152	Unknown natural		48 50	123 45	600
12	60160	S.O.# 126 - unimproved	1993	50 27	127 09	188
13	60352	S.O.# 126 - SMP	1998	50 30	126 53	95
14	60351	S.O.# 126 - full-sibs	1998	50 02	125 52	121
15	6883	S.O.# 133 - unimproved	1990	50 00	124 30	140
16	61060	S.O.# 133 - SMP	1999	50 12	125 08	300
17	60624	S.O.# 133 - full-sib	1997	50 00	125 00	300
18	60183	S.O.# 136 - unimproved	1993	49 38	126 08	169
19	60224	S.O.# 156 - unimproved	1995	52 43	131 40	241
20	60319	S.O.# 130 - unimproved	1995	48 47	124 18	595
21	60374	S.O.# 143 - unimproved	1998	49 54	124 29	745
22	60067	S.O.# 127 - unimproved		50 37	127 19	616

Table 2.2. Mean heterozygosity, (both expected Hardy-Weinberg (H_E) and Direct count (H_o)), sample size (SS), mean number of alleles per locus (NA) and percentage of loci that are polymorphic (%P) for 11 natural populations, 11 seed orchards of *T. heterophylla*.

						Mean H	
#	Population	code	SS	NA	%P*	H_o	H_E**
1	HOLERG	2685	49	2 (0.1)	50.0	0.143 (0.071)	0.171 (0.051)
2	TOBA	2753	49	1.9 (0.2)	50.0	0.149 (0.072)	0.146 (0.045)
3	UCONA	3471	49	1.9 (0.2)	50.0	0.131 (0.069)	0.145 (0.045)
4	FLEET	4088	50	1.6 (0.1)	28.6	0.117 (0.071)	0.132 (0.054)
5	SOMBRIO	4538	50	1.6 (0.2)	42.9	0.143 (0.075)	0.147 (0.052)
6	CAMPER	4692	50	1.4 (0.1)	28.6	0.119 (0.075)	0.114 (0.051)
7	SECHELT	4787	52	1.6 (0.2)	42.9	0.16 (0.076)	0.148 (0.049)
8	NANAIMO	7832	50	1.9 (0.2)	35.7	0.14 (0.072)	0.131 (0.046)
9	UBC	9789	49	1.7 (0.2)	28.6	0.112 (0.073)	0.137 (0.054)
0	KAOUK	18784	50	1.6 (0.2)	42.9	0.149 (0.075)	0.147 (0.053)
11	UNKNOWN	46152	49	1.5 (0.1)	42.9	0.161 (0.079)	0.131 (0.048)
12	SO126UN	60160	50	1.8 (0.2)	42.9	0.146 (0.074)	0.154 (0.051)
13	SO126SMP	60352	49	1.8 (0.2)	42.9	0.167 (0.076)	0.158 (0.053)
14	SO126FS	60351	50	1.6 (0.2)	35.7	0.154 (0.079)	0.149 (0.057)
15	SO133UN	6883	50	1.9 (0.2)	35.7	0.143 (0.076)	0.135 (0.047)
16	SO133SMP	60106	50	1.8 (0.2)	35.7	0.141 (0.074)	0.158 (0.056)
17	SO133FS	60624	50	1.6 (0.2)	42.9	0.12 (0.071)	0.123 (0.042)
18	SO136UN	60183	50	1.8 (0.2)	42.9	0.145 (0.072)	0.147 (0.048)
19	SO156UN	60224	50	1.9 (0.1)	35.7	0.126 (0.07)	0.14 (0.049)
20	SO130UN	60319	52	1.7 (0.2)	35.7	0.136 (0.072)	0.145 (0.053)
21	SO143UN	60374	50	1.7 (0.2)	28.6	0.12 (0.073)	0.121 (0.049)
22	UNKNOWN	60067	48	1.8 (0.2)	35.7	0.142 (0.074)	0.15 (0.051)

** A locus is considered polymorphic if the frequency of the most common allele does not exceed 0.95*

*** Unbiased estimate for Hardy Weinberg expectation (Hdy Wby) (see Nei, 1978)*

Table 2.3. Average mean heterozygosity (both expected Hardy-Weinberg (H_E) and Direct count (H_o)), average mean sample size per locus (SS), average mean number of alleles per locus (NA), average percentage of loci that are polymorphic (%P) and the genetic diversity index, Ht for 10 natural populations, 6 unimproved (UN) seed orchards,2 supplimental mass pollinated (SMP) seed orchards and 2 full-sib (FS) orchards of *T. heterophylla.*

				Mean H		
Population	SS	NA	%P*	H_o	H_E**	Ht
TOTAL MEAN	49.1	1.7	38.9	0.139	0.142	
NATURAL MEAN		1.7	40.3	0.136	0.142	0.146
S.0. UN MEAN		1.8	36.9	0.136	0.14	0.142
S.O. SMP MEAN		1.8	39.3	0.154	0.158	0.164
S.O. FS MEAN		1.6	39.3	0.137	0.136	0.141

**A locus is considered polymorphic if the frequency of the most common allele does not exceed 0.95*

***Unbiased estimate estimate for Hardy Weinberg expectation (Hdy Wby) (see Nei, 1978)*

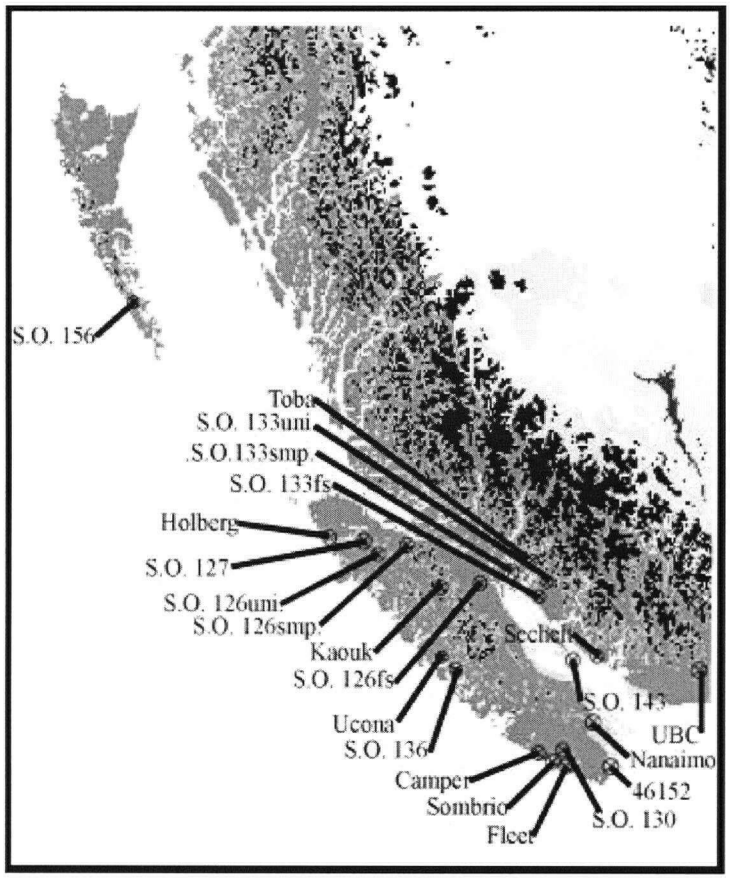


Figure 2.1. Locations of *Tsuga heterophylla* source populations in British Columbia, Canada.

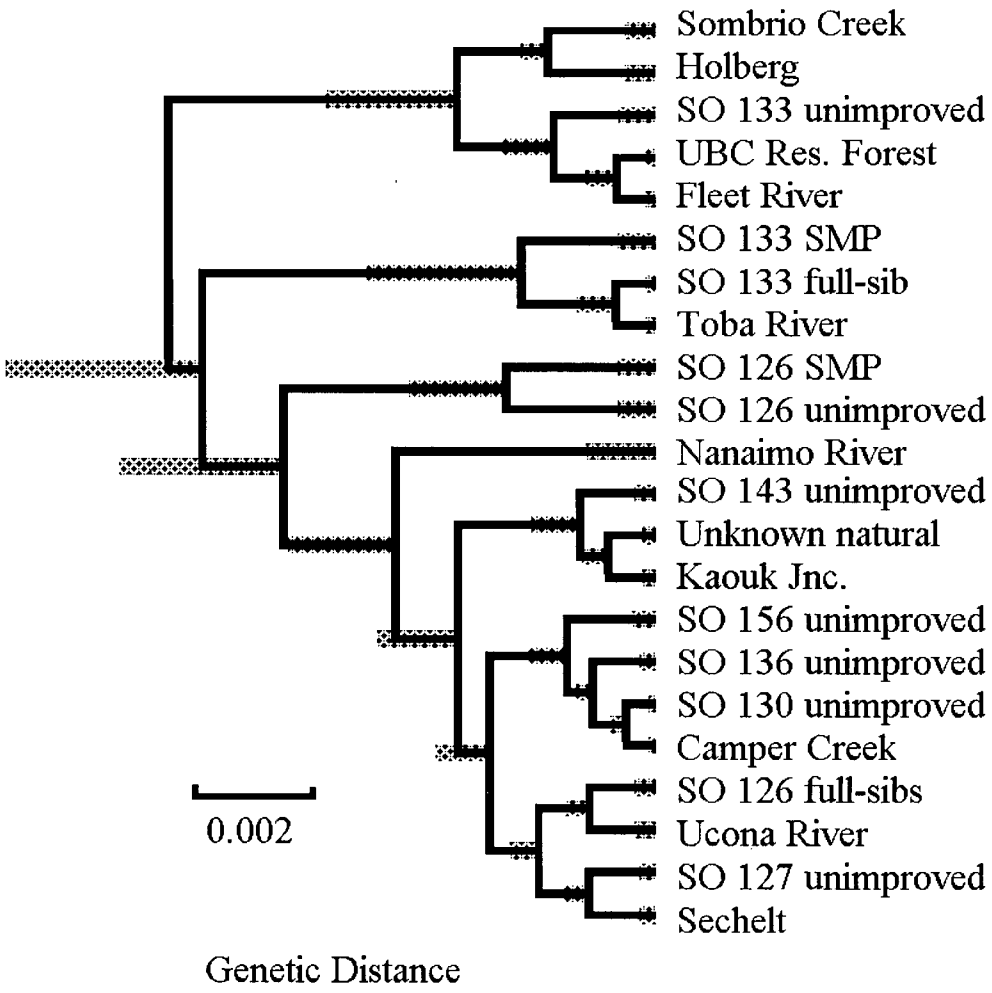


Figure 2.2. Dendogram of 22 seed orchard and natural *T. heterophylla* populations.

Appendix 2.A. Allozyme frequencies for the 22 seed orchard and natural populations of Western hemlock collected in British Columbia.

		Population										
Locus	Allele	1	2	3	4	5	6	7	8	9	10	11
Aat-1	(N)	50	50	50	50	50	50	50	50	50	50	50
	A	0.99	0.99	0.98	0.98	1	1	0.981	0.94	1	0.98	0.98
	B	0.01	0.01	0.02	0.02	0	0	0.019	0.06	0	0.02	0.02
Aat-2	(N)	50	50	50	50	50	50	53	50	50	50	50
	A	0.9	0.93	0.91	0.97	0.87	0.98	0.837	0.98	0.96	0.97	0.98
	B	0.1	0.07	0.09	0.03	0.13	0.02	0.163	0.02	0.04	0.03	0.02
Gdh	(N)	40	40	49	50	50	50	52	50	48	50	50
	A	0.6	0.8	0.837	0.6	0.7	0.8	0.808	0.88	0.583	0.82	0.86
	B	0.4	0.2	0.163	0.4	0.3	0.2	0.192	0.12	0.417	0.18	0.14
Pgi	(N)	50	50	50	50	50	50	50	52	50	50	50
	A	0.01	0.01	0.03	0.01	0	0	0	0.03	0	0	0.01
	B	0.99	0.99	0.97	0.99	1	1	1	0.97	1	1	0.99
	C	0	0	0	0	0	0	0	0	0	0	0
Pgm-1	(N)	45	45	45	45	50	50	52	50	40	30	38
	A	0.956	0.944	0.978	1	0.98	1	1	1	0.975	0.95	0.934
	B	0.044	0.056	0.022	0	0.02	0	0	0	0	0.05	0.066
	C	0	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	0	0	0	0	0	0.025	0	0
Pgm-2	(N)	45	45	45	45	50	50	52	50	40	30	38
	A	1	1	1	1	1	1	1	1	0.975	1	1
	B	0	0	0	0	0	0	0	0	0	0	0
	C	0	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	0	0	0	0	0	0.025	0	0
Skdh-1	(N)	50	49	49	50	50	50	52	50	50	50 1	50
	A	0.07	0.173	0.102	0	0.02	0	0.019	0.02	0.05	0.13	0.18
	B	0.88	0.786	0.878	0.9	0.9	0.9	0.846	0.86	0.91	0.72	0.74
	C	0.05	0.041	0.02	0.1	0.08	0.1	0.135	0.12	0.04	0.15	0.08
Skdh-2	(N)	50	49	49	50	45	35	37	50	49	40 2	40
	A	0.49	0.214	0.378	0.4	0.556	0.443	0.351	0.58	0.439	0.463	0.475
	B	0.51	0.786	0.622	0.6	0.444	0.557	0.649	0.42	0.561	0.538	0.525
	C	0	0	0	0	0	0	0	0	0	0	0
Pgd-1	(N)	50	50	50	50	50	50	52	50	50	50	50
	A	0.02	0	0	0	0	0	0	0	0	0	0
	B	0.98	1	1	1	1	1	1	1	1	1	1
Pgd-2	(N)	50	50	50	50	50	50	52	50	50	50	50
	A	0.02	0	0	0	0	0	0	0	0	0	0
	B	0.98	1	1	1	1	1	1	1	1	1	1

Appendix 2.A. Allozyme frequencies for the 22 seed orchard and natural populations of Western hemlock collected in British Columbia.

		Population										
Locus	Allele	12	13	14	15	16	17	18	19	20	21	22
Aat-1	(N)	50	50	50	50	50	50	51	52	50	50	50
	A	0.97	1	0.99	0.99	1	0.98	0.931	0.981	0.98	1	1
	B	0.03	0	0.01	0.01	0	0.02	0.069	0.019	0.02	0	0
Aat-2	(N)	50	50	50	50	50	50	51	52	50	50	50
	A	0.94	0.96	0.96	0.94	0.91	0.92	0.961	0.962	0.96	0.97	0.88
	B	0.06	0.04	0.04	0.06	0.09	0.08	0.039	0.038	0.04	0.03	0.12
Gdh	(N)	50	50	50	50	40	50	51	52	50	50	50
	A	0.88	0.8	0.82	0.54	0.825	0.76	0.784	0.75	0.82	0.84	0.88
	B	0.12	0.2	0.18	0.46	0.175	0.24	0.216	0.25	0.18	0.16	0.12
Pgi	(N)	50	50	50	50	50	50	51	52	50	50	50
	A	0.01	0	0.02	0.01	0	0.02	0.01	0.01	0.01	0.02	0
	B	0.99	1	0.98	0.99	1	0.98	0.99	0.99	0.99	0.96	1
	C	0	0	0	0	0	0	0	0	0	0.02	0
Pgm-1	(N)	50	50	50	50	50	50	51	52	44	40	50
	A	0.97	0.99	0.99	0.98	1	1	0.961	0.981	1	0.988	1
	B	0.01	0.01	0.01	0.02	0	0	0.039	0.019	0	0.013	0
	C	0.02	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	0	0	0	0	0	0	0	0
Pgm-2	(N)	50	50	50	50	50	50	51	52	44	40	50
	A	1	1	1	1	1	0.99	0.98	1	1	1	1
	B	0	0	0	0	0	0.01	0.02	0	0	0	0
	C	0	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	0	0	0	0	0	0	0	0
Skdh-1	(N)	50	50	50	50	50	49	50	52	50	50	50
	A	0.04	0.06	0.02	0.07	0.02	0.143	0.1	0.096	0.03	0.18	0
	B	0.69	0.63	0.87	0.77	0.87	0.776	0.88	0.788	0.92	0.74	0.91
	C	0.27	0.31	0.11	0.16	0.11	0.082	0.02	0.115	0.05	0.08	0.09
Skdh-2	(N)	50	50	33	50	50	29	30	30	50	39	50
	A	0.56	0.49	0.333	0.41	0.18	0.241	0.45	0.467	0.48	0.385	0.4
	B	0.44	0.51	0.652	0.59	0.82	0.759	0.55	0.533	0.52	0.615	0.6
	C	0	0	0.015	0	0	0	0	0	0	0	0
Pgd-1	(N)	50	50	50	50	50	49	50	52	50	50	50
	A	0	0	0	0	0	0	0	0	0	0	0
	B	1	1	1	1	1	1	1	1	1	1	1
Pgd-2	(N)	50	50	50	50	50	49	50	52	50	50	50
	A	0	0	0	0	0	0	0	0	0	0	0
	B	1	1	1	1	1	1	1	1	1	1	1

Appendix 2.A. Allozyme frequencies for the 22 seed orchard and natural populations of Western hemlock collected in British Columbia.

		Population										
Locus	Allele	1	2	3	4	5	6	7	8	9	10	11
Mdh	(N)	50	50	50	50	50	50	52	50	50	50	50
	A	0.9	0.91	0.95	1	0.95	0.98	0.904	0.96	0.98	0.94	0.9
	B	0.1	0.07	0.05	0	0.05	0.02	0.096	0.04	0.02	0.06	0.1
	C	0	0.02	0	0	0	0	0	0	0	0	0
Fdp	(N)	50	50	50	50	50	50	52	50	50	50	50
	A	0.5	0.5	0.49	0.48	0.5	0.5	0.5	0.5	0.5	0.5	0.5
	B	0.5	0.5	0.51	0.52	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Idh-1	(N)	50	50	50	50	50	50	53	50	50	50	50
	A	1	0.98	0.98	1	1	1	1	0.98	1	1	1
	B	0	0	0	0	0	0	0	0.02	0	0	0
	C	0	0	0.02	0	0	0	0	0	0	0	0
	D	0	0.02	0	0	0	0	0	0	0	0	0
Idh-2	(N)	50	50	50	50	50	50	52	50	50	50	50
	A	0.02	0	0.01	0	0	0	0	0.01	0	0	0
	B	0.04	0.02	0.05	0.04	0.04	0	0.019	0.01	0.02	0	0.04
	C	0	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	0	0	0	0	0	0	0	0
	E	0.94	0.98	0.94	0.96	0.96	1	0.981	0.98	0.98	1	0.96

		Population										
Locus	Allele	12	13	14	15	16	17	18	19	20	21	22
Mdh	(N)	50	50	50	50	50	49	50	52	50	50	50
	A	0.83	0.94	0.96	0.97	0.96	0.918	0.99	0.942	0.99	0.9	0.88
	B	0.17	0.06	0.04	0.03	0.04	0.082	0.01	0.058	0.01	0.1	0.12
	C	0	0	0	0	0	0	0	0	0	0	0
Fdp	(N)	50	50	50	50	50	49	50	52	50	50	50
	A	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
	B	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Idh-1	(N)	50	50	50	50	50	49	50	52	50	50	50
	A	1	1	1	1	1	1	1	1	1	1	1
	B	0	0	0	0	0	0	0	0	0	0	0
	C	0	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	0	0	0	0	0	0	0	0
Idh-2	(N)	50	50	50	50	50	49	50	52	50	50	50
	A	0	0	0.01	0	0.02	0	0	0	0	0	0
	B	0	0.02	0.07	0.04	0.06	0.02	0.02	0	0.04	0.04	0.02
	C	0	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	0	0	0	0	0	0	0	0
	E	1	0.98	0.92	0.96	0.92	0.98	0.98	1	0.96	0.96	0.98

CHAPTER 3: Inbreeding in Western hemlock [*Tsuga heterophylla* (Raf.) Sarg.] Selective Breeding

3.1 Introduction

Western hemlock, *Tsuga heterophylla* (Raf.) Sarg., forms some of the most productive conifer stands in the world (Packee 1990). This member of the pine family (Pineaceae) is found in the coastal regions of British Columbia, Washington, Oregon and California. Like many conifers, western hemlock has a continuous geographic distribution, and extensive cross-pollination and pollen travel; therefore it should show very little population differentiation (Mosseler et al. 1994). King (1991) notes that gene flow both by gamete and zygote are so extensive in Western hemlock that there would be very little chance for genetically distinct demes to develop. In general, molecular genetic studies suggest that conifers have low levels of population differentiation, with less than 10% of genetic variation occurring among populations (Hamrick and Godt 1990).

Because of competition from other species, natural restocking of western hemlock following harvesting is not very successful. The practical option is to reforest stands using nursery-grown seedlings. Many seed orchards now provide seedlings for western hemlock reforestation, and in addition, first-generation breeding programs have been initiated for this species.

Paradoxically, realized gain trials of Western hemlock, carried out by the British Columbia Ministry of Forests, have found seed from unimproved orchards outperforms seed collected from wild stands by 5-8% (Charlie Cartwright, pers. comm.). This is much high than the theoretical 2% expected from these trees, which were "plus" trees selected in the wild. Alternative hypothesis for this anomalous gain are (1) wild collections contain inbred seed, either via self-fertilization or localized genetic differentiation, whereas orchard seed is outbred (outbreeding effects), (2) seed raised in the orchard is healthier and gives rise to more vigorous seedling (maternal effects) or (3) first generation selection of wild "plus" trees is more efficient than the expected 2%.

In this study, we test for the validity of the "outbreeding effect" (hypothesis 1), using estimates of inbreeding and population differentiation based upon isozyme loci.

This hypothesis seems the most likely, as early collection methods were crude, with seeds collected from the lower crown, where selfing is more likely. There currently is little information on population genetic structure in Western hemlock. Malavasi and Perry (1993), in their study of shade tolerance, identified variability within populations, and suggested that Western hemlock's competitive variability is a function of high genetic variation within populations, but no genetic markers were assayed. Thus our study also forms a baseline for levels of genetic differentiation in this species.

3.2 Methods and Materials

The seed orchard source population's geographic origins are listed in Table 3.1. From each of the 21 populations approximately 50 individuals were sampled. All natural population material was collected by the British Columbia Ministry of Forests. The unimproved seed orchards under went open-pollination. The only control that exists in these orchards is a small buffer of other species that inhibits the flow of outside pollen. The full-sib orchard seed samples originated from completely controlled pollination. Female strobili were isolated from one parent with a paper/plastic bag and pollen from only one other parent was then applied. To meet British Columbia's diversity requirement, 20 isolation bags are used on one ramet of one clone and each approximately covers 20 plus cones. The supplemental mass pollination orchards had pollen from best trees applied to receptive buds. Best pollen can be applied in the hopes that contamination from outside sources is minimized.

Starch gel electrophoresis was chosen as the molecular genetic technique for this study. Fourteen loci were identified using nine isozyme stains. Two electrophoresis systems were used to optimize the resolution of each locus. These included one continuous system, morpholine-citrate pH 8.0 (Clayton and Tretiak 1972) and one discontinuous system, sodium-borate pH 8.0 (Poulik 1957). The morpholine system had a running time of 4 hour sat 160V and included stains: shikimate dehydrogenase (SKDH), 6-phosphogluconate dehydrogenase (PGD), malate dehydrogenase (MDH), fructose-1,6-diphosphatase (FDP) and isocitric dehydrogenase (IDH). The Sodium-borate had a running time of 4 hours at 230V and included stains: aspartate

aminotransferase (AAT), glutamate dehydrogenase (GDH), phosphoglucoisomerase (PGI) and phosphoglucomutase (PGM). The histochemical staining solutions were obtained from Murphy et al. (1996) and Acquaah (1992).

All banding patterns were recorded visually. It can be noted that one tetramer (MDH) was found. MDH was scored as a monomer for 2 positions, hence two loci for the stain. The diversity index (H_t) and N_m were calculated using the program, POPGENE version 1.32 (Yeh et al. 1997). Wrights F-statistics, F_{IT} and F_{ST} and there corresponding standard errors were calculated using Mark (Ritland, unpub.). The F_{IT} is calculated from the mean and distribution of individual inbreeding coefficient estimates. The F_{ST} is calculated by averaging the pairwise relatedness for each individual averaged over all other individuals in a population. The standard error is obtained as the square root of the variance of F estimates, divided by the square root of the number of individuals.

3.3 Results

The level of local inbreeding (F_{IT}) and genetic differentiation between populations (F_{ST}) is listed for each of the 22 populations in Table 3.2. The average F_{IT} and F_{ST} in the natural, full-sib, unimproved and SMP orchard groups are given in Table 3.3. The gene flow (N_m) and heterozygosity (H_t) are also given for each of the orchard groups and natural populations in Table 3.3. F_{IT} ranged from 0.283 for S.O.126un to 0.090 for S.O.133un. The lowest was seen in the full-sib seed orchard (Table 3.3). The highest F_{IT} was seen in the natural population (Table 3.3).

F_{ST} ranged from 0.21 for S.O.133fs to 0.008 for both population 46152 and S.O.126un (Table 3.2). The full-sib orchards showed a much higher F_{ST} than any of the other orchard techniques and the natural populations (Table 3.3). The SMP and unimproved orchards showed relatively similar F_{ST} values when compared with the natural populations (Table 3.3). The SMP orchard showed the highest heterozygosity H_t (Table 3.3). The natural populations and unimproved and full-sib orchards showed similar H_t values with 0.146, 0.142 and 0.141 respectively (Table 3.3).

3.4 Discussion

As hypothesized, total inbreeding (F_{IT}) was highest in the natural populations (0.156, Table 3.3), with the full-sib seed orchards showing a significant reduction (0.120, Table 3.3). The SMP orchard also showed a significant reduction but it was not as great as the reduction in the full-sib orchard (0.145, Table 3.3). In contrast the unimproved orchard did not show a significant decrease in F_{IT} (Table 3.3). This result suggests that the controlled treatments are removing inbreeding that existed in the natural populations. This may be expected given that controlled pollination diminishes the occurrence of selfing. In addition, seed orchards likely contain a variety of material collected from several populations, so their total inbreeding due to among-orchard differences should be smaller. SMP and full-sib treatments have also been shown to improve both parental balance and reproductive synchrony, which will in turn reduce inbreeding (El-Kassaby 1995). Besides removal of inbreeding effects, SMP and full-sib pollination has the advantage of increasing the true genetic gain, by the introduction of desirable genetic material.

F_{ST}, the inbreeding due to among-population differentiation, was 0.057 for the natural populations (Table 3.3). This is slightly lower than that found in long-lived woody perennials with outcrossing and wind pollination (average F_{ST} =0.077, Hamrick et al., 1992). Hamrick et al. (1992) does estimate that widespread range species would have an approximate average F_{ST} of 0.033, bringing the overall estimate down. This may be expected given the high estimate of gene flow in the Western hemlock natural populations (8.15, Table 3.3). However, the gene flow value may be overestimated as Whitlock and McCauley (1999) suggest that there is non-equilibrium when N_m is calculated from F_{ST}. These results agree with the physiological work in hemlock that identified high variability with very little structure (Malavasi and Perry 1993).

In comparison, the unimproved and full-sib seed orchards showed a significantly higher estimate of F_{ST} (Table 3.3). This could be the result of controlled parental fertilization which minimizes contamination (El-Kassaby and Davidson, 1990) but tend to have just one male parent. The SMP orchard did not show such a high F_{ST} (Table 3.3).

This is likely due to pollen contamination combined with the larger number of males used in the pollen mix.

Total inbreeding (F_{IT}) decreased by 2-3% in the seed orchards. While significant and expected, this is probably an insufficient decrease to explain the anomalous genetic gain recorded in realized gain trials. Rather, non-genetic effects such as maternal effects (healthier seed) may underlie the anomalous genetic gain seen in unimproved seed orchards. As well, efficient plus tree selection may be another component of this genetic gain. Selection intensity is quite high at the advanced stages of best tree selection. The most highly selected individual have a selection proportion defined as 1/30 (King et al. 1997).

Gene diversity is compromised when only the best genotypes are used for pollination. Gene diversity must be considered in orchard management as changes in selection pressure as a result of factors such as climate change may require a large amount of trait diversity for a species survival. A balance then exists between the loss of diversity and selection intensity. Selection intensity has to be high enough to obtain substantial genetic gain while the breeding population must be maintained at a sufficient size to retain allelic diversity (Johnson 2001). In terms of orchard structure, controlled pollination techniques can be implemented to conserve genetic diversity while maintaining selection intensity.

Compared to wild populations, total inbreeding (F_{IT}) was lower in all orchard populations but only significantly so in the full-sib orchard (Table 3.3). The full-sib orchard also had a significantly higher F_{ST} when compared with all other orchards (Table 3.3). Given the apparent decrease in inbreeding, the full-sib orchard showed a lower value of diversity when compared with other orchard techniques. Depending on the goal of seed orchard management, this lack of diversity may outweigh any decrease in inbreeding. In contrast the SMP orchard showed a much higher heterozygosity (H_t) when compared with both the natural populations and the other orchard populations (Table 3.3). Therefore one can conclude that supplemental mass pollination may offer the balance of strong selection intensity while still maintaining sufficient gene diversity.

3.5 References

Acquaah, G. 1992. *Practical Protein Electrophoresis for Genetic Research.* Theodore R. Dudley, general editor. Dioscorides Press. Portland Oregon.

Clayton, J.W. and D.N. Tretiak. 1972. Amine-citrate buffers for pH control in starch gel electrophoresis. Journal of Fish Research. Board Canada. **29**: 1169 – 1172.

El-Kassaby, Y.A. and R. Davidson. 1990. Impact of crop management practices on the seed crop genetic quality in a Douglas-fir seed orchard. Silvae Genetica **39**: 230 – 237.

El-Kassaby, Y.A.. 1995. Evaluation of the tree-improvement delivery system: factors affecting genetic potential. Tree Physiology. **15**: 545-550.

Foster, G.S. and D.T. Lester. 1983. Fifth-year height variation in western hemlock open -pollinated families growing on four test sites. Canadian Journal Forest Research **13**: 251 – 256.

Fowells, H.A. 1965. Silvics of forest trees of the United States. U.S.D.A., Wash., D.C. Agricultural Handbook. No. 271.

Fujimori, T. 1971. Primary productivity of a young *Tsuga heterophylla* stand and som e speculations about biomass of coast communities on the Oregon coast. USDA Forest service, Research paper PNW – 123. Pacific Northwest Forest and Range Experiment Station, Portland, Oregon. Pp 11.

Hamrick, J.L. and M.J.W. Godt. 1990. 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, Sunderland, Mass. 43 – 63.

Hamrick, J.L., M.J.W. Godt and S.L. Sherman-Broyles. 1992. Factors influencing levels of genetic diversity in woody plant species. New Forest **6**: 95 – 124.

Hamrick, J.L., J.B. Mitton and Y.B. Linhart. 1981. Levels of genetic variation in trees: influence of life history characteristics. *In* Isozymes of North America forest trees and forest insects. *Edited by* M.T. Conkle. USDA Forest Service General Technical Representatives. PSW-48: 35 – 41.

Hepting, G.H. 1971. Disease of forest and shade trees of the United States. U.S.

Department of Agriculture, Agriculture Handbook 386. Washington, DC. 658 p.

Hosie, R.C. 1975. Native trees of Canada. Department of Environment, Can. For. Serv., Ottawa, Ont.

Isaac, L.A. 1930. Seed flight in the Douglas-fir Region. Journal of Forestry **28**: 492 – 499.

Johnson, J., B. St.Clair and S. Lipow. 2001. Genetic conservation in applied tree breeding programs. In: Proceedings ITTO conference on in situ and ex situ conservation of commercial tropical trees pp. 215-230.

King J.N. 1991. The significance of geographic variation patterns for western hemlock genetic improvement. Technical Report, B.C. Ministry of Forests Research Branch.

King, J. N., C. Cartwright, and D. Cress. 1997. Western hemlock tree improvment: selection of P-1 parents. B.C. Ministry of Forests, Victoria B.C.

Kuser, J.E. and K.K. Ching. 1980. Provenance variation in phenology and cold hardiness of western hemlock seedlings. Forest Science **26**: 463 – 470.

Kuser, J.E. and K.K. Ching. 1981. Provenance variation in seed weight, cotyledon number, and growth rate of western hemlock seedlings. Forest Science **26**: 463 – 470.

Malavasi, U.C. and D.A. Perry. 1993. Genetic variation in competitive ability of some shade-tolerant and shade-intolerant pacific Coast (USA) conifers. Forest Ecology and Management **56**: 69 – 81.

Meagher, M.D. 1976. Studies of variation in hemlock (*Tsuga*) populations and individuals from southern British Columbia. Ph. D. thesis, University of British Columbia, Vancouver, B.C.

Mosseler, A., K.N. Egger and S.M. Carr. 1994. Molecular biology and genetic diversity in tree populations. *In.* Recent progress in Forest Biotechnology in Canada. Ed. Charest and Duchense. 114 – 124.

Murphy, R.W., J.W. Sites Jr., D.G. Buth, and C.H. Haufler. 1996. Proteins: Isozyme Electrophoresis. In: *Molecular Systematics*. 2nd Ed. Hillis, David M., Craig Moritz, Barbara K. Mable (eds.). Sinauer Associates, Inc. Sunderland, Massachusetts. 51 – 120.

Owens, J.N. and M. Molder. 1984. The reproductive cycles of western and mountain hemlock. Biology Dept, UVIC, Victoria B.C.. Ministry of Forests. Pp 30.

Packee, E.C. 1976. The ecology of western hemlock. In Proceedings, Western Hemlock Management Conference. P. 10 – 25. W. A. Atkinson and R. J. Zasoski, eds. University of Washington, College of Forest Resources, Seatle.

Packee, E.C. 1990. *Tsuga heterophylla* (Raf.) Sarg: Western hemlock. In *Silvics of North America: Volume 1, conifers.* Burns, R. M.; Honkala, B. H. (tech. coords). Agriculture Handbook 654. Washington, DC: U.S. Department of Agriculture, Forest Service: 260 – 267.

Poulik, M.D. 1957. Starch gel electrophoresis in a discontinuous system buffers. Nature 180: 1477 – 1479.

Ritland, K. 2000. Forest renewal B.C. operational improvement proposal. Forest Genetic Council of British Columbia. Pp 2.

Ritland, K. 2002. Extensions of models for the estimation of mating systems using n independent loci. Heredity 88: 221-228.

Whitlock, M.C. and D.E. McCauley. 1999. Indirect measures of gene flow and migration: $F_{st} \neq 1/(4Nm + 1)$ Heredity **82**: 117 – 125.

Williamson, R.L. 1976. Natural regeneration of western hemlock. *In* Western Hemlock Management, Inst. Of For. Prods., Univ. Wash., Seattle Wash. Pp. 166 – 169.

Yeh, F., Yang, C., Boyle, R.C., Timothy, B.J., Ye, Z.H., and J.X. Mao. 1997. POPGENE, the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada.

3.6 Tables and Figures

Table 3.1. Populations, year of collection and breeding zone of *T. heterophylla* seed.

Seedlot	ID or location	Year	Latitude	Longitude	Elevation (m)
2685	Holberg	1975	50 38	128 05	92
2753	Toba River	1975	50 30	124 10	366
3471	Ucona River	1978	49 40	126 00	475
4088	Fleet River	1979	48 38	124 04	370
4538	Sombrio Creek	1978	48 32	124 18	365
4692	Camper Creek	1982	48 34	124 30	300
4787	Sechelt	1976			
7832	Nanaimo River	1987	49 00	124 10	550
9789	UBC Res. Forest	1985	49 17	122 33	275
18784	Kaouk Jnc.	1982	50 05	126 59	60
46152	Unknown natural		48 50	123 45	600
60160	S.O.# 126 - unimproved	1993	50 27	127 09	188
60352	S.O.# 126 - SMP	1998	50 30	126 53	95
60351	S.O.# 126 - full-sibs	1998	50 02	125 52	121
6883	S.O.# 133 - unimproved	1990	50 00	124 30	140
61060	S.O.# 133 - SMP	1999	50 12	125 08	300
60624	S.O.# 133 - full-sib	1997	50 00	125 00	300
60183	S.O.# 136 - unimproved	1993	49 38	126 08	169
60224	S.O.# 156 - unimproved	1995	52 43	131 40	241
60319	S.O.# 130 - unimproved	1995	48 47	124 18	595
60374	S.O.# 143 - unimproved	1998	49 54	124 29	745
60067	S.O.# 127 - unimproved		50 37	127 19	616

Table 3.2. Levels of total inbreeding, F_{IT}, and genetic differentiation between populations, F_{ST}, as calculated for 11 natural (nat), populations and 2 full-sib (fs), 2 supplemental mass pollination (smp) and 6 unimproved (un) seed orchard populations of *T. heterophylla*.

Locus	Sample Size	F_{IT}	F_{IT} SE	F_{ST}	F_{ST} SE
Holerg (nat)	49	0.144	0.04	0.023	0.017
Toba (nat)	49	0.105	0.048	0.058	0.024
Ucona (nat)	49	0.239	0.119	0.033	0.008
Fleet (nat)	50	0.165	0.026	0.042	0.017
Sombrio (nat)	50	0.122	0.035	0.02	0.014
Camper (nat)	50	0.123	0.027	0.195	0.022
Sechelt (nat)	52	0.096	0.044	0.028	0.014
Nanaimo (nat)	50	0.11	0.036	0.131	0.025
UBC (nat)	49	0.237	0.038	0.059	0.02
Kaouk (nat)	50	0.173	0.055	0.027	0.013
46152 (nat)	49	0.205	0.069	0.008	0.014
SO 126 (un)	50	0.283	0.075	0.008	0.012
SO 126 (smp)	49	0.113	0.045	0.058	0.024
SO 126 (fs)	50	0.128	0.039	0.053	0.027
SO 133 (un)	50	0.09	0.037	0.094	0.017
SO 133 (smp)	50	0.177	0.039	0.035	0.033
SO 133 (fs)	50	0.112	0.031	0.21	0.028
SO 136 (un)	50	0.191	0.065	0.048	0.017
SO 156 (un)	50	0.1	0.021	0.075	0.012
SO 130 (un)	52	0.157	0.041	0.062	0.011
SO 143 (un)	50	0.106	0.022	0.125	0.023

Table 3.3. Total inbreeding, F_{IT}, and genetic differentiation between populations, F_{ST}, with standard errors and gene flow, N_m, and gene diversity, H_t as calculated for the natural and seed orchard groups of *T. heterophylla*.

Locus	Sample Size	F_{IT}	F_{IT} S.E.	F_{ST}	F_{ST} S.E.	H_t	*N_m
Natural	547	0.156	0.0148	0.057	0.0052	0.146	8.15
S.O. UNIMP	302	0.155	0.0177	0.068	0.0063	0.142	
S.O. SMP	99	0.145	0.0297	0.0465	0.0201	0.164	
S.O. FS	100	0.120	0.0247	0.1315	0.0194	0.141	

*N_m = *Gene flow estimated from* $F_{ST} = 0.25(1 - F_{ST})/ F_{ST}$

CHAPTER 4: Microsatellite Markers in Western hemlock [*Tsuga heterophylla* (Raf.) Sarg]

4.1 Primer Note

Western hemlock [*Tsuga heterophyla* (Raf) Sarg.], a conifer with a rapid growth rate, has gained economic importance in recent years. With the threat of depleted stands in Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco] plantations by various diseases, western hemlock have been identified as an alternative species (Hansen et al. 2000). It is known to yield some of the most productive stands in the world (Packee 1990). This species is recognized as an all-purpose raw material in the forest industry and has been defined as good to excellent in its pulping characteristics (Packee 1990). Mountain hemlock [*Tsuga mertensiana* (Bong.) Carr.], a minor commercial tree species, used in small-dimension lumber, will also be considered for interspecific amplification. Mountain hemlock also serves as a slope stabilizer and provides a component of wildlife habitat (Taylor and Taylor 1980). Characterized by disjunct populations and low density stands at the edges of its range, mountain hemlock is considered by some as a threatened species (Farjon et al. 1993).

Genetic improvement has been recognized as a cost effective way to improve on productivity in western hemlock plantations (Jayawickrama 2002). Given selection pressure present in hemlock breeding programs it is important to monitor changes in genetic variation. Aggressive selection for specific traits can result in individuals with very low genetic diversity which may cause subsequent inbreeding depression in future generations. The high level of polymorphism and its codominant nature make microsatellites very useful in describing changes in genetic variation. It may be possible to use microsatellites to predict specific combining ability among the top selected seedling (on the basis of coancestry).

Microsatellite loci were isolated from western hemlock genomic DNA using modifications of published biotin-enrichment strategies (Bérubé et al. 2003). Clones probed and selected for possessing AG repeats were stored in glycerol at – 80 °C. One hundred and ninety-two clones were amplified and sequenced directly from glycerol stocks using SequiTherm EXCEL™ II Long-Read DNA Sequencing kits-LC (Epicentre

Technologies) on a LiCor 4200 (LiCor Inc. Lincoln, NB). Of the 192 clones 87 were suitable for designing primers. The remainder had either very short repeats (less than 10) or repeats at one end of the insert, precluding adequate primer sites. The primers were designed using OLIGO 6.3 (Molecular Biology Insight Inc., Cascade, CO), and they were then tested and resolved on 5-7% polyacrylamide gels (Long Ranger™) using a LiCor 4200 automated sequencer (LiCor Inc.). From the 87 primers sets, 15 sets were finalized. Seventy-two primer sets were excluded as they were either irresolveable using the gel based system or were monomorphic. The size and highly repetitive nature of the conifer genome have been reported to be a major cause of low recovery of microsatellites from genomic DNA (Fischer and Bachmann 1998).

Foliage samples of western hemlock and mountain hemlock were collected from wild stands to determine microsatellite polymorphism. Between 8 to 24 western hemlock individuals and 10 or 195 mountain hemlock individuals were tested per primer. Total genomic DNA was extracted from both species using a modified CTAB method (Doyle and Doyle 1990).

Polymerase chain reactions (PCR) were carried out in 10 μl volume using an MJ Research PTC-100 thermal cycler (MJ Research, Inc.). Each reaction was composed of 10 to 80 ng of total genomic DNA (Table 4.1), 1.0 ul of 2.0mM dNTP, 1× Taq Buffer (10 mM Tris-HCl, 1.5 mM $MgCl_2$, 50 mM KCl, pH 8.3; Roche), 0.15 U of Taq DNA Polymerase (Roche), and 0.4 - 0.8pmol of M13 Infrared Label Primer (LiCor Inc.) (Table 4.1). The reaction mixture for mountain hemlock contained 1.0ul of 0.01% Erica Haddelberg buffer (10mM Tris-HCl pH=8.0, 50mM KCl, 2.5mM $MgCl_2$, 0.01% gelatin and 0.16ug/ul BSA), 1.0uL of 2.0mM dNTP, 0.5 pmol each of forward and reverse primers, 0.5 pmol M13 IRD-labeled primer, 1Unit Taq DNA Polymerase, and 30-50ng of genomic DNA template.

Samples were amplified using a basic or touch-down protocol. The basic program is as follows: Initial denaturing for 3 min at 95°C; subsequent cycling for 45 sec at 95°C, 45 sec at the respective annealing temperature (see Table 4.1, 4.2), 1 min at 72 °C, repeated 35 times; a final extension cycle of 4 min at 72 °C. The touch-down protocol is as follows: initial denaturation of 3 min at 95 °C; cycling (5 times) with 45 seconds at 95 °C, 45 sec at the higher respective annealing temperature (see Table 4.1, 4.2), 1 min at 72 °C, back to 95 °C; then cycling (30 times) with 45 sec at 95 °C, 45 sec at the lower

respective annealing temperature (see Table 4.1, 4.2), 1 min at 72 °C; ending with a final extension cycle of 4 min at 72 °C. Nine of fifteen dinucleotide microsatellite loci were used to determine genetic polymorphism.

Allelic diversity was generally high in western hemlock individuals ranging from 7 (EE 12) to 15 alleles (EE 10) with a mean per locus of 11. The mean observed heterozygosity (H_o) averaged 0.76 and the mean expected heterozygosity (H_e) averaged 0.88 (determined as by Nei 1978) (Table 4.1). Mean observed heterozygosity is higher in this study than that reported for previously designed western hemlock microsatellites primers (Amarasinghe et al. 2002). No null alleles were found (EH03-2 and EC 10) based on analysis of controlled crosses (1 female: 1 male) for the limited testing done. The presence of highly conserved flanking regions regularly reported in microsatellites (Zane et al. 2002) allowed the cross-amplification of mountain hemlock. This cross-species amplification was successful with 6 primer pairs tested (Table 4.3). Three of these loci were used to calculate expected heterozygosity (Table 4.3). Allelic diversity in mountain hemlock individuals was higher than in western hemlock, ranging from 5 (AE05) to 30 (EE12) (Table 4.3). The mean observed heterozygosity (H_o) was 0.87 and the mean expected heterozygosity (H_e) was 0.89 (Table 4.3).

4.2 References

Amerasinghe V, Brown GR, Mank JE, Carlson JE (2002) Microsatellite DNA loci for Western Hemlock [*Tsuga heterophylla* (Raf.) Sarg]. *Molecular Ecology Notes* **2**, 236-238.

Bérubé Y, Ritland CE, Ritland KM (2003) Isolation, characterization, and cross-species utility of microsatellites in yellow cedar (*Chamaecyparis nootkatensis*). Genome, **46**, 353-361.

Doyle JJ & Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus*, **23**, 13-15.

Farjon A, Page CN, Shellevis N (1993) A preliminary world list of threatened conifer taxa. Biodiv. Conserv. **2**, 304-326.

Fisher D & Bachmann K (1998) Microsatellite enrichment in organisms with large genomes (*Allium cepa* L.). *Biotechniques*, **24**, 796-798.

Hansen EM, Stone JK, Capitano BR, Rosso P, Sutton W, Winton L, Kanaskie A,

McWilliams MG (2000) Incidence and impact of Swiss needle cast in forest plantations of Douglas-fir in coastal Oregon. *Plant Diseases*, **84**, 773-778.

Jayawickrama KJS (2002) Genetic improvement and deployment of western hemlock in Oregon and Washington: Review and future prospects. *Silvae Genetica*, **in press**.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a smaller number of individuals. *Genetics*, **89**, 583-590.

Packee EC (1990) *Tsuga heterophylla* (Raf.) Sarg: Western hemlock. In *Silvics of North America: Volume 1, conifers.* Burns, R. M.; Honkala, B. H. (tech. coords). Agriculture Handbook 654. Washington, DC: U.S. Department of Agriculture, Forest Service, 260 – 267.

Taylor, Roy L, and Sylvia Taylor (1980) Tsuga *mertensiana* in British Columbia. *Davidsonia* **11**(4):78-84.

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology*, **11**, 1-16.

Table 4.3 Tables and Figures

Table 4.1. Characteristics of nine developed microsatellite loci in *Tsuga heterophylla*.

Locus		Primer sequence (5' to 3')	Motif	T_a (°C)	M13 (*ul*)	DNA (ng)	No. of alleles	Size range (bp)	N	H_o	H_e	Genbank accession no.
AA01	F:	GGACTCTCTCATGTATTGCTATG	$(CT)_{41}$	57/54 (TD)	0.8	80	11	329 - 303	17	0.9375	0.899	AY279983
	R:	CGCAAAGGGCAACCAAGGAAGAC										
AA06	F:	AGCACACACACTTACCTCTCAAG	$(CT)_{18}$	62 (B)	0.8	40	8	296 - 266	8	1	0.926	AY279984
	R:	AGTACACAACAATATATCTTGGG										
EC07	F:	GACCATGATCAATCTGGGAGTG	$(CT)_{29}$	57/54 (B)	0.8	40	11	162 - 124	9	1	0.954	AY279988
	R:	AGTTCTCCTAGTGATACACG										
EC10	F:	CCTTAGGACTACTCTCTCT	$(CT)_{26}$	55/52 (TD)	0.5	10	15	410 - 306	24	0.708	0.904	AY279989
	R:	CAGCATCAAGGAAGATTTC										
EC12	F:	CATTAAATTTGGGATATGCAAG	$(GT)_{21}$	51/49 (TD)	0.4	40	15	213 - 159	23	0.609	0.854	AY279990
	R:	CCCAAAGGATCAAAATCTATTT										
EE06	F:	GGGTGTGTAGAGATCTAGTGTAG	$(GA)_{17}$	52 (B)	0.5	15	12	308 - 238	14	0.429	0.926	AY279992
	R:	GYAGCATAAACAATGTAKAGATG										
EE10	F:	CACCTTCCAATTTTCAACTCT	$(CT)_{17}$	58/55 (TD)	0.4	10	11	141 - 107	24	0.542	0.864	AY279993
	R:	GCCCAAGGAGATGGCTTTTGC										
EE12	F:	AAACAACCCATGTTGCTTTCA	$(AC)_{14}$	58/55 (TD)	0.4	40	7	190 - 158	17	0.706	0.784	AY279994
	R:	CCGCTGGGAACCGATAGGAGG										
EH032	F:	TGGGAAAGGAGGGTTAAATAG	$(AG)_{24}$	63/60 (TD)	0.7	40	8	236 - 202	17	0.941	0.845	AY279997
	R:	AGACCACCTTCTACCTCAAGC										

T_a, anealling temperature; B, Basic amplification protocol; TD, Touch-down amplification protocol; H_o, observed heterozygosity; He, expected heterozygosity; N, number of individuals tested

Table 4.2. Characteristics of five microsatellite loci that require further development in *Tsuga heterophylla*.

Locus		Primer sequence (5' to 3')	Motif	T_a (°C)	M13 (*ul*)	DNA (ng)	Size range (bp) with tail	Genbank accession no.
AD05	F:	CTTGCTTGAATCTATTAGTTGAG	$(AG)_{18}$	49 (B)	0.5	30	173	AY279986
	R:	ACCTAAACTCCATTTTCT						
AE05	F:	CCAAAACCAACATGCCCTAGTTC	$(AG)_{48}$	59/56 (TD)	0.8	30	359	AY279987
	R:	TAAACATTGCTCTTCCTGCCCAC						
ED08	F:	AAGAGATCATAACCCAAATAC	$(AC)_{24}$	55/50 (TD)	0.5	40	213	AY279991
	R:	CCCCATAGAAAATTGTGAGAC						
EF01	F:	CACCACCCCTGTCTTTAACTCT	$(CT)_{22}$	62 (B)	0.5	40	301	AY279995
	R:	TGAAGGGTGGATTAGGGAGAT						
EG12	F:	GAGGCTAGAGGCATGCATGGC	$(CT)_{22}$	56/53 (TD)	0.8	30	163	AY279996
	R:	CATGTGTAGACAAGATRAGGG						

T_a, anealling temperature; B, Basic amplification protocol; TD, Touch-down amplification protocol;

Table 4.3. Characteristics of 6 microsatellite cross species amplification in *Tsuga mertensiana* (mountain hemlock).

Locus	T_a (°C)	No. of alleles	Size range (bp) without tail	N	H_o	H_e
AE05	58/55 (TD)	5	254-266	10	—	—
EC10	52 (B)	11	390-446	195	—	—
EE06	58/52 (TD)	23	239-303	195	0.8667	0.8924
EE10	58 (B)	18	94-140	195	0.8821	0.8399
EE12	58/55 (TD)	30	125-257	195	0.8564	0.9248
EH03-2	65/62 (TD)	10	214-298	10	—	—

Ta, annealing temperature; B, Basic amplification protocol; TD, Touch-down amplification protocol; H_o, observed heterozygosity; H_e, expected heterozygosity; N, number of individuals tested

CHAPTER 5: Heterosis-Fitness Correlations in Western hemlock *Tsuga heterophylla* [*Tsuga heterophylla* (Raf.) Sarg.]

5.1 Introduction

Selection for more fit individuals has long been the primary objective of most tree breeding programs. Fitness has long been associated with factors that are of prime interest to breeders such as height, volume and density. These traits are the consequence of the frequency of favorable alleles. Breeders may know what traits they are selecting for but they do not know what alleles lead to their expression. They also do not know the distribution of individuals that harbor that specific trait in the natural environment (Johnson et al. 2001). As a consequence breeding programs must maintain adequate genetic resource for advanced selection of desirable traits. While breeders must maintain extensive genetic entries they also must endure lengthy assessment periods with respect to selective field testing. Performance indicators such as height are often not indicative of genotypic expression for up to 5 years (Arcade et al. 1996). Therefore, predictive techniques such as HFC (heterozygosity fitness correlation) could assist in reducing the expensive and labour-intensive task of selection trials (Vaillancourt et al. 1995).

The genetic distance between two markers may be indicative of the closeness of their relationship and as such may be predictive of firstly, the number of deleterious alleles shared by the two parents and secondly the degree of expected heterozygosity of the progeny. These two factors, termed inbreeding depression, can conclude in reduced fitness with increasing genetic relatedness (Wright 1921, Crow 1948). This can be contrasted against outbreeding depression. Outbreeding depression results in reduced fitness with extreme outcrossing and is attributable to the randomization of locally adapted traits (Templeton 1986).

The heterozygosity-fitness correlation can be used to predict the fitness of offspring before mating, based on the genetic distance between their parents. The use of HFC could ultimately reduce the number of parents that would have to be tested. If a strong relationship can be defined for HFC, heterozygosity can be maintained through selection for greater genetic distance while increasing the overall fitness of the next

generation population. A strong correlation may allow for the prediction of F1 performance in Western hemlock offspring greatly reducing the number of individuals that would have to be tested in selection trials.

Western hemlock (*Tsuga heterophylla* (Raf.) Sarg.) is found along the pacific coast from Alaska south through northern California (Packee 1990). Western hemlock is recognized as an all-purpose raw material for construction purposes with over 7.9 million sowing requests in 1997 (Webber 2000). An extensive breeding program exists for Western hemlock with original selections made in the early 70's (King et al. 1997). To date more than 200 families are represented in orchards from Oregon to British Columbia. The goal of this study is to test HFC in elite western hemlock families and identify if it may be a good predictor of higher fitness in their progeny.

5.2 Materials and Methods

5.2.1 Material Collection

The top 30 elite families of western hemlock were selected from the Western Hemlock Tree Cooperative, HEMTIC bank of orchard material. Of the 30 elite families, 6 were taken from each of the 5 programs spanning throughout the coastal regions of western hemlocks range. They were defined as elites based on the performance of their progeny and were quantified using a breeding value (Table 5.1) Their locations and elevations are also given in Table 5.1, and in Figure 5.1. The breeding value of a family is a reflection of the average performance of its tested progeny. Material for all 30 elite families was collected from established seed orchards located at the B.C. Ministry of Forests, Cowichan Valley Research Station, Vancouver Island B.C.. Needle and bud tissue were sampled from the upper crown of 10 year old seedlings in late fall. Tissue was stored at -80°C prior to DNA extraction.

Measurements of height at 5 years, deaths (mortality rate), presence of forks and frost damage were collected for progeny that were the result of crosses between the 30 elite families. The measurements were performed by Charlie Cartwright of the B.C. Ministry of Forests. The progeny trials were located at 5 different locations including,

Michelson (MKL), Stove (STV), Tlupana (TLU), BR 265 (B265) and Volmer (VOLMR) given in Figure 5.2. Approximately 30 progeny exist in a randomized block design for each cross at each of the 5 progeny trials.

5.2.2 DNA Extraction and Genotyping

Genomic DNA was isolated from *Tsuga heterophylla* foliage using the hexadecyltrimethyl-ammonium bromide (CTAB) method described by Doyle and Doyle (1990). Over 1 gram of tissue was needed to get usable concentrations of DNA. Further purification was also performed for some samples. This was done using the lysates isolation method described for the Genomic Tip 20 kit (Qiagen Inc.). Seven microsatellites (AC07, EC02, EC10, EC12, EE06, EE10 and EE12) developed for *Tsuga heterophylla* were used for genotyping (Chapter 4).

Polymerase chain reactions (PCR) were carried out in 10 μl volume using an MJ Research PTC-100 thermal cycler (MJ Research, Inc.). Each reaction was composed of 10 to 50 ng of total genomic DNA, 1.0 ul of 2.0mM dNTP, 1× Taq Buffer (10 mM Tris-HCl, 1.5 mM $MgCl_2$, 50 mM KCl, pH 8.3; Roche), 0.15 U of Taq DNA Polymerase (Roche), and 0.4 - 0.8pmol of M13 Infrared Label Primer (LiCor Inc.).

Samples were amplified using a "touch-down" protocol which is as follows: initial denaturation of 3 min at 95 °C; cycling (5 times) with 45 seconds at 95 °C, 45 sec at the higher respective annealing temperature, 1 min at 72 °C, back to 95 °C; then cycling (30 times) with 45 sec at 95 °C, 45 sec at the lower respective annealing temperature, 1 min at 72 °C; ending with a final extension cycle of 4 min at 72 °C (Chapter 4).

After the amplification was complete 3ul of loading dye (stop dye) was added to each reaction. The microsatellites then underwent electrophoresis on a LI-COR 4200 sequencer using 6% w/v polyacrylamide gels (Long Ranger™, BioWhittaker Molecular Applications, Rockland, Maine). They were then scored using a ladder based on size using the program, RFLPscan Plus (Scanalytics Inc. FairFax VA).

5.2.3 Statistical Analysis

Two coefficients were chosen as a measure of relationship between pairs of individuals. The first is an unbiased pairwise measure of genetic relatedness, 'r' (Lynch and Ritland 1999). This coefficient was chosen over Ritland's (1996) measure as it yields lower sampling variances when loci are highly variable, such is the case with microsatellites. Genetic Relatedness 'r' was calculated using SPAGeDi 1.1 (Hardy and Veckemans 2002). The second, 'mean d^2' (Coulson et al. 1998) is a measure of genetic distance that is estimated from alleles at a locus for each individual averaged over loci. Mean d^2 is calculated as "the squared distance in repeat units between the two alleles an individual had at a microsatellite locus, averaged over all loci at which an individual was scored" (Coulson et al. 1998). Mean d^2 is therefore a measure of the genetic distance between gametes that formed the individual (Coulson et al. 1998). This measure was chosen as it measures evolutionary distance, not pedigree relatedness, and will include differences in allele length between populations due to the stepwise mutation model which is the general model assumed when using microsatellites. Genetic Distance 'mean d^2' was calculated using MARK (Ritland 2004). Mean d^2 is slightly different than what is found in Coulson's (1998) paper as each locus is divided by the variance of the allele size distribution. This was done to standardize for differences in mutation rates among loci (Ritland pers. comm. 2004).

Allele frequencies and expected heterozygosity for each microsatellite and a pairwise spatial distance (Euclidian Distance) based on spatial coordinates (Latitude and Longitude) for all crosses of elite parents were also calculated using SPAGeDi 1.1 (Hardy and Veckemans 2002). All histograms and there corresponding regressions were calculated using Excel XP (Microsoft 2002). All other regressions (Height at age 5 vs. Genetic Relatedness, Genetic Distance and Geographic distance) were performed using the Regression and Graph wizards in SigmaPlot 8.02 (SigmaPlot 2002). It can be noted that a number of assumptions associated with regression analysis must be met to avoid biased results. Therefore each regression was tested for homoskedasticity, independence, normality and lack of fit and results were not listed when any assumption were violated.

The ANOVA analysis used to calculate if there was a difference in within and between variances was calculated using SAS version 8.02 (SAS 2001).

5.3 Results

All 7 Loci showed high polymorphism with the number of alleles ranging from 10 for AC07/EC02 to 23 for EC12/EE06 and expected heterozygosity ranging from 0.8345 for EC02 to 0.9503 for EC12 (Table 5.2). Allele frequencies graphs are given in Appendix 5.A.1 – a. to g. EC02, EC12, EE10, EE06 and EE12 all show a bell shaped distribution. AC07 and EC10 show more of a bimodal distribution.

The mean height at year 5 for approximately 30 F1 progeny for each cross was first regressed against the genetic relatedness 'r' (Lynch and Ritland 1999) calculated between the parents (Figure 5.3). This was calculated for each progeny trial separately and then combined for all progeny trials. Each of the progeny trials were also graphed separately (Appendix 5.B.1- 7). Basic coefficients for each of the 6 regressions are then given in Table 5.3. The R^2 for all regression are very low ranging from 0.0000003 to 0.0030006 and 0.00002483 when treated as a group (Table 5.3). All of the relationships were not significant with *P* values ranging from 0.9953 to 0.5215 and a group *P* value of 0.8961, so any trend must be interpreted cautiously (Table 5.3). The slope is positive for the VOLMR and B265 trials, with 12.05 and 3.63, respectively, and negative for trials MKL STV and TLU ranging from -0.12 to -10.15 (Table 5.3). It can be noted that both trials with a positive correlation between mean height and genetic relatedness in their parents are found at the southern end of the trial locations (Figure 5.2). The grouped regression concluded in a slightly negative slope (Figure 5.3).

The mean height at year 5 for approximately 30 F1 progeny for each cross was then regressed against mean d^2 (Coulson et al. 1998) calculated between the parents (Figure 5.4). This was also calculated for each progeny trial separately and then combined for all progeny trials. Each of the progeny trials were again graphed separately (Appendix 5.C.1-7). Similarly, basic coefficients for each of the 6 regressions are then given in Table 5.4. In contrast to the genetic relatedness measure, two progeny trials were found to have a significant relationship at 95% for genetic distance, with *P* values of

0.0264 and 0.0208 for MKL and STV, respectively. However the R^2 values were quite low for both 0.0352284 and 0.0378565, respectively (Table 5.4). All other progeny trials show non-significant relationships for genetic distance with *P* values ranging from 0.9664 to 0.7419 and a group *P* value of 0.1970 (Table 5.4). The most northern and only significant progeny trials, MKL and STV resulted in a negative slope with -8.04 and -8.08, respectively, where as the B265, TLU and VOLMR trials were basically 0, ranging from -1.03 to 0.9 (Table 5.4, Figure 5.2). The grouped regression also resulted in a slightly negative slope (Figure 5.4).

The mean height at year 5 for the 30 F1 progeny was then regressed against the geographic distance between the parents (top elites) for each progeny trial (Figure 5.5). All basic coefficients related to Figure 5.7 are given in Table 5.6. Again, none of the progeny trials showed a significant relationship with *P* values ranging from 0.4279 to 0.1246. The slopes were not very strong for any of the progeny trials ranging from 1.605 to -0.9835 (Table 5.5). The R^2 values were a bit higher for these regressions ranging from 0.01689 for the STV trial to 0.00820 for the VOLMR trial. The relationship between geographic distance and height at age 5 strengthens as the trials move northward (Figure 5.2). The MKL trial is the only exception to this trend.

The geographic distance was then regressed against both genetic distance (mean d^2) and relatedness ('r') (Figures 5.6 and 5.7). The coefficients are placed within the figures. The R^2 values were low at 0.03475 and 0.00359 for genetic relatedness and genetic distance respectively (Figures 5.6 and 5.7). The slope was negative for relatedness and positive for genetic distance.

The genetic relatedness ('r') and genetic distance (mean d^2) between the top 30 elites was then regressed against the frequency of deaths of the F1 progeny for each of the Progeny trials in Figures 5.8 and 5.9 respectively. The frequency of forks and frost damage was also regressed against both relatedness and distance. The regressions are used as an indication of outbreeding depression. All coefficients related to Figures 5.8 and 5.9 are then given in Tables 5.8 and 5.9 respectively. Each trial is also graphed separately for simpler interpretation (Appendix 5.B.1-7, 5.C.1-7). The R^2 values were much higher in these regressions ranging from 0.0173 for presence of forks to 0.4975 for frequency of deaths (TLU progeny trial) in respect to genetic relatedness and ranging

from 0.0159 for frequency of death (B265 progeny trial) to 0.1433 for frequency of deaths (VOLMR progeny trial) in respect to genetic distance (Tables 5.6 and 5.7). Given the higher R^2 values only 3 progeny trials, STV, MKL and TLU, showed a significant relationship with P values of 0.0133, 0.0275 and 0.0007 respectively (Table 5.6). The significant relationships only existed while using the genetic relatedness measure. The significant progeny trials were also noted as being the most northern progeny trials (Figure 5.2). The presence of forks and frost damage showed a smaller relationship than frequency of deaths to genetic relatedness (Table 5.6). This relationship did not hold for genetic distance (Table 5.7). The slope was close to zero for all relationships with the genetic relatedness regressions ranging from 0.0013 to -0.0023 for number of forks and frequency of deaths in the MKL progeny trial, respectively, and the genetic distance regressions ranging from 0.0009 to -0.0015 for frequency of deaths in both the STV and B265 progeny trials and frequency of deaths in the VOLMR progeny trial respectively (Tables 5.6 and 5.7).

Given that most regressions were non-significant, an ANOVA test was performed to test if there was a difference among geographic groups and a histogram was used to determine if the distribution of data based on 25cm height classes was normal. The results are listed in Table 5.8 and Appendix 5.D.1, respectively. The variance within and among geographic groups was found not to be significantly different in any of the 5 progeny trials with *P* values ranging from 0.863 to 0.3284. The within site variance was also high ranging from 77.2 to 94 % (Table 5.8). It can be noted that this model incorporated the F1 height-5 values from crosses within geographic areas of only 5 to 7 crosses. The distribution of data for the B265 progeny trial was found to be normal.

5.4 Discussion

The HFC was employed in this study of western hemlock as a possible predictor of fitness in first generation offspring. Microsatellites were the marker of choice for this study with their high polymorphism and small amount of DNA template required. Given that the microsatellites used have not been utilized in any population genetics based study to date, allelic frequency was calculated for each of the 7 loci (Appendix 5.A). The loci

generally showed bell shaped distributions indicating expected evolutionary patterns under the stepwise mutation model (Estoup and Cournet 1999). The stepwise mutation model, as compared with the infinite allele model, which describes mutation of microsatellite alleles by loss or gain of a single tandem repeat, is thought to best reflect the distribution of microsatellite alleles in a population (Estoup and Cournet 1999). Some loci, namely, EC10 and EE06, didn't show bell shaped distributions. These loci showed most of their alleles as rare (< 0.05) in frequency (Appendix 5.A). This may have an impact on the results given the low numbers used in the genetic distance/relatedness measures.

Genetic Relatedness ('r') between elites was found not to be significantly correlated in their progeny to height at year 5 at the 95% level in any of the 5 progeny trials (Table 5.3). Genetic Distance ('mean d^2') between elites was also found not to be significantly correlated in their progeny to height at year 5 in 3 of 5 progeny trials (Table 5.4). The MKL and STV progeny trials were found to be significantly correlated at 95% with *P* values of 0.0264 and 0.0208, respectively (Table 5.4). The two significant cases, MKL and STV, both showed negative slope indicating that as genetic distance increase, height decreases. This is indicative of outbreeding depression, which is generally attributed to the disruption of local adaptation or co-adapted groups of genes (Templeton 1986). Interestingly these progeny trails are located at the northern end of the distribution of trials (Figure 5.2). They therefore may endure more extreme environmental stress and as a consequence better express any depression. It can be noted that although there is significance the R^2 (the coefficient of determination) values are quite low with 0.035 and 0.0.038 for both regressions, respectively (Table 5.3). This indicates that the relationship, though significant, is not very strong.

The low R^2 values obtained are consistent with the results of correlating the phenotypic data for all progenies (Table 5.3, 5.4). Neither coefficient resulted in a significant relationship when regressed against mean height at year 5 for all progeny trials. The *P* values of 0.896 for genetic relatedness and 0.197 for genetic distance indicate that even with the larger sample size no significant relationship exists. With no significant correlation identified between height at year 5 and both genetic distance and relatedness, a geographic distance correlation was then tested for congruency. The

relationship with height at year 5 was also not significant across all progeny trials (Figure 5.5). The *P* values ranging from 0.428 to 0.125 indicate no significant relationship between geographic distance between parents and fitness in their progeny based on the phenotypic trait, height at year 5 (Table 5.5).

In carrying out these analyses it became apparent that any correlation was quite dependent on the level of environmental variance for a given trait. The phenotypic variance was too great to result in statistically significant relationships. For instance, a single cross (441 x 581) from one progeny trial (B265) randomly selected showed a range in height for its 30 progeny ranging from 48 to 352 cm. The variance was also quite pronounced when comparing across all crosses with height at year 5 in progeny ranging from 27 to 433 cm. Appendix 5.D.1 illustrates the distribution of heights in progeny B265. Given the large distribution the data does appear to be normal. This study concentrated on phenotypic measurement of the progeny of crosses which inherently result in variation. The within site variability accounted for more than 77.2 % of the total variability in all of the 5 progeny trials (Table 5.8). This is thought to be a conclusion of extensive site heterogeneity. It is therefore hard to discern progeny variation from environmental variation. It may be that height is more influenced by additive factors than dominance by this point in the seedlings growth. This has been identified in some physiological studies with Western hemlock (Kuser and Ching 1981, Foster and Lester 1983).

This null result may be expected given the mechanisms that are hypothesized to conclude in HFC. Due to the neutral nature of SSR's, indirect processes, that is, correlations between neutral marker heterozygosity and heterozygosity at fitness loci elsewhere in the genome are required for HFC. Linkage or identity disequilibrium with quantitative trait loci are the mechanisms that are needed to have any predictive value using HFC (Charcosset and Essioux 1994). Identity disequilibrium occurs when there is partial inbreeding within a population as a result of historical inbreeding rates. Identity disequilibrium is not influenced by physical linkage and as a result fitness loci contribute the same to any marker concluding in a general effect (David 1998). In contrast linkage disequilibrium, the nonrandom association of alleles at different loci in gametes, has a local effect as it decreases with physical distance along a chromosome. Possible causes of

linkage equilibrium include small population sizes, bottlenecks and admixture (Hansson et al. 2004).

Both identity and linkage equilibrium are not expected given the reproductive characteristics of Western hemlock. As a result, reproductive characteristics may be a major aspect of the lack of correlation in heterozygosity and fitness. Given its wide pollen and seed dispersal and prolific seed production, Western hemlock has high gene flow and therefore one would not expect much identity disequilibrium. Also given its continuous range one wouldn't expect much genetic drift and therefore population structure. As a conclusion, extensive linkage disequilibrium is not expected either. Therefore any observed correlation would have to be the result of intrinsic heterozygote advantage or functional overdominance (Savolainen and Hedrick 1995). To date a number of studies have shown that it is very unlikely that functional overdominance would lead to HFC in panmictic populations (Strauss and Libby 1987; Houle 1989; Bush and Smouse 1991; Savolainen and Hedrick 1995). Given that SSR's are generally neutral markers, functional overdominance would not be detectable and would therefore not be factor when analyzing these results.

Sampling factors also become a major focus when trying to identify HFC. Each family was sampled throughout the large geographic distribution of Western hemlock. Selection from wild stands was based on visual indicators, mainly height and diameter (King et al. 1997). Further selection for elites used was the results of phenotypic performance in progeny trials. Given the relatively high selection pressure large genotypic distance is still expected. Therefore degrees of relatedness between elite parents may not have been close enough to result in inbreeding depression. Geographic sampling (i.e. wide crosses) may have been too large to results in any extensive genotypic structure (Figure 5.1).

Given the influence that both reproductive characteristics and sampling play in determining HFC, measures of genetic distance and relatedness were therefore regressed against geographic distance (Figures 5.6 and 5.7). The relationship was also quite week, with a R^2 of 0.03475 and 0.00359 for genetic relatedness and distance, respectively. Given the poor relationship the slopes did indicate the expected result of increasing genetic relatedness with increasing geographic distance. Further analysis was conducted

to test if heterosis existed in the parent families. This was done by identifying if there was a difference within and between the phenotypic variances of the offspring based on the geographic location of their parents. Results suggest that very little structure exists in the Western hemlock parents based on geography (Table 5.8). This lack in relationship can be attributed to reproductive factors such as large seed dispersal and continuous range that has led to very little structuring.

Sample size is also a major factor when trying to identify a HFC (David 1998). When sample size of parental lines is limited, variation of genetic distance is restricted (Arcade et al. 1996). Therefore any correlation is restricted to a smaller range and becomes more difficult to significantly portray. Even when the HFC is 0 an average value of 1/(sample size)-1 is expected by chance (David 1998). David (1998) concludes that thousands of individuals are needed to result in the order of magnitude that is required to discern signal over noise.

One factor that also may influence a HFC is the effect of outbreeding depression. Heterosis has been shown in some annual plants to increase with genetic divergence, then decrease at higher levels of divergence (Waser and Price 1989). This response is said to be a product of the balance between inbreeding depression and outbreeding depression (Lynch 1991). Outbreeding depression was therefore tested by looking at the correlation between frequency of deaths (mortality rate), forks and frost damage in progeny compared with genetic relatedness and distance in the parents (Figures 5.8 and 5.9). The relationship between genetic distance and relatedness and presence of both forks and frost in the VOLMR progeny trial was not very strong (Tables 5.6 and 5.7). The relationship between frequency of deaths and genetic distance (mean d^2) was also quite weak across all progeny trials (Table 5.7). In contrast, the genetic relatedness vs. frequency of deaths was correlated (Table 5.6). The most northwestern progeny trials MKL and TLU showed a significant relationship with P values of 0.0275 and 0.0007 respectively and R^2 values of 0.3333 and 0.4975, respectively (Figure 5.2 and Table 5.6). Both TLU and MKL showed a slightly negative slope indicating that as genetic relatedness increases, frequency of deaths decreases. This therefore may indicate that some outbreeding depression is occurring in the most northwestern progeny trials. This result could be the

product of more stressful environmental variables localized to the more northwestern progeny trials.

In conclusion genetic distance has proven not to be a good predictor of heterosis in *Tsuga heterophylla*. This may be more indicative of the lack of structure that exists throughout this species range as very little heterosis was shown to exist. This conclusion is consistent with the many null result papers that have been published to date (eg. Houle 1989; Booth et al. 1990; Elliot & Pierce 1992; Whitlock 1993; Vaillancourt 1995; Savolainen and Hedrick 1995). The heterozygosity fitness correlation may be more suitable as predictor of phenotypic performance in more structured and heavily selected population. This is consistent with the few significant relationships that have been identified in trees (*Populus tremuloides*, Grant 1980; *Astrocaryum mexicanum*, Eguiarte et. al. 1992).

Extensive linkage disequilibrium is not expected in populations with large population sizes, high gene flow and very little inbreeding. Therefore HFC is expected as being the result of intrinsic heterozygote advantage or overdominance (Savolainen and Hedrick 1995). A large comprehensive study using *Drosophila melanogaster* identified no relationship and their continued review lead to a conclusion that overdominance was a very unlikely explanation for HFC (Houle 1989). Therefore a strong HFC is more likely to be the result of dominant effects and therefore would not be useful in panmictic populations such as Western hemlock.

5.5 References

Arcade, A., P. Faivre-Rampant, B. Le Guerroue, L. E. Paques, and D. Prat. 1996. Heterozygosity and hybrid performance in Larch. Theor. App. Genet. **93**: 1271-1281.

Booth, C. L., D. S. Woodruff and S. J. Gould. 1990. Lack of significant associations between allozyme heterozygosity and phenotypic traits in the land snail Cerion. Evolution. **44**: 210-213.

Bush, R. M. and P. E. Smouse. 1991. The impact of electrophoretic genotype on life history traits in *Pinus taeda*. Evolution **45**: 481-498.

Charcosset A. and L. Essioux. 1994. The effects of population structure on the relationship between heterosis and heterozygosity at marker loci. Theor. Appl. Genet. **89**: 336-343.

Coulson, T.N., J. M. Pemberton, S. D. Albon, M. Beaumont, T. C. Marshall, J. Slate, F. E. Guiness and T. H. Clutton-Brock. 1998. Microsatellites reveal heterosis in red deer. Proc. R. Soc. Lond. B **265**: 489-495.

David, P. 1998. Heterozygosity-fitness correlation: new perspectives on old problems. Heredity **80**: 531-537.

Elliot, A. C. and A. Pierce. 1992. Size, growth rate, and multiple-locus heterozygosity in the land snail (*Otala lacteal*). J. Hered. **83**: 270-274.

Estoup, E. and J. M. Cornuet. 1999. Microsatellite evolution: inferences from population data. In. Microsatellites: evolution and applications (ed. D. B. Goldstein and C. Schlotterer). pp. 49-65. New York NY. Oxford University Press.

Foster, G. S., and D. T. Lester. 1983. Fifth-year height variation in western hemlock open-pollinated families growing on four test sites. Canadian Journal Forest Research **13**:251 - 256.

Hansson, B., H. Westerdahl, D. Hasselquist, M. Akesson and S. Bensch. 2004. Does linkage disequilibrium generate heterozygosity-fitness correlations in great reed warblers. Evolution **58**: 870-879.

Hardy, O. J. and X. Vekemans. 2002. SPAGeDi : a versatile computer program to analyse spatial genetic structure at the individual or population levels. Molecular Ecology Notes **2**: 618-620.

Houle, D. 1989. Allozyme-associated heterosis in *Drosophila melanogaster*. Genetics **123**: 789-801.

Johnson, J., B. St.Clair and S. Lipow. 2001. Genetic conservation in applied tree breeding programs. In: Proceedings ITTO conference on in situ and ex situ conservation of commercial tropical trees pp. 215-230.

King, J. N., C. Cartwright, and D. Cress. 1997. Western hemlock tree improvment: selection of P-1 parents. B.C. Ministry of Forests, Victoria B.C.

Kuser, J. E., and K. K. Ching. 1981. Provenance variation in seed weight, cotyledon number, and growth rate of western hemlock seedlings. Forest Science **26**:463 - 470.

Lynch, M. 1991. The genetic interpretation of inbreeding depression and outbreeding depression. Evolution **45**: 622-629.

Lynch, M. and K. Ritland. 1999. Estimation of pairwise relatedness with molecular markers. Genetics **152**: 1753-1766.

Microsoft 2002. Microsoft Excel SP3. Microsoft Corporation. Seattle Wa.

Packee, E. C. 1990. Western hemlock, Tsuga heterophylla (Raf.) Sarg. In Silvics of North America. Vol 1. Conifers (ed. R. M. Burns and B. H. Honkala). U.S. Dep. Agric. Handb. No 654. pp. 613-622.

Ritland, K. 1996. Estimators for pairwise relatedness and inbreeding coefficients Genetical Research **67**: 175 – 186.

Ritland, K. 2004. MARK : A computer program to measure pairwise relatedness and individual inbreeding analysis. http://genetics.forestry.ubc.ca/ritland/programs.html.

SAS Institute Inc. 2001. SAS Release 8.02 TS Level 02MO. Cary, NC, USA.

Savolainen, O., and P. Hedrick. 1995. Heterozygosity and Fitness: No Association in Scots Pine. Genetics **140**: 755-766.

Shull, G. H. 1952. Beginnings of the heterosis concept. Pages 14-48 *in* J. W. Gowen, editor. Heterosis. Iowa State College Press, New York.

SigmaPlot. 2002. Windows version 8.02. SPSS inc. Chicago IL.

Strauss, S. H., and W. J. Libby. 1987. Allozyme heterosis in radiate pine is poorly explained by overdominance. Am. Nat. **130**: 879-890.

Templeton, A. R. 1986. Coadaptation and outbreeding depression. In Conservation biology. The science of scarcity and diversity (ed. M.E. Soule), pp. 105-116. Sunderland, MA: Sinauer Associates.

Vaillancourt, R. E., B. M. Potts, M. Watson, P. W. Volker, G. R. Hodge, J. B. Reid, and A. K. West. 1995. Detection and prediction of heterosis in *eucalyptus globulus*. Forest Genetics **2**: 11-19.

Waser, N. M., and M. V. Price. 1989. Optimal outcrossing *Ipomopsis aggregata*: seed set

and offspring fitness. Evolution **43**: 1097-1109.

Webber, J. E. 2000. Western hemlock: a manual for tree improvment seed production. Working paper 44, B.C. Ministry of Forests, Victoria B.C.

Wellman, H., E. Pritchard, A. Benowicz, D. Ally and C. Ritland. 2003. Microsatellite markers in western hemlock [Tsuga heterophylla (Raf.) Sarg]. Molecular Ecology Notes 3, **4**: 592-594

Whitlock, M. 1993. Lack of correlation between heterozygosity and fitness in forked fungus beetles. Heredity. **70**: 574-581.

Wright, S. 1921. Systems of mating II. The effects of inbreeding on the genetic composition of a population. Genetics **6**: 124 - 143.

5.6 Tables and Figures

Table 5.1. Origin, breeding value and code for each elite *Tsuga heterophylla* family.

#	Family	Geographic Origin	elev. (ft)	Lat.	Long.	Breeding Value
1	31	Head Bay	30	49 47	126 29	20
2	41	Kaouk River	30	50 05	126 59	16
3	441	Frederick Arm	100	50 29	125 15	22
4	459	St Vincents	330	49 52	124 06	25
5	687	Beaver Cove	150	50 35	126 55	17
6	1004	Eve River	130	50 24	126 14	10
7	2020	Wentworth Lake	110	47 59	124 33	11
8	2156	Beaver Lake	700	48 07	124 15	9
9	2157	Sappho	440	48 04	124 16	9
10	2166	Forks	250	47 59	124 23	12
11	2232	Elk Creek	400	47 48	124 12	9
12	2360	Knob Point B.C.	250			14
13	4161	Bernard Creek	300	47 02	123 57	14
14	4191	Burnt Hill	520	47 20	123 55	12
15	4230	Dekay Road	240	47 04	123 57	12
16	4296	Rehab Test Site	120	47 05	123 60	11
17	4492	Humptulips Gaurd Station		47 19	123 46	14
18	5030	Macafee Hill	360	47 20	124 03	10
19	6574	Shweeash	600	46 01	123 50	13
20	6581					8
21	6828	Bean Creek				13
22	6847	Salmon Creek				8
23	6850	Salmon Creek	350	46 24	123 42	10
24	6610	Fishawk Creek	900	45 59	123 39	9
25	7103	Buzzard Butte	1,280	45 17	123 54	11
26	7111	Cougar Mountain	1,200	44 58	123 51	14
27	7130	Sand Lake	1,320	45 16	123 55	7
28	7166	East Beaver Creek	1,920	45 20	123 45	8
29	7175	Three Rocks	450	45 03	123 56	9
30	7254	North Creek Campground	1,000	44 53	123 54	9

Table 5.2. Number of Alleles and expected heterozygosity for each locus.

Locus	# alleles	He
AC07	10	0.8429
EC02	10	0.8345
EC10	16	0.9311
EC12	23	0.9503
EE06	23	0.9452
EE10	13	0.8983
EE12	12	0.878

Table 5.3. Coefficients related to Genetic Relatedness ('r') between top elites vs. mean height at age 5 among 30 progeny that were the result of crosses between the top elites organized into each progeny trial shown in Figure 5.3.

Coefficients	B 265	MKL	STV	TLU	VOLMR	All progeny's
b[0]	263.106	240.4257	198.0409506	239.8992	250.8941	238.0321
b[1] (slope)	3.628379	-10.14585	-0.12507906	-4.058257	12.05011	-1.8518
R^2	0.0002846	0.0015532	0.0000003	0.0003065	0.0030063	0.00002483
P value	0.8477	0.6439	0.9953	0.8391	0.5215	0.8961

Table 5.4. Coefficients related to Genetic Distance ('mean d^2') between top elites vs. mean height at age 5 among 30 progeny that were the result of crosses between the top elites organized into each progeny trial shown in Figure 5.5.

Coefficients	B 265	MKL	STV	TLU	VOLMR	All progeny's
b[0]	263.0321	240.2439	197.4813	240.0465	250.4109	237.8931
b[1] (slope)	0.926057	-8.043285	-8.08807	0.140905	-1.033948	-3.0479
R^2	0.0006766	0.0352284	0.0378565	0.0000132	0.0007943	0.00242179
P value	0.7672	0.0264	0.0208	0.9664	0.7419	0.1970

Table 5.5. Coefficients related to geographic distance between top elites vs. mean height at age 5 among 30 progeny that were the result of crosses between the top elites organized into each progeny trial shown in Figure 5.7.

Coefficients	B 265	MKL	STV	TLU	VOLMR
b[0]	260.48966	238.86942	194.48829	236.91838	252.67806
b[1] (slope)	1.12590	0.85753	1.60567	1.40240	-0.98357
R^2	0.01122	0.00456	0.01689	0.01497	0.00820
P value	0.2268	0.4279	0.1246	0.1544	0.2890

Table 5.6. Trendline equations and R-squared values and *P* values for the frequency of deaths vs. genetic relatedness (Lynch and Ritland's 'r') histogram given as Figure 5.10.

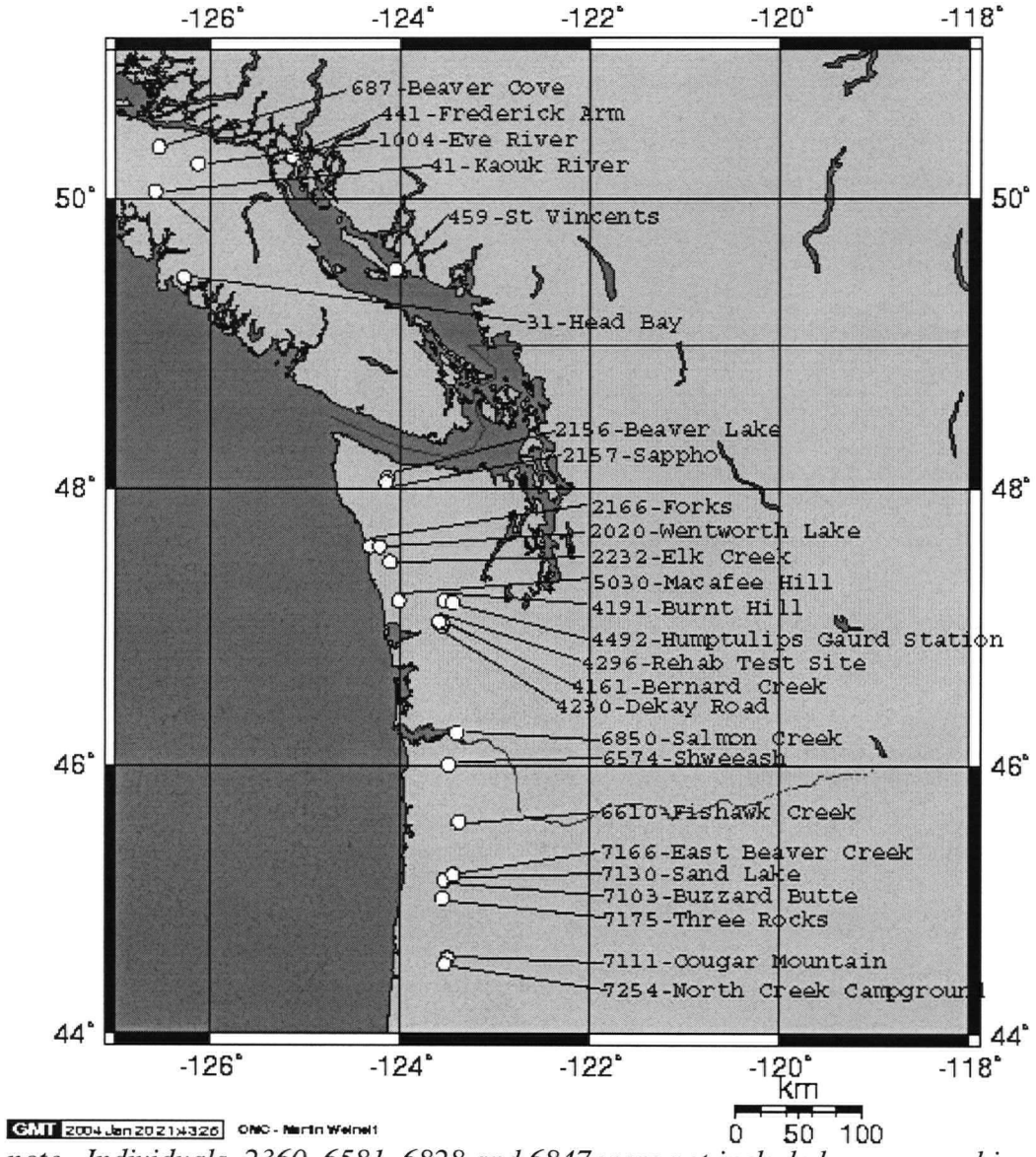
Progeny Trials	Equation	R^2	*P* value
B265	Y = -0.0016x + 0.1269	0.0556	0.8590
STV	Y = 0.001x + 0.0315	0.2045	0.0133
MKL	Y = -0.0023x + 0.0703	0.3333	0.0275
TLU	Y = -0.0018x + 0.0744	0.4975	0.0007
VOLMR	Y = -0.0018x + 0.051	0.129	0.5390
Fork (VOLMR)	Y = 0.0013x + 0.1404	0.0173	0.4949
Frost (VOLMR)	Y = 0.0005x + 0.0085	0.0346	0.7794

Table 5.7. Trendline equations and R-squared values and *P* values for the frequency of deaths vs. genetic distance (mean d^2) histogram given as Figure 5.11.

Progeny Trials	Equation	R^2	*P* value
B265	Y = 0.0009x + 0.113	0.0159	0.6773
STV	Y = 0.0009x + 0.0254	0.1228	0.6000
MKL	Y = -0.0003x + 0.0558	0.0126	0.1418
TLU	Y = -0.0005x + 0.0666	0.0126	0.3544
VOLMR	Y = -0.0015x + 0.0537	0.1433	0.1255
Fork (VOLMR)	Y = -0.0015x + 0.1614	0.0357	0.2983
Frost (VOLMR)	Y = 0.0002x + 0.0076	0.0257	0.4663

Table 5.8. Test of within and between group variation of western hemlock based on geographic origin using measurements of height at year 5 of F1's based on crosses of individuals within a provenance from 5 progeny trials.

Progeny	Model DF	Between Group Sum of Squares	Mean Square	F Value	Pr > F	Within to Total Variance Ratio (%)
B265	4	1521.892743	380.473186	1.23	0.3284	80.2
STV	4	660.46596	165.11649	0.32	0.863	94.0
MKL	4	2378.56461	594.64115	1.48	0.2461	77.2
TLU	4	1128.596347	282.149087	1.07	0.3996	82.4
VOLMR	4	1102.729377	275.682344	0.88	0.4913	85



note. Individuals, 2360, 6581, 6828 and 6847 were not included as geographic origin was not known.

Figure 5.1. Geographic origin of *Tsuga heterophylla* elite families.

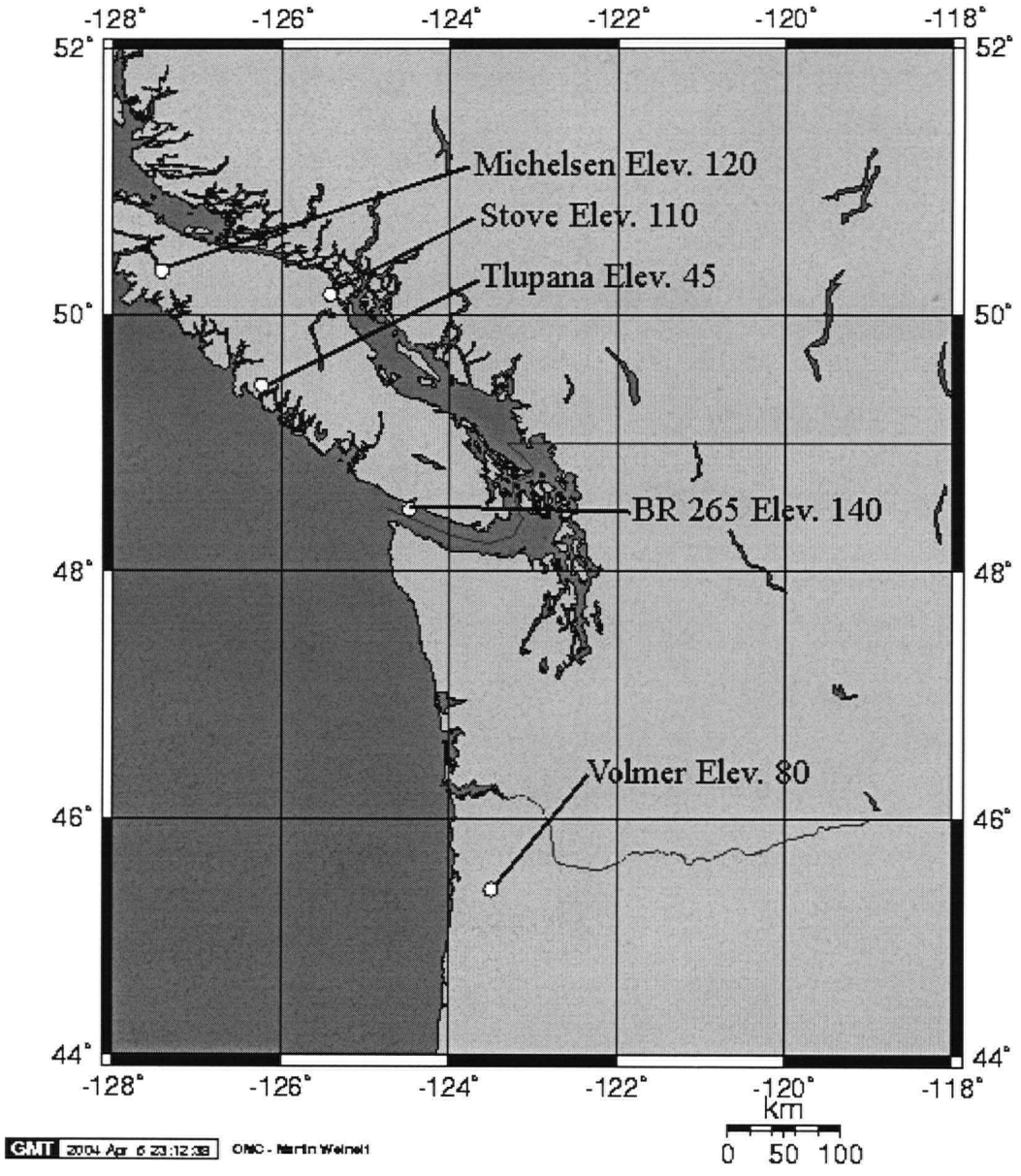


Figure 5.2. Geographic location of *Tsuga heterophylla* elite cross progeny trials.

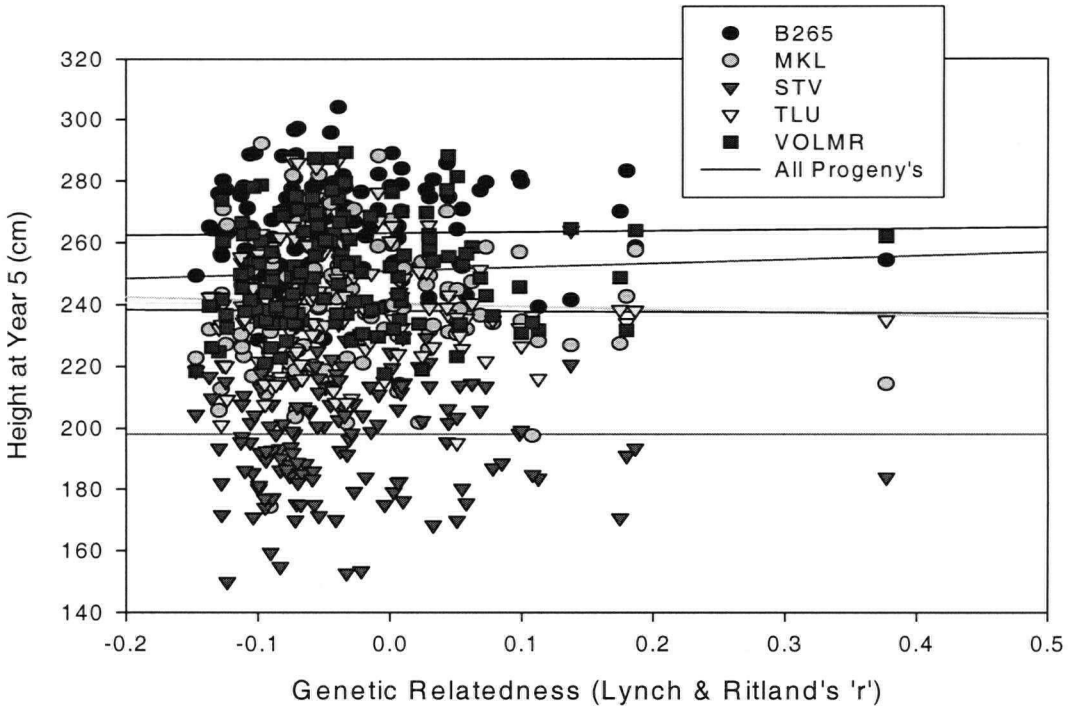


Figure 5.3. Genetic Relatedness ('r') between top elites vs. mean height at age 5 among 30 progeny that were the results of crosses between the top elites organized into each progeny trial and across all progeny trials.

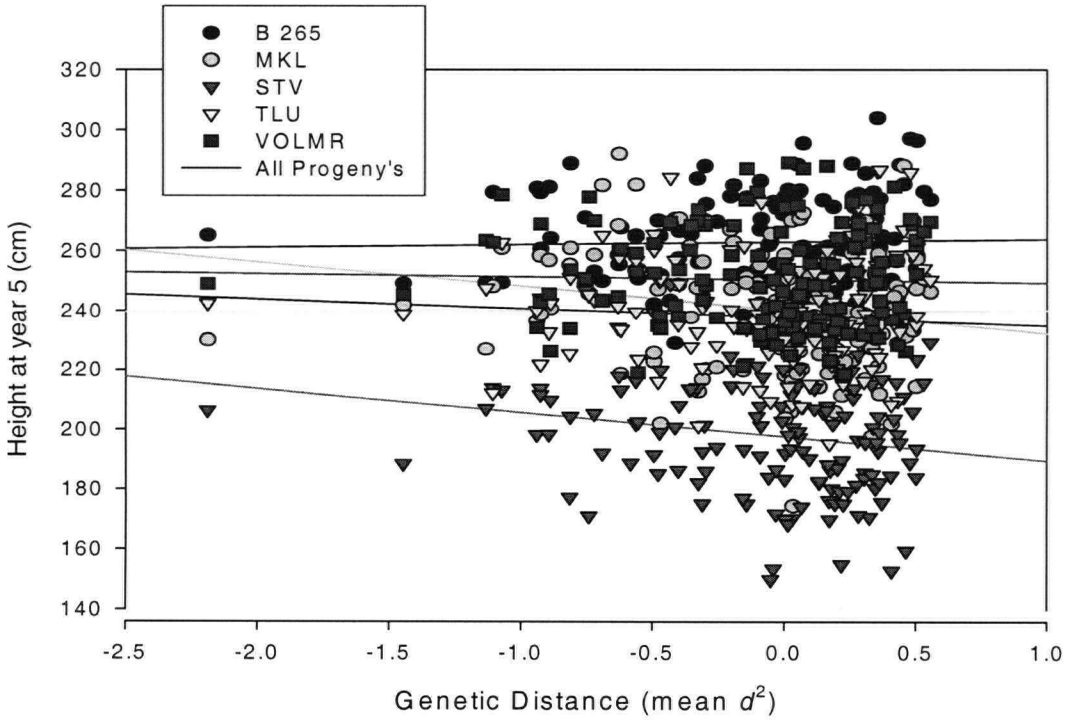


Figure 5.4. Genetic Distance (mean d^2) between top elites vs. mean height at age 5 among 30 progeny that were the results of crosses between the top elites organized into each progeny trial and across all progeny trials.

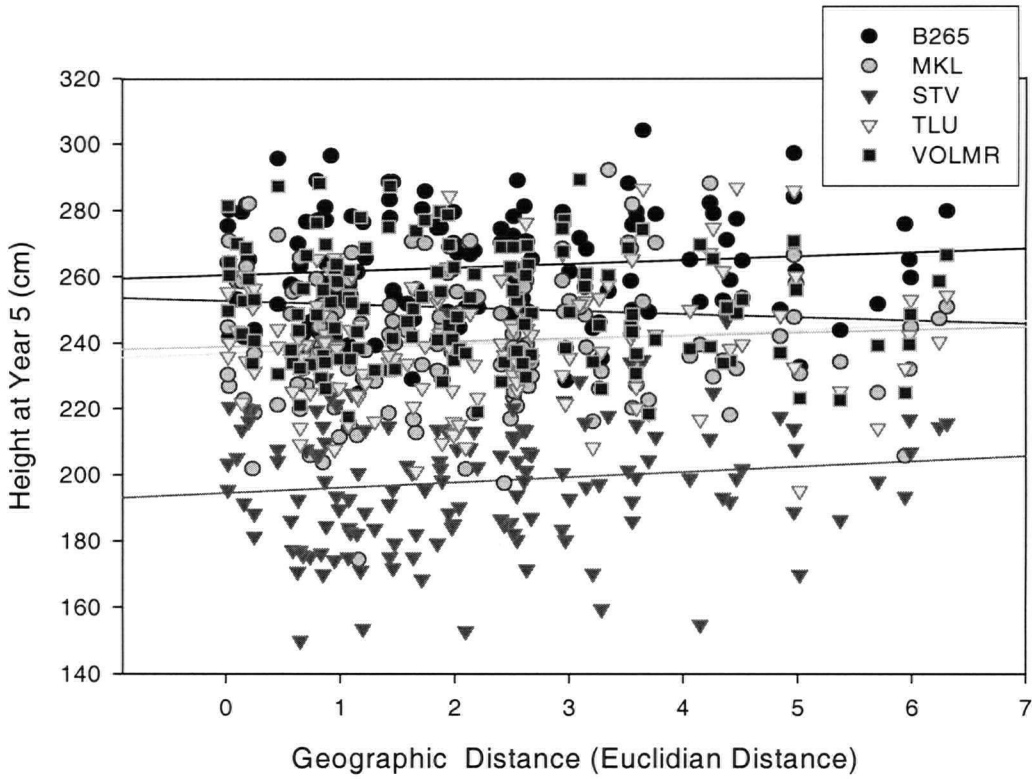


Figure 5.5. Geographic distance between top elites vs. mean height at age 5 among 30 progeny that were the results of crosses between the top elite organized into each progeny trial.

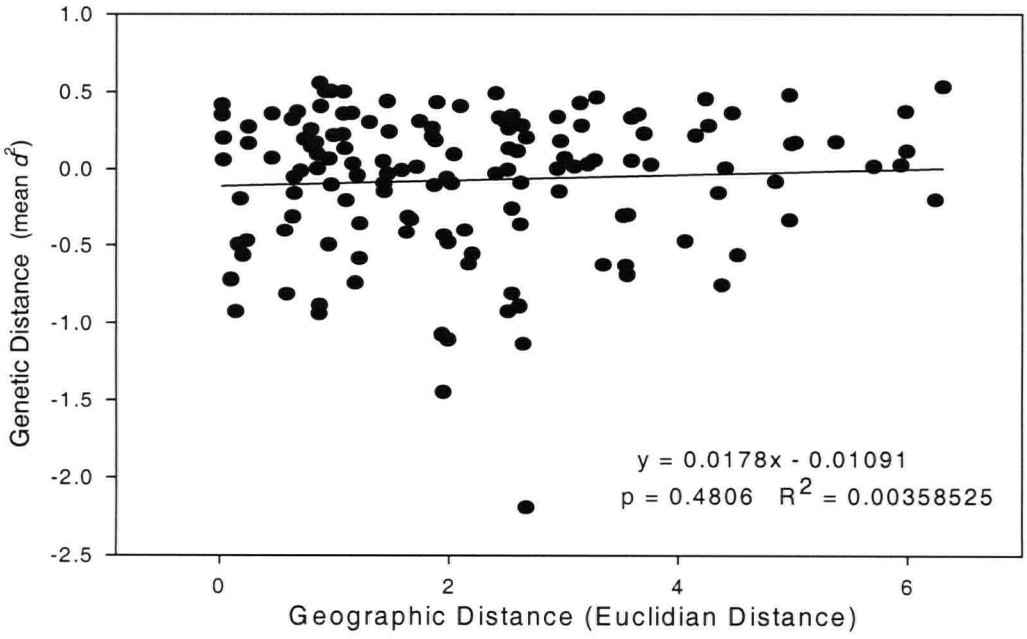


Figure 5.6. Genetic Distance (mean d^2) between elite crosses vs. geographic distance between the same elite crosses of *Tsuga heterophylla* families.

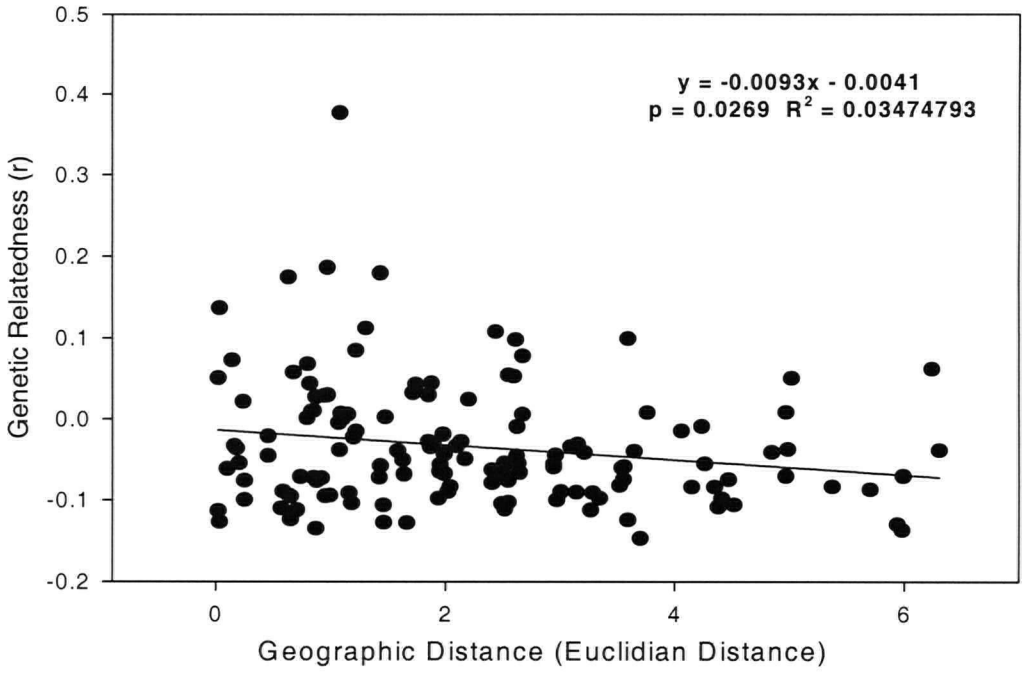


Figure 5.7. Genetic Relatedness ('r') between elite crosses vs. geographic distance between the same elite crosses of *Tsuga heterophylla* families.

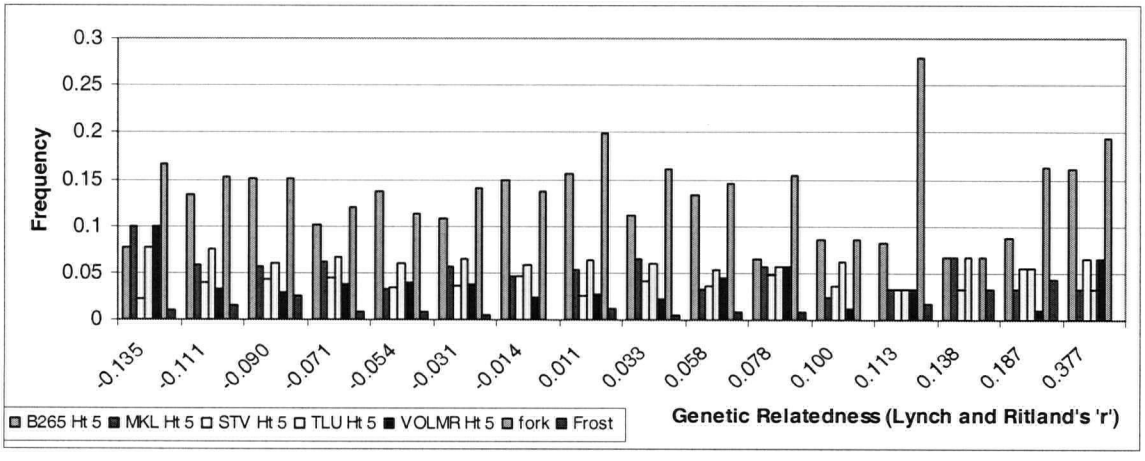


Figure 5.8. Frequency of Deaths of F1's in 5 progeny trials and frequency of frost and forks in the VOLMR trial across differing amounts of genetic relatedness (Lynch and Ritland's 'r') in their parents.

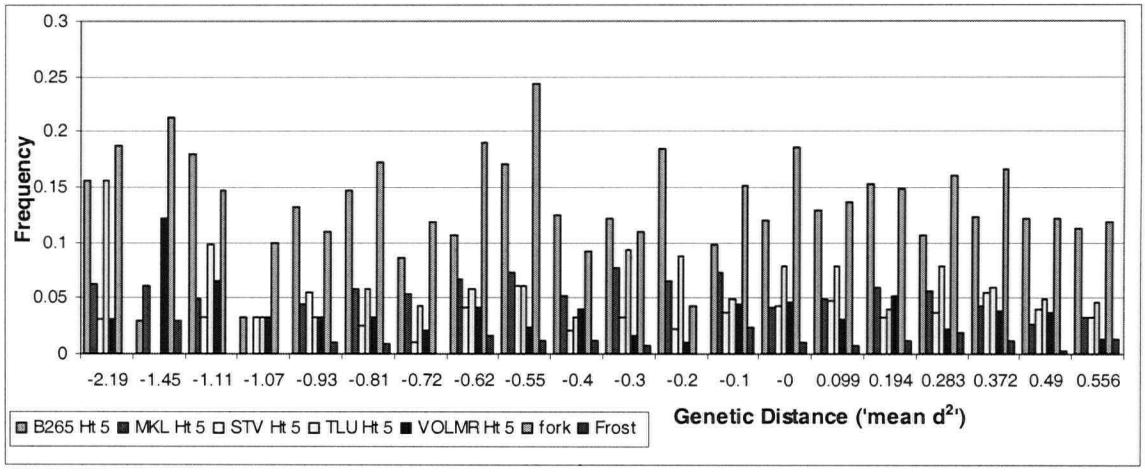
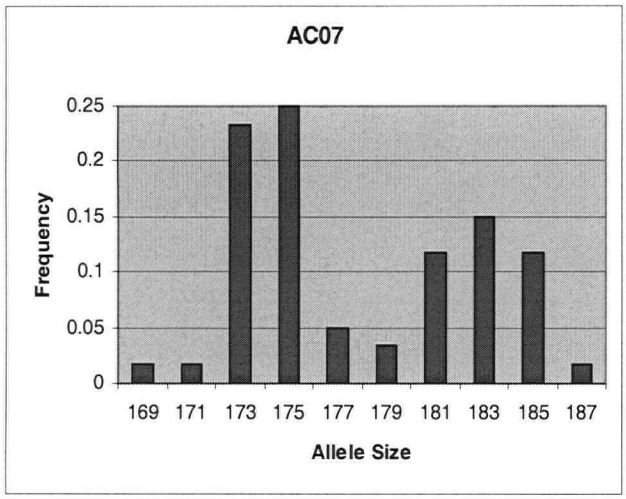


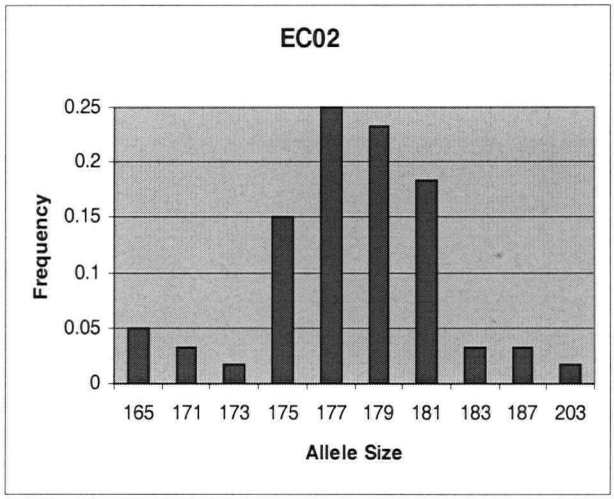
Figure 5.9. Frequency of Deaths of F1's in 5 progeny trials and frequency of frost and Forks in the VOLMR trial across differing amounts of genetic distance (mean d^2) in their parent.

5.7 Appendix

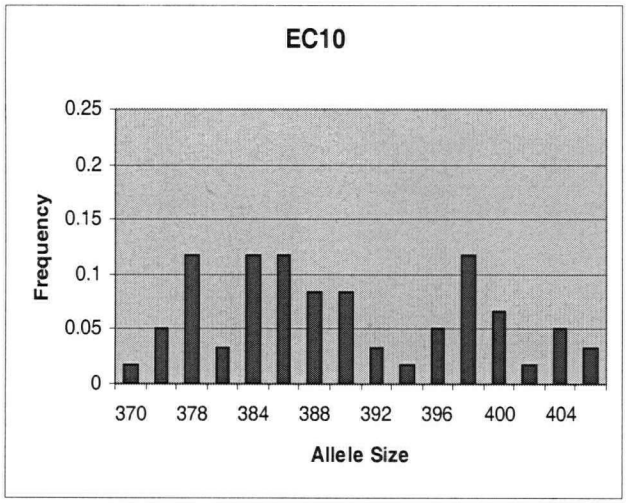
a.



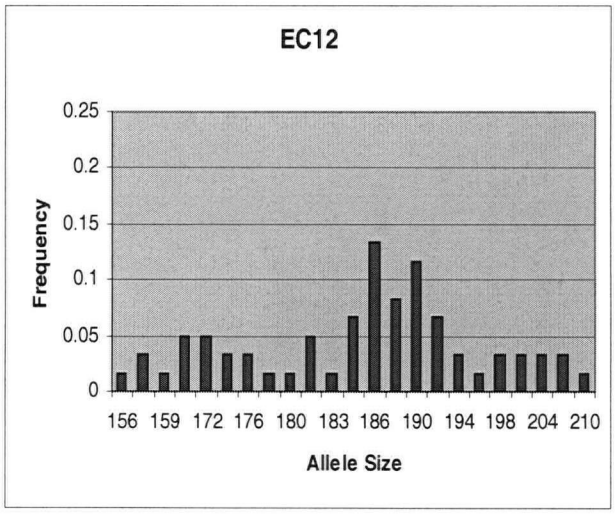
b.



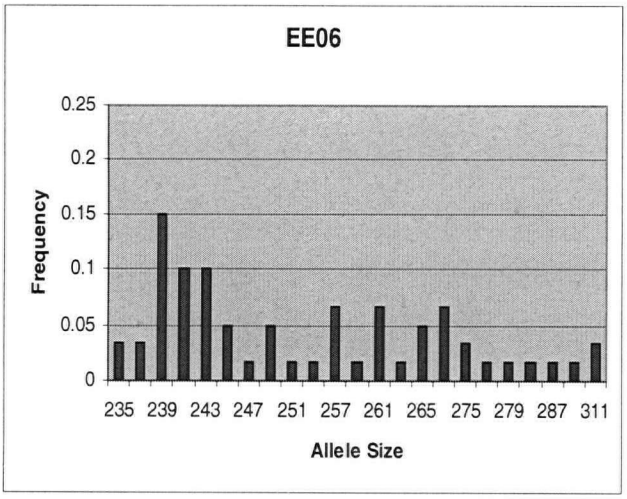
c.



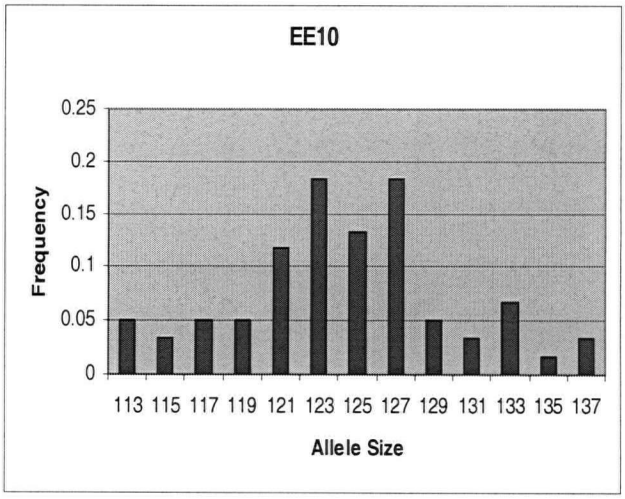
d.



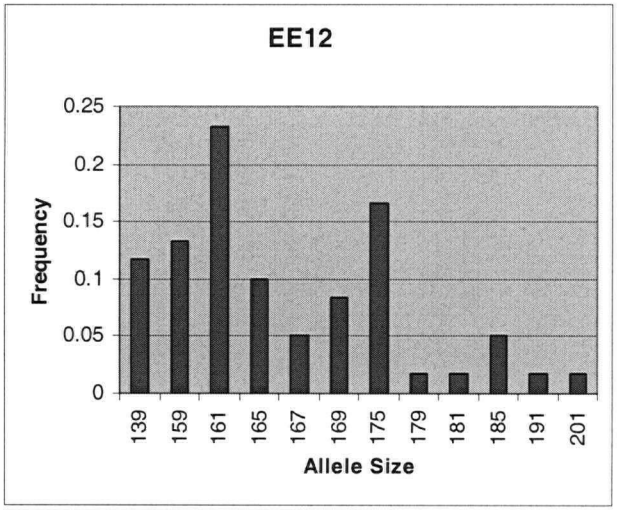
e.



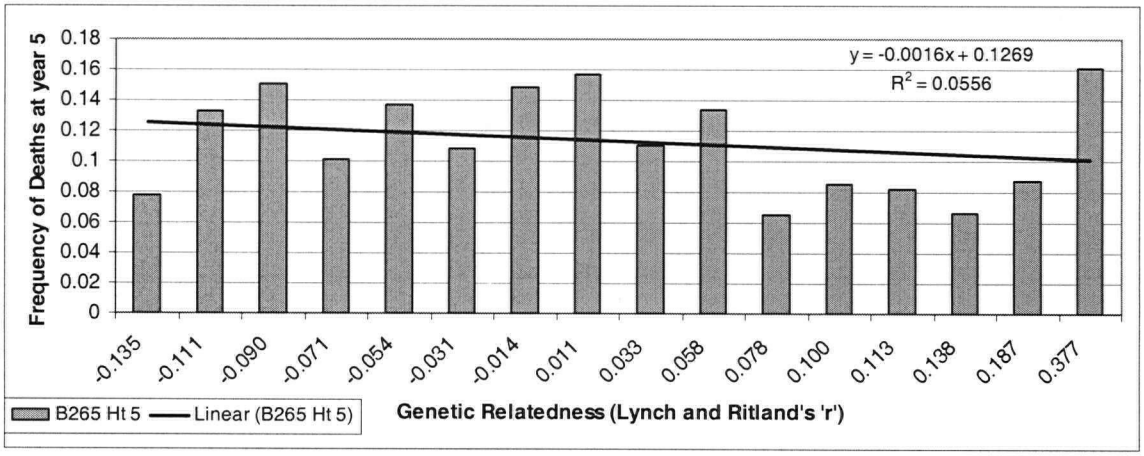
f.



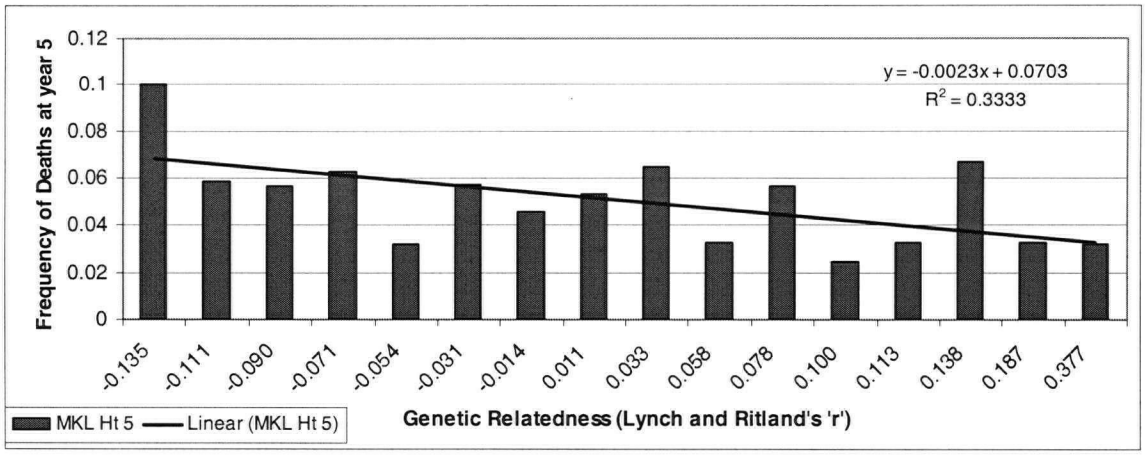
g.



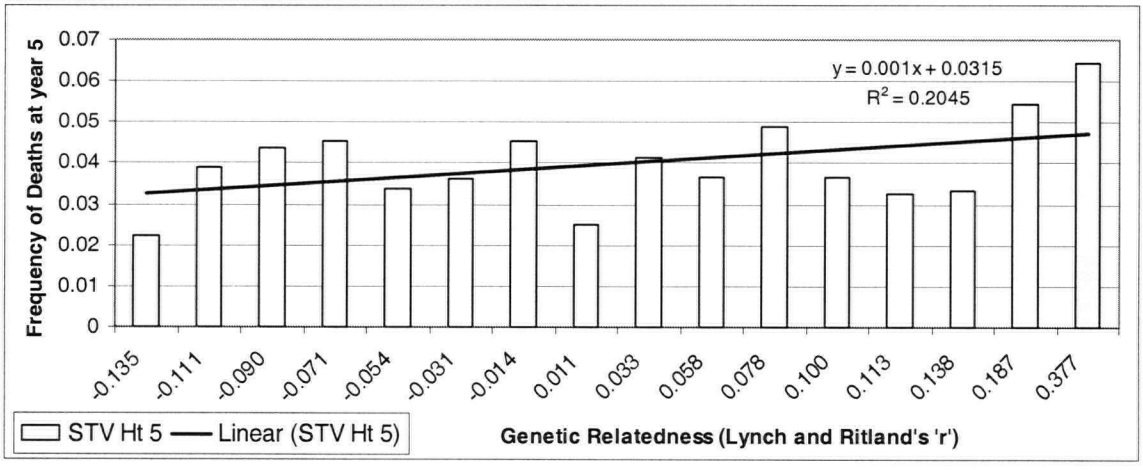
Appendix 5.A.1. Allele frequencies for all loci used in this study. (a) Loci AC07, (b) loci EC02, (c) loci EC10, (d) loci EC12, (e) loci EE06, (f) loci EE10 ,(g) loci EE12.



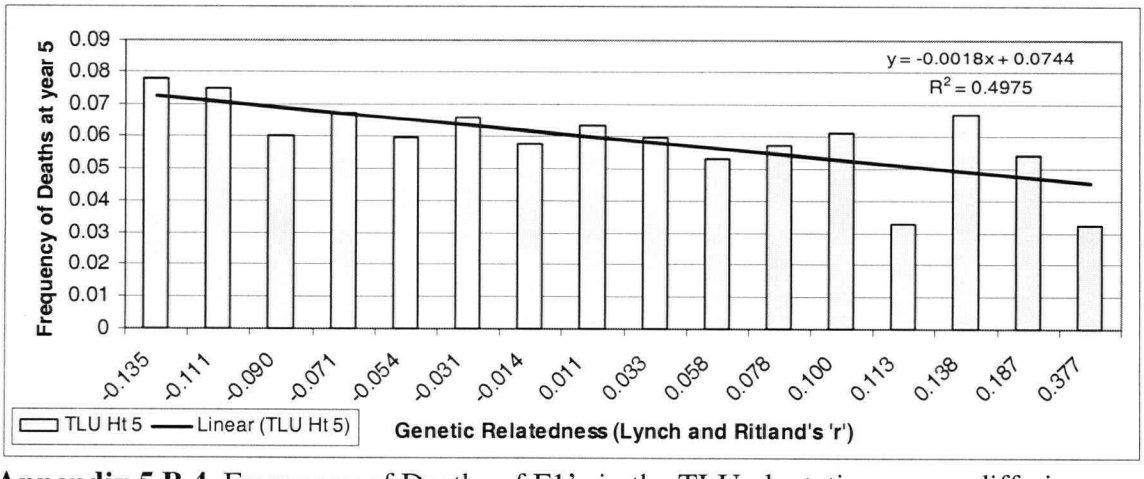
Appendix 5.B.1. Frequency of Deaths of F1's in the B265 plantation across differing amounts of genetic relatedness ('r') in their parents.



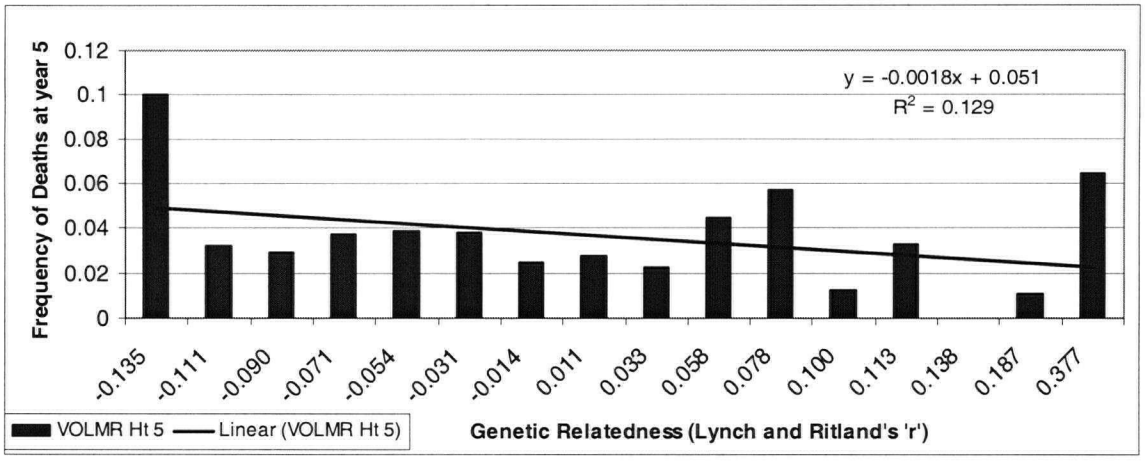
Appendix 5.B.2. Frequency of Deaths of F1's in the MKL plantation across differing amounts of genetic relatedness ('r') in their parents.



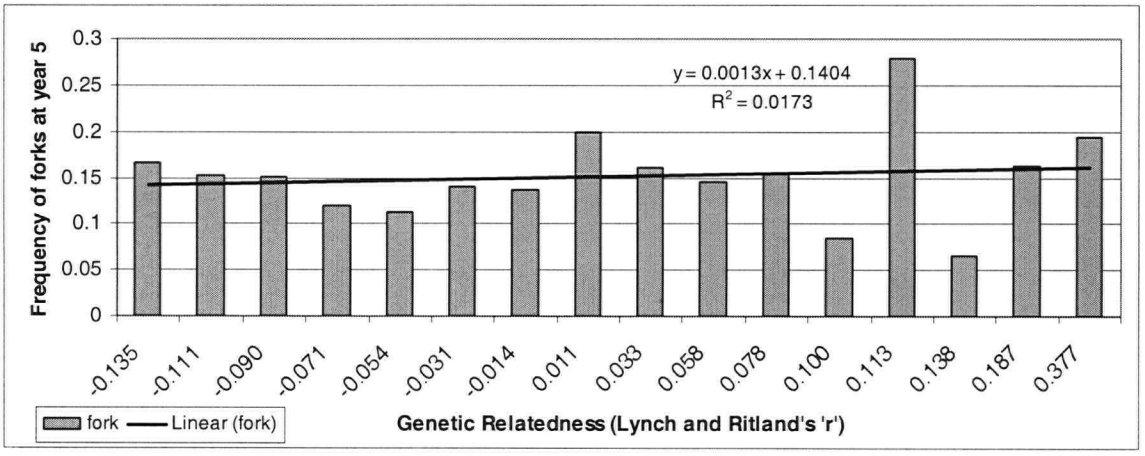
Appendix 5.B.3. Frequency of Deaths of F1's in the STV plantation across differing amounts of genetic relatedness ('r') in their parents.



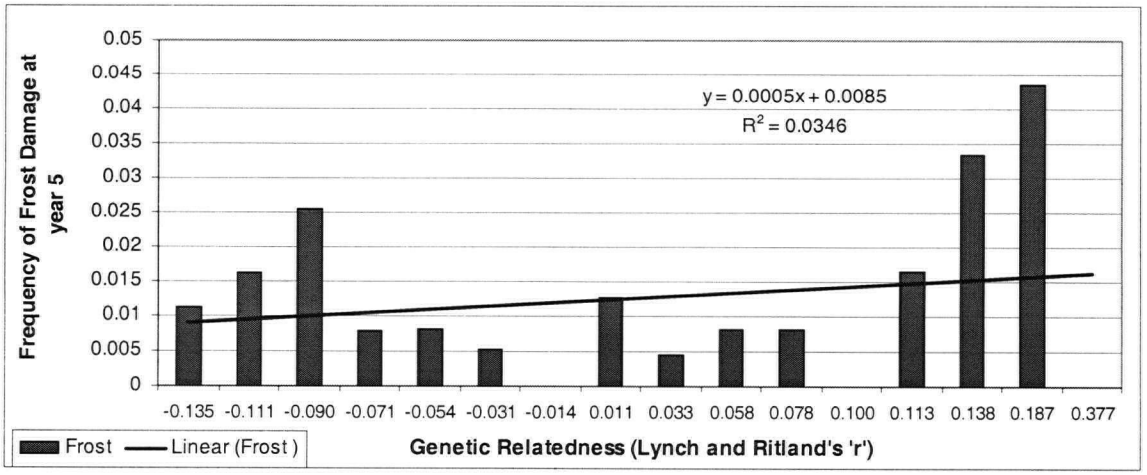
Appendix 5.B.4. Frequency of Deaths of F1's in the TLU plantation across differing amounts of genetic relatedness ('r') in their parents.



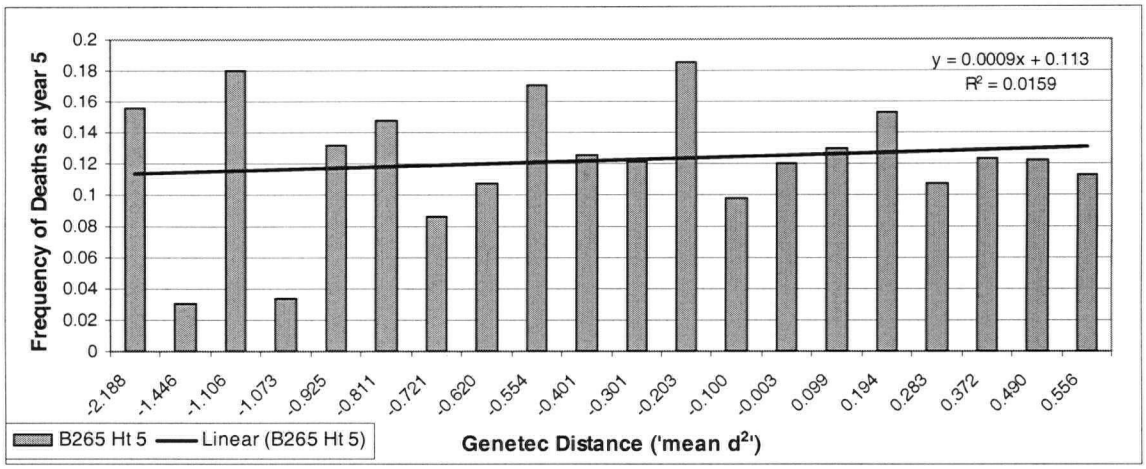
Appendix 5.B.5. Frequency of Deaths of F1's in the VOLMR plantation across differing amounts of genetic relatedness ('r') in their parents.



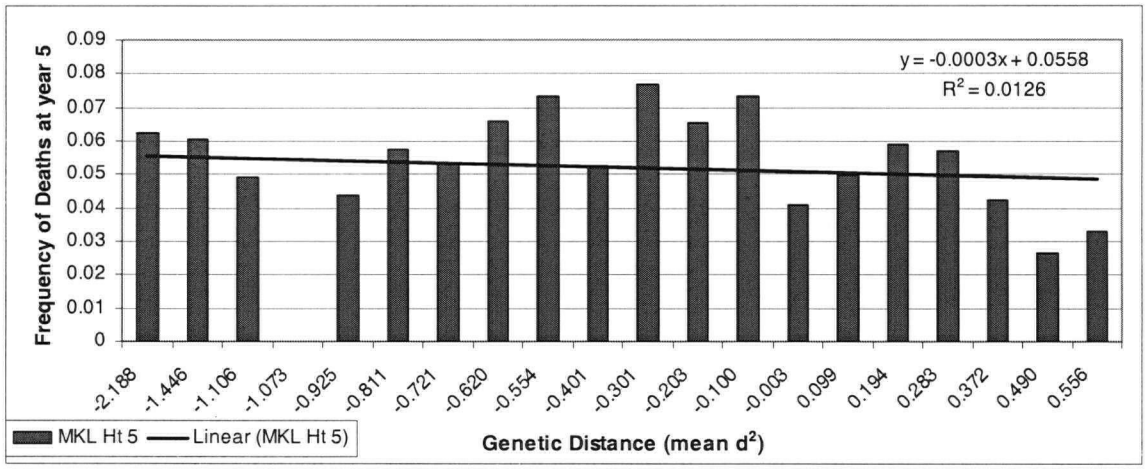
Appendix 5.B.6. Frequency of Forks in F1's in the VOLMR plantation across differing amounts of genetic relatedness ('r') in their parents.



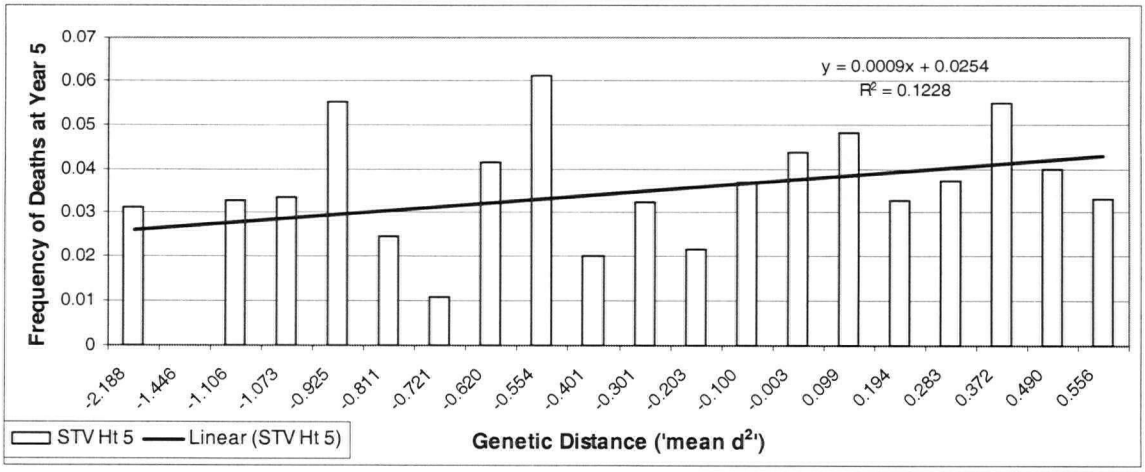
Appendix 5.B.7. Frequency of frost damage in F1's in the VOLMR plantation across differing amounts of genetic relatedness ('r') in their parents.



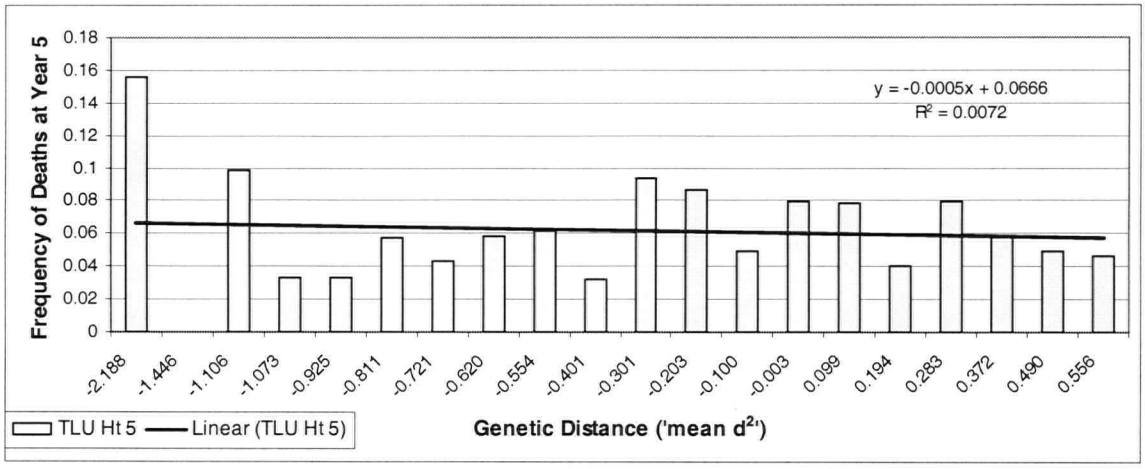
Appendix 5.C.1. Frequency of Deaths of F1's in the B265 plantation across differing amounts of genetic distance in (mean d^2) their parents.



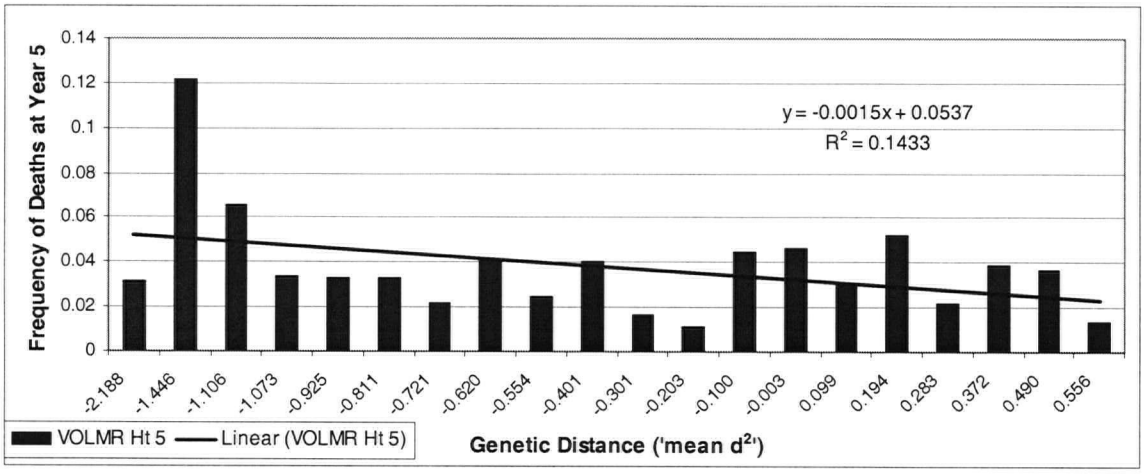
Appendix 5.C.2. Frequency of Deaths of F1's in the MKL plantation across differing amounts of genetic distance (mean d^2) in their parents.



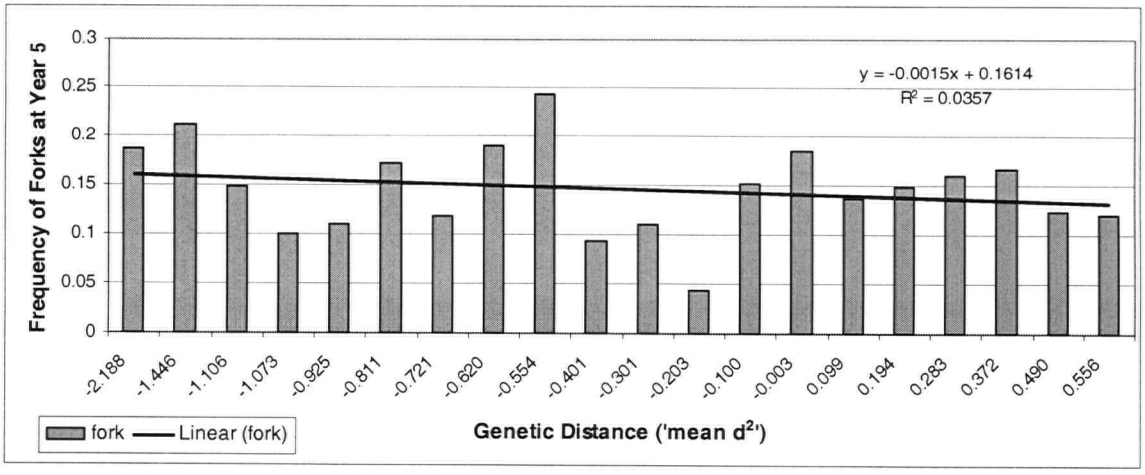
Appendix 5.C.3. Frequency of Deaths of F1's in the STV plantation across differing amounts of genetic distance (mean d^2) in their parents.



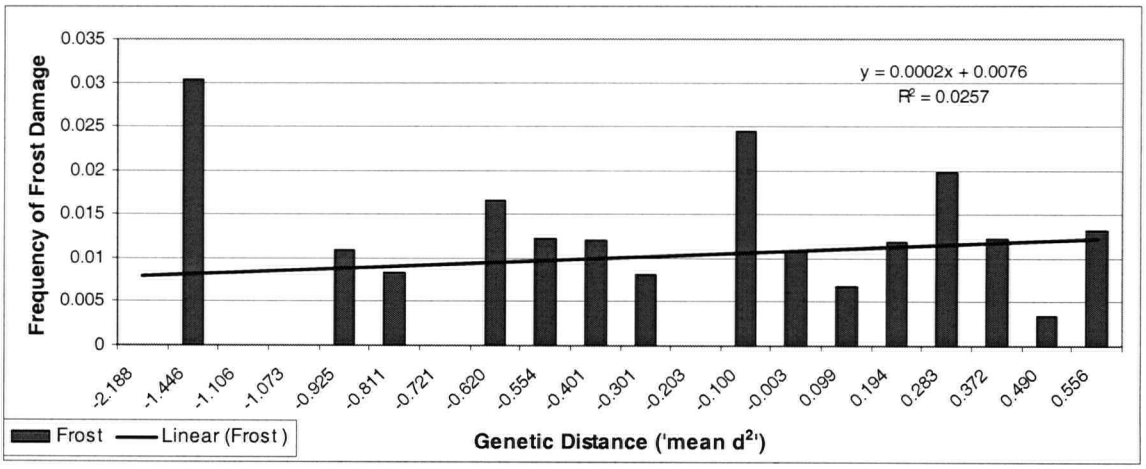
Appendix 5.C.4. Frequency of Deaths of F1's in the TLU plantation across differing amounts of genetic distance (mean d^2) in their parents.



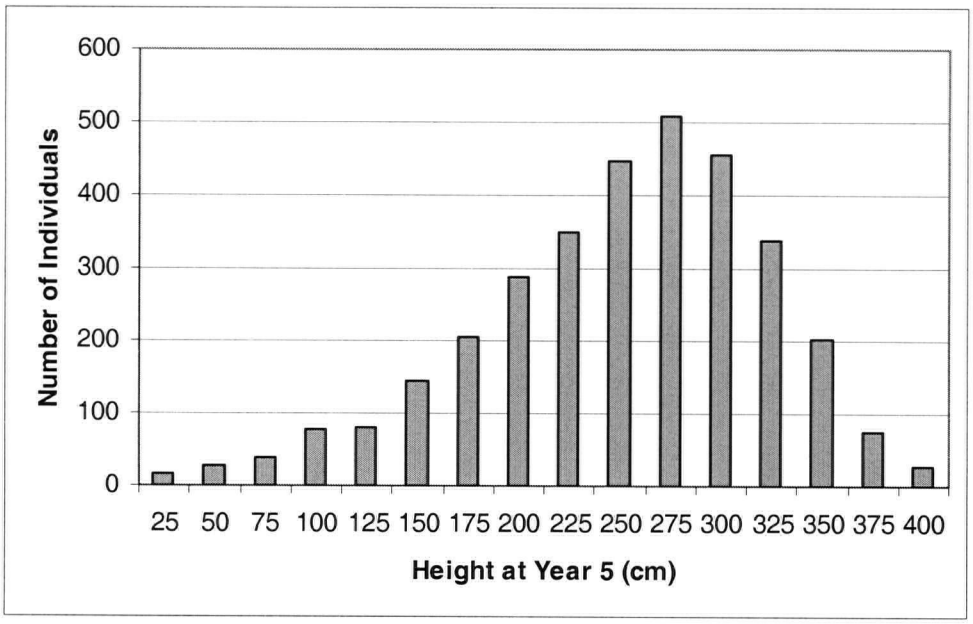
Appendix 5.C.5. Frequency of Deaths of F1's in the VOLMR plantation across differing amounts of genetic distance (mean d^2) in their parents.



Appendix 5.C.6. Frequency of Forks in F1's in the VOLMR plantation across differing amounts of genetic distance (mean d^2) in their parents.



Appendix 5.C.7. Frequency of frost damage in F1's in the VOLMR plantation across differing amounts of genetic distance (mean d^2) in their parents.



Appendix 5.D.1. Number of individual Western hemlock seedlings divided as based on height at year 5 in the B265 progeny trial.

CHAPTER 6: General Discussion and Conclusions

6.1 Main Findings

Rates of genetic diversity were inferred from allozyme allelic variation and used to compare differing seed crops that had originated from seed orchards under various seed production conditions with natural stands of western hemlock (*Tsuga heterophylla* (Raf.) Sarg). The main goal of chapters 2 and 3 were to measure the effectiveness of seed orchard management methods in maintaining the level of genetic variation observed in natural populations therefore minimizing inbreeding. A measure of gene diversity, H_t was found to be fairly high for both the natural populations and that of the various seed crops produced from seed orchards. Large amounts of genetic variation accords with the life history and patterns of morphological and physiological variability in Western hemlock. This trend was also noted in terms of H_o (observed heterozygosity) when supplemental mass pollination (SMP) was used as a pollen management option for seed orchards. The SMP technique was found to maximize both H_t and H_o. This is expected as SMP has been shown to minimize self-fertilization, increase the genetic gain through the introduction of desirable genotypes and improve parental balance. These results indicate that high selection intensity can be incorporated into breeding programs while still maintaining genetic diversity. An interesting finding was the loss of 8 private alleles in the seed orchard populations. This is most likely the result of different sampling ranges. As a conclusion homogenization of allelic diversity, as a result of genetic drift within the orchards has not been shown to lead to less genetic diversity, at least in first generation seed orchards.

Results of the inbreeding analysis have suggested that when compared with natural populations, inbreeding decreased in the wind pollinated unimproved and controlled full-sib seed crops of first generation seed orchards. This is reflected in F_{IT}, the total inbreeding which was shown to be smaller in the seed crops produced from the unimproved, SMP and full-sib seed orchards than in the natural populations. This is expected given the fertilization advantage of applied pollen compared with that of random natural pollen. In terms of genetic structure, F_{ST} was found to decrease

significantly when seed were produced under SMP mating. Hence some measure of genetic diversity may have been compromised with the high selection intensity found in the SMP orchard treatment. Therefore, some trade offs must be realized, as genetic diversity will be compromised when the breeding pool is restricted.

Fifteen microsatellite markers were isolated from western hemlock genomic DNA and six of these markers were optimized for use in mountain hemlock [*Tsuga mertensiana* (Bong) Carr.]. The mean expected heterozygosity (H_e) was 0.88 and 0.89 for western and mountain hemlock, respectively. Allelic diversity was high for both hemlock species, ranging from 7 to 15 and 5 to 30 alleles per locus for western and mountain hemlock respectively. The allelic markers identified in this project will allow for better monitoring of changes in genetic variation as a result of the domestication process in western hemlock.

The microsatellite markers developed were then implemented in a test of HFC (heterozygosity-fitness correlation) in western hemlock elite families. In general no significant relationship was identified between the genetic distance/relatedness of top western hemlock families and the fitness of their progeny based on phenotypic indicators. One exception were the MKL and STV progeny trials which were found to be significantly correlated at 95% with *P* values of 0.0264 and 0.0208 respectively. They both showed negative slope indicating that as genetic distance increase, height decreases. This is indicative of the effects of outbreeding depression which is generally attributed to the disruption of local adaptation or co-adapted groups of genes (Templeton 1986). These progeny trails are located at the northern end of the distribution of trials and therefore may endure more extreme environmental stress and as a consequence better express any depression. It can be noted that although there is significance the R^2 (the coefficient of determination) values are quite low with 0.035 and 0.0.038 for both regressions respectively (Table 5.3).

The null result can be explained by looking at the mechanism that was expected to result in HFC, associative overdominance. Either identity or linkage disequilibrium are required to cause marker loci to correlate with heterozygosity at agent loci and result in overdominance. The wide pollen flow and seed dispersal as well as the species prolific seed production all have resulted in high gene flow and therefore very little identity

disequilibrium is expected. Also given its continuous range one wouldn't expect much genetic drift and therefore population structure. As a conclusion, extensive linkage disequilibrium is not expected either. Another reason that may also have also had an influence is the degree of relatedness between elite parents which may not have been close enough to result in any inbreeding depression (Vaillancourt et al. 1995). Due to the limited parental lines used, the sample size became a major factor as the distance variation was restricted (Arcade et al. 1996). Therefore any correlation is restricted to a smaller range and becomes more difficult to detect.

6.2 Implications for Breeding Strategies

A compromise between selection intensity and the maintenance of allelic diversity must be recognized to best manage the differing objectives of breeders. A certain level of selection intensity is required to result in substantial gain to justify breeding in a species with high natural regeneration while a large breeding population is needed to maintain genetic diversity. As a result differing seed crop management methods have been developed to maximize both of these objectives in production populations (seed orchards). The SMP technique resulted in the highest genetic diversity in the F1 generation while still maintaining enough selection intensity for greater gains. In contrast the full-sib seed production and crops produced under wind-pollination from unimproved first generation orchards lead to a lower diversity index when compared with natural populations. Therefore these finding suggest that the SMP technique best maintains genetic diversity and minimizes inbreeding while retaining the required selection intensity for genetic gains.

Genetic distance/relatedness has proven not to be a good predictor of heterosis in *Tsuga heterophylla*. This may be due to a lack of structure identified in Western hemlock populations. The heterozygosity-fitness correlation (HFC) could be a more suitable predictor of phenotypic performance in more structured and heavily selected breeding populations in species such as *Pinus radiata* and *Eucalyptus*. A strong HFC is more likely to be the result of dominant effects and therefore would not be useful in panmictic

populations such as Western hemlock. Given a more developed breeding population (eg. F3, F4) and therefore more closely related breeding pools, HFC may prove useful.

6.3 Recommendations for Further Research

In order to maintain a high level of genetic variability in intensive breeding programs, the dynamics of rare alleles in domesticated populations require further investigation. Genetic markers can provide sufficient power to identify trends. Isozymes may not be that suitable as they are well known to encompass a very restricted portion of the total genome of an organism. In a Douglas-fir study isozymes were shown to not fully measure the losses of variation that occur in the initial stages of domestication (El-Kassaby and Ritland 1996). The ideal markers would lie adjacent to loci controlling physiologically and ecologically important characters, and lacking knowledge of the locations of these genes, would at least require the use of numerous and highly polymorphic markers. This is the case with SSR's and given the development of Western hemlock specific primers, further evaluation of the domestication process is warranted.

HFC may prove applicable with a much greater parental pool. Sample sizes in the thousands have been defined as a requirement for a strong HFC (David 1998). This may prove unrealistic in most conifer breeding programs. A more developed breeding program in which individuals are more closely related may also result in a significant HFC. This may also not prove applicable in a system in which the maintenance of genetic diversity is a major objective. Further study into the application of HFC should only be initiated at further stages in selection (advanced breeding populations), when the genetic pool has been greatly reduced and a significant amount of inbreeding has accumulated.

6.4 References

Arcade, A., P. Faivre-Rampant, B. Le Guerroue, L. E. Paques, and D. Prat. 1996. Heterozygosity and hybrid performance in Larch. Theor. App. Genet. **93**: 1271-1281.

David, P. 1998. Heterozygosity-fitness correlation: new perspectives on old problems. Heredity **80**: 531-537.

El-Kassaby, Y.A. and K. Ritland. 1996. Impact of selection and breeding on the genetic diversity in Douglas-fir. Biodiversity and Conservation **5**: 795 – 813.

Templeton, A. R. 1986. Coadaptation and outbreeding depression. In Conservation biology. The science of scarcity and diversity (ed. M.E. Soule), pp. 105-116. Sunderland, MA: Sinauer Associates.

Vaillancourt, R. E., B. M. Potts, M. Watson, P. W. Volker, G. R. Hodge, J. B. Reid, and A. K. West. 1995. Detection and prediction of heterosis in *eucalyptus globulus*. Forest Genetics **2**: 11-19.