Winter 1991

Molecular variation in Picea rubens and Picea mariana

Michael S. Bobola
University of New Hampshire, Durham

Order Number 9209580

Molecular variation in *Picea rubens* and *Picea mariana*

Bobola, Michael S., Ph.D.

University of New Hampshire, 1991

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

MOLECULAR VARIATION IN PICEA RUBENS AND PICEA MARIANA

by

Michael S. Bobola

B.A., University of Maine, Orono, 1987

Dissertation

Submitted to the University of New Hampshire

in Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

in

Biochemistry

December, 1991

This dissertation has been examined and approved.

Dissertation director
Anita S. Klein
Associate Professor of Biochemistry/Molecular Biology

Robert T. Eckert
Associate Professor of Natural Resources

Clyde L. Denis
Associate Professor of Biochemistry/Molecular Biology

John J. Collins
Assistant Professor of Biochemistry/Molecular Biology

Peter W. Garrett
Professor of Genetics

Yun-Tzu Kiang,
Professor of Genetics

11/22/91
Date

Acknowledgements

This dissertation has been brought to you in part by the genetic contributions of Patricia Bobola, maternal donor, and Stephen Bobola, paternal donor, thank you mom and dad. Special thanks to my family for their love and support over the years. A very special thank you to my wife Etienna Bachand for all her support and love during my time at UNH. Thanks to all my fellow lab mates and friends, both past and present, for their help and support. Special thanks to Leila, Mike, Yu, Denis, Tian, Tom, Dan, Jim, Phil, and Dirk for their help in and out of the lab. Thanks to Ding Bin-Bin, Tian Jingxiang, Lui Hai-yan, and Zang Xiaoming for their help in the final days. Many thanks to my dissertation committee for their guidance, support and unsolicited signatures. Thanks to Bob Eckert for all the help with data analysis and sample collections. Thanks to my dissertation advisor, Anita Klein, for all her help and guidance over the years.

Permission to collect samples from the Massabesic Experiment Station was granted by Dr. Peter Garrett (USFS), Durham NH. Dr. Peter Brym and John Carlson of University Computing and Information Services instructed me on the use of the SAS package. Dirk Smith provided the ITS 1 probe, identified the mitochondrial polymorphism and assisted with DNA isolations. Denis Guenette identified the chloroplast polymorphism and assisted with DNA isolations. Suzan Dziennis, Kim Hillenberg and Steve Gendreau assisted with DNA isolations and data collections. A special thank you to the Mt. Washington collection crew, it was an experience. Thanks to Karen Stapelfeldt for classifying the Mount Washington samples using a morphological index. Thanks to Dr. Thomas Laue for access to the densitometer for data analysis. The 18S and 26S rDNA soybean probes were obtained from Dr. Elizabeth Zimmer in the form

of plasmid pXBr 1 and pGmr 1. Two maize clones containing the mitochondrial 18S-5S rDNA and the 26S rDNA unit were obtained from Dr. C. S. Levings III. A petunia clone containing the chloroplast rDNA was obtained from Dr. David Neale. Red on Black control-cross hybrids were provided by Dr. James Hanover of Michigan State University. This work was supported in part by grants from the UNH Central University Research Fund (Michael S. Bobola), and New Hampshire Agricultural Experiment Station MacIntire Stennis Project MS27 (Anita S. Klein).

Table of Contents

List of Tables

List of Figures

Abstract

Molecular variation in Picea rubens and Picea mariana

by

Michael S. Bobola

University of New Hampshire, December, 1991

Restriction fragment length polymorphism (RFLP) variation was examined among samples from the entire range of red spruce (Picea rubens Sarg.), the eastern complex of black spruce (Picea mariana (Mill.) B.S.P.), control-cross red-on-black hybrids, and natural populations of red and black spruce. Within-species and population variation was examined. In addition an accurate species index capable of identifying red spruce, black spruce and hybridization between the two species was developed.

The nuclear rDNA repeat unit size in Picea ranged from a minimum of 32 kbp to greater than 40 kbp, two to three fold larger than the typical angiosperm rDNA unit. At a size greater than 32 kbp and a concentration averaging 1.2-1.3 X 10^4 copies per pg genomic DNA, the rDNA repeat constitutes approximately 4% of the spruce genome.

The rDNA repeat units were found to be polymorphic within an individual genome with up to five distinct rDNA repeat unit types (alleles) evident. The RFLPs observed in the rDNA repeat were not species specific; however, noticeable trends in internal allelic frequencies were noticed which were useful for between-species differentiation. One marker (EMW 4.35) displayed a significant relationship with

geographic origins and habitat suggesting that the observed between-species variation for this marker may be due to selection rather than the result of a true species difference.

Variation in the nuclear rDNA repeat could not accurately differentiate hybrids from black spruce. Additional markers were required to identify hybrids.

RFLPs were identified for the organelle genomes of red spruce, and black spruce. The organelle inheritance pattern was deduced using controlled-cross hybrids. Organelle markers were combined with allelic data from the nuclear rDNA repeat to derive a simple three character index capable of identifying red spruce, black spruce and hybridization between the two species.

Significant gene flow was observed between red and black spruce populations located at Head Harbor, Isle au Haut, Maine but not on Mount Washington or Mount Lafayette. These findings suggest that hybridization and introgression between red and black spruce is influenced by not only proximity of the two species but also by habitat type.

Introduction

Red spruce (Picea rubens Sarg.) is predominantly a New England species whose range extends from North Carolina, at high elevations in the mountains, to coastal Nova Scotia (Fowells 1965). Red spruce will grow on steep rocky slopes at high elevations (up to 1490 m), and is not tolerant of boggy or arid sites (Dallimoreet al. 1967).

Red spruce populations have been displaying increased mortality and exhibiting reduced growth rates in the northeastern American forests since the mid 1960's (Volgelmannet al. 1985; Scottet al. 1984). Dieback and mortality of red spruce has been observed as far south as North Carolina and is most obvious in high elevation forests (Johnson 1988). Dieback and mortality occurs in both young and mature stands as well as in both mixed and pure stands and on a variety of soils (Schier 1985). The decline in red spruce populations, starting in the early 1960's, corresponds to drought conditions in the range of red spruce (Johnson 1988). Regeneration and establishment of new spruce has failed to replace dead trees.

Decline symptoms observed in red spruce appear to be associated with winter injury damage (Friedlandet al. 1984). Similar declines have not been reported for black spruce (Picea mariana (Mill.) B.S.P.) or white spruce (Picea glauca (Moench) Voss), the only other spruce species sympatric with red spruce.

Red spruce has the lowest heterozygosity level of the northeastern spruces (Eckert 1989), supporting a hypothesis that low genetic diversity may be a factor in red spruce decline. Comparing 13 loci, red spruce had a relatively low level of average heterozygosity, 7.8%, as compared to 20.3% for black spruce and 22.8% for white spruce (Eckert 1989). The low genetic diversity in red spruce may be due to its restricted range

or to genetic bottlenecks resulting from glaciation events and/or over-logging during colonial settlement. Its occasional hybridization with black spruce, a cold site species, may afford opportunity for enhancing genetic variation and increasing cold tolerance in red spruce.

Hybridization is reported to occur between red and black spruce (Morgenstern and Farrar 1964; Morgensternet al. 1981; Gordon 1976; Manley and Ledig 1979; Berlynet al. 1990) which may lead to potential gene flow between the two species. It is, therefore, important to develop methods which can be used to accurately assess the degree of gene flow between these two species.

Taxonomy of Picea rubens and Picea mariana.

The only other spruce species which occur naturally in the red spruce range are white and black spruce. White spruce, more distantly related to red spruce (Smith and Klein in preparation), can be easily identified using morphological traits (Gordon 1976). Black spruce is considered to be more closely related to red spruce (Manley 1971; Gordon 1976; Smith and Klein in preparation). The two species have been sympatric extending back to the Pleistocene or earlier. Field studies suggest that the species are distinguishable on two extreme site types; red spruce predominant on upland sites, and black spruce in bogs (Manley 1972). While red and black spruce occupy different microecological niches, they occasionally overlap.

Hybrid seed between the two species may be produced by controlled pollination (Wright 1955; Fowleret al. 1970). The ability to produce hybrids, without complications, prompted some researchers to suggest that the taxa were not separate species (Stern and Roche 1974). Gordon (1976) determined that the germination rate of spruce hybrids is

exceptionally low, suggesting that the two parent species have undergone selection for reproductive isolation. Gordon's comprehensive study of Picea rubens (1976) makes a strong case that P. rubens and P. mariana are distinct species, not a species pair.

Morphological traits have been used to distinguish red and black spruce. Gordon (1976) employed factor analysis using 24 morphological characters to study variation within red and black spruce and assess the degree of hybridization between the two species. He concluded that the parent species were discrete but that there was considerable variation within each species. Putative hybrids accounted for only 0.4% of his large sample populations (10,000 individuals) and were judged to be very rare. Morphological characters with the strongest predictive values were reproductive tissues and new growth tissues: cone shape, color and size, bud color, 1st year twig color, etc. In field sampling, it is not always feasible to score these morphological traits, eg. reproductive tissues aren't necessarily present. Environmental variation influences and/or alters these morphological traits. Therefore, red and black spruce may not always be distinguished using morphological parameters and are sometimes misclassified in mixed stands.

Classification of putative hybrids is even more difficult. Morphological indexes used to identify hybrids are very controversial. The frequency and persistence of hybrids has been the subject of considerable debate (Morgenstern and Farrar 1964; Gordon 1976; Manley and Ledig 1979; Wilkinson 1990; Berlynet al. 1990). A number of researchers have questioned the purity of provenance seed source samples (Morgenstern and Farrar 1964; Morgensternet al. 1981; Wilkinson 1990), suggesting that a number of provenances contain hybrid and introgressed individuals.

The difficulty in classifying red spruce, black spruce, and hybrids may, in part, be due to introgression and gene flow between red and black spruce during glaciation events.

Prior to the last glacial retreat northeastern spruce was centered in the southern Appalachians (Daviset al. 1980), the present location of southern red spruce populations. Species migration in response to glaciation has influenced genetic variation in conifers (Critchfield 1984) and may have resulted in many hybridization events between red and black spruce.

Manley and Ledig (1979) showed, in laboratory studies, that red x black hybrids exhibit negative heterosis, with hybrids exhibiting reduced levels of photosynthesis in comparison with parents. They predicted that naturally occurring hybrids would be at a severe disadvantage as compared to parental species and that the hybrids would be only likely to succeed in habitats opened by fire and logging. Hybrids are most likely to survive in areas where competition from parental species has been removed or reduced (Gordon 1976).

In contrast to laboratory studies where hybrids have been reported to be inferior to the parental species, a number of researchers have reported hybrid superiority in the field. Hybrids have been reported to be less prone to spruce budworm infestation than red spruce (Osawa 1986; Osawa, Spies, and Dimond 1986). A number of researchers have also reported that hybrids and introgressed spruce show less winter drying damage than red spruce (Morgenstern 1969; Roche 1969; Wilkinson 1990) and that introgressed spruce display superior growth rates particularly in height (Wilkinson 1990).

The degree of hybridization and between-species gene flow is controversial as is the habitat which may enhance hybridization and introgression. A significant relationship was reported between elevation and degrees of introgression on a number of New England mountains including Mt. Washington (Berlynet al. 1990). Enhanced hybridization has also been reported between black spruce located in bog land and red

spruce located upland of the black spruce within a number of populations collected in Maine (Osawa 1986; Thorpe 1986; Osawa, Spies, and Dimond 1986).

Identification of field samples have traditionally been based on morphology. It has been assumed that red spruce surveys used in the decline studies have been pure red spruce; however, because of the difficulty involved with classifying red and black spruce based on morphology, the controversy surrounding the presence of hybrids and the relationship between hybridization and habitat, this may not be the case. A number of researchers have suggested that plot surveys may contain introgressed spruce and/or pure black spruce (Berlynet al. 1990; Eckert 1989).

Molecular markers for Picea taxonomy.

More recently isozymes have been used to distinguish red and black spruce. Eckert (1989) developed a discriminant function, based on thirteen polymorphic isozyme loci, which could distinguish red and black spruce and their hybrids. The discriminant function correctly reclassified trees at a high degree of accuracy: 92% red spruce, 97% for black spruce and 78% (7 out of 9 samples) for controlled cross hybrids.

Isozyme markers have limitations for phylogenetic and taxonomic studies. Isozyme analysis is an indirect method, identifying only those mutations in genes which change the electrophoretic mobility or the stability of enzymes. With DNA markers, either restriction fragment length polymorphisms or DNA sequence variants, a much larger fraction of the genome can be surveyed for within-species and between-species differences.

Berlyn and coworkers (1990) used nuclear genome content as an index for red spruce, black spruce, hybrids and introgression. In their study, samples were collected from elevational transects of Mount Washington (New Hampshire), Camels Hump

(Vermont), and Whiteface Mountain (New York). Using the nuclear genome size and a morphological index, modified from Manley (Manley 1971, as described in Berlynet al. 1990) all trees collected from Whiteface Mountain were classified as pure red spruce, trees sampled on Camels Hump were classified as introgressed or hybrid (no pure black or red spruce were detected), and trees on Mount Washington displayed introgression along an elevation gradient with pure black spruce above 1220 m, hybrids occurring between 1000 m and 1220 m, and pure red spruce at the lower elevations. Berlyn and coworkers (1990) concluded that hybridization between red and black spruce was frequent and that levels of introgression were related to elevation.

A better understanding of the genetic relationship between red spruce, black spruce and their hybrids along with a better understanding of the conditions leading to hybridization may help in determining the factors contributing to declines in red spruce. Hybridization events, even if rare, could result in gene flow from black spruce, a cold site species, to red spruce, sufficient to increase levels of cold tolerance in red spruce. To test this hypothesis molecular markers are needed to distinguish the species and measure the degree of introgression between the species.

Taxonomy based on the structure of the nuclear ribosomal DNA repeat.

The structure of the rDNA repeat has been determined for many angiosperm species (reviewed in Appels and Honeycutt 1986; Jorgensen and Cluster 1988). The rDNA repeat has been useful in systematic and phylogenetic studies among some groups within the angiosperms (Saghai Maroofet al. 1984; Doyle and Beachy 1985; Sytsma and Schaal 1985; King and Schaal 1989; Allardet al. 1990; Saghai Maroofet al. 1990). Both restriction site polymorphisms and repeat length variation were observed in several species

of soybean (Glycine max (L.) Merr.) and its relatives; however, no variation in restriction sites or length was observed within an individual (Doyle and Beachy 1985). A phylogenetic tree for the Lisianthius Skinneri (Gentianaceae) species complex was constructed using restriction site polymorphisms in the rDNA repeat (Sytsma and Schaal 1985). Again, no variation was observed in the rDNA repeat units within an individual. Length variation was absent in the rDNA repeat unit of Rudbeckia missourienses (Englm.); however, restriction site variation occurred among rDNA repeat units within all individuals studied (King and Schaal 1989). Both length and site variation have been observed within individuals of the genera Zea and Tripsacum (Zimmer, Jupe and Walbot 1988). Spacer length variants were observed in wild and cultivated barely (Hordeum vulare) distributed over two loci (Allardet al. 1990). Significant correlations occurred between rDNA alleles at the two loci and between rDNA alleles and specific factors of the physical environment (Saghai Maroofet al. 1990). It was concluded that natural selection plays a major role in the development and maintenance of observed patterns of molecular and genetic organization of rDNA variability (Saghai Maroofet al. 1990).

Little is known about the rDNA repeat unit in gymnosperms. Cullis and coworkers (1988) characterized the rDNA repeat unit in Pinus radiata D. Don and compared it to angiosperms. The rDNA repeat unit size of Pinus was more than twice the size of any angiosperm repeat unit, approximately 27 kbp in length. The distribution of arrays in the nucleus was also found to differ between angiosperms and Pinus radiata. Pinus radiata has ten major rDNA arrays and a number of minor arrays. Minor rDNA arrays have not been noted in angiosperms.

Among the Pinaceae, genome size has been shown to vary considerably within species (Mergen and Thielges 1967; Miksche 1967; Miksche 1968; Miksche, 1971; Dhir

and Miksche 1974). The rDNA gene number was found to increase with latitude in Douglas-fir (Pseudotsuga menziesii [Mifb.] Franco) (Strauss and Tsai 1988), corresponding with observed trends of increasing genome size in angiosperms (Bennett 1976; Levin and Fuderburg 1979) and Douglas fir (El-Lakany, and Sziklai 1971). Several investigators have suggested that variation in the copy number of the rDNA repeat unit could account for a disproportionate amount of the variation in genome size (Grime and Mowforth 1982; Strauss and Tsai 1988).

The repetitive nature of the nuclear rDNA repeat unit may be useful when evaluating within-species and between-species variation in red and black spruce and assist with assessing the degree of natural hybridization and introgression between these Picea species. Additional molecular markers, for example derived from the organelle genomes, may be helpful in determining the direction of hybridization.

Inheritance of conifer organelle genomes.

The inheritance patterns of the conifer organelle genomes are different from that observed in angiosperms. Maternal inheritance of both mitochondria and chloroplasts is common among angiosperms. Conifers have maintained the trend toward uniparental inheritance of organelles (Neale and Sederoff 1988); however, paternal inheritance of chloroplast is prevalent (Stine and Keathley 1990; Neale, Marshall and Sederoff 1989; Wagneret al. 1987). Most conifers exhibit maternal inheritance of mitochondria; however, paternal inheritance of mitochondria has been observed in Sequoia sempervirens (D. Don Endl) (Neale, Marshall and Sederoff 1989). Neale and Sederoff (1988) provide a more complete review of organelle inheritance and evolution in the Coniferales. The inheritance pattern observed in the Coniferales may be useful when identifying putative

Picea hybrids and assessing the degree of natural hybridization.

The primary goals of my research were to:

1) Characterize the nuclear rDNA repeat unit for Picea rubens and Picea mariana.

2) Examine within-species and between-species variation in nuclear rDNA allelic frequencies.

3) Develop an accurate index for identifying red spruce, black spruce and hybridization events.

In the first chapter, I present the characterization of the nuclear rDNA repeat unit in Picea rubens and Picea mariana. Polymorphic restriction sites are identified and the rDNA copy concentration is analyzed to test for significant differences in the nuclear rDNA copy concentration between red and black spruce, and the relationship between rDNA copy concentration and geographic origins of spruce populations is examined.

In the second chapter, I use provenance (seed source) plantation samples to investigate within-species and between-species variation in nuclear rDNA allelic frequencies.

In the third chapter restriction fragment length polymorphisms for both the mitochondria and chloroplast genomes are identified for red and black spruce. Data collected from controlled-cross hybrids are used to deduce the inheritance patterns of the organelles. Mitochondrial and chloroplast RFLP data are combined with quantitative RFLP differences in the nuclear rDNA repeat yielding methods useful in distinguishing red spruce, black spruce and identifying hybridization events between the two species.

In the fourth chapter samples from elevational transects of Mount Washington and

Mount Lafayette are examined. Trees from Mt. Washington are classified using both a morphological index (Eckert 1990) and the molecular DNA indexes presented in chapter three. Samples collected from Mt. Lafayette are classified using the molecular indexes. Nuclear rDNA markers are examined for significant variation related to elevation and for differences between populations from these two mountains.

Finally, in chapter five, a population of red and black spruce from Isle au Haut, ME is examined for interactions between red and black spruce populations. The population consists of black spruce, localized to a bog, red spruce upland and inland of the black spruce, one region between the bog and upland site, and a region between the bog and the ocean. Samples are classified using the three character index described in Chapter 3. Levels of hybridization and introgression are examined for relationships with habitat type.

Methods and Materials

Foliage collections

Red spruce samples were collected from a provenance (seed source) test plantation, located in Coleman State Forest, Stewartstown, New Hampshire. The red spruce plantation is maintained by the State of New Hampshire. Foliar tissue samples were collected from 58 individuals representing 12 provenances. Four to seven trees were sampled per provenance. Black spruce samples were collected from a provenance test plantation, maintained by the US Forest Service (Northeastern Forest Experiment Station), in the Massabesic Experiment Forest, Alfred Me. Black spruce samples were collected from 31 provenances (102 individuals) of the eastern complex of Picea mariana, with one to four individuals from each provenance. Duplicate samples were collected randomly from 20% of the red spruce and 10% of the black spruce. Duplicate samples normally consisted of one sample taken from the lower branches and the second from near the crown. An additional collection was made in 1991 to examine black spruce samples misclassified with the molecular indexes described in Chapter 3.

Needle samples from red-on-black controlled-cross hybrids were donated by Dr. James Hanover of Michigan State University. Red and black spruce used to create hybrids were from natural stands. The hybrids were made using a red spruce pollen mix from trees located in Michigan and Pennsylvania. Black spruce trees used to produce hybrids were from a stand located in northern Michigan. The relationship among black spruce and among red spruce parents is unknown.

Foliar tissue samples were collected from 80 spruce trees up an elevational transect on Mount Washington. Collections were made at eight elevations: 530 m (1740

ft.) (n=10), 610 m (2000 ft) (n= 10), 670 m (2200 ft) (n=10), 790 m (2600 ft) (n=10), 880 m (2900 ft) (n=15), 1060 m (3500 ft) (n=8), 1220 m (4000 ft) (n=8), and 1435 m (4700 ft) (n=8). Sample collections from Mount Washington were made in March of 1991. Mount Lafayette samples were collected, up an elevational transect, from 94 individuals at three elevations: 710 m (2320 ft) (n=29), 1190 m (3900 ft) (n=39), and 1430 m (4685 ft) (n=26).

Foliar tissue samples were collected from 87 spruce trees located at Head Harbor on Isle Au Haut Maine. Collections were made at four sites: Bog (site 2) (n=26), upland and inland of the bog (site 4) (n= 17), intermediate region between sites 2 and 4 (site 1) (n=26), and one site between the bog and ocean (site 3) (n=18). A number of trees located within site 1 were identified as putative hybrids using a modification of the index described by Gordon (1976) (Smoot Major personal communication). Other trees collected from site 1 and trees collected from sites 2-4 were originally identified based on gross morphological traits: foliar color, twig color, cone size and shape, crown shape, and twig ridge shape, no attempts were made to quantitatively classify samples, based on morphology, in the lab. Site 2, the bog, contained trees which displayed predominantly black spruce traits whereas site 4 contained trees which displayed predominantly red spruce traits. Site 1 contained trees which displayed predominantly black spruce traits and a number of individuals displaying hybrid characteristics, and site 3 contained trees which displayed predominantly red spruce traits (Smoot Major personal communication).

Between 5 to 10 grams of current year needles were collected from each tree. Samples were held in either dry ice or wet ice in the field and then stored at -70°C.

Morphological classification of trees

Samples collected from Mt. Washington were classified using the morphological index developed by Eckert (1990). Samples were classified in Dr. Robert T. Eckert's laboratory by Karen D. Stapelfeldt. Foliage for morphological examination was left on branches and maintained in cold storage (10°C) in ziploc plastic bags until evaluation, three days from time of collection. Scoring was done on second year growth on the primary stem of the branch held at reading distance, except on twig ridges which was done under 8X magnification with a binocular dissecting microscope in the laboratory. Color determination was done with Munsell plant color plates.

Several variables described by Gordon (1976) have been proven reliable indicators of species differences in earlier efforts to differentiate red and black spruce, and their hybrids (Eckert 1990). Hue, value, and chroma of abaxial and adaxial foliar surfaces, overall color of foliage according to a subjective scale developed by Gordon (1976), terminal bud color and the cross-sectional shape of twig-ridges were scored for the Mt. Washington samples.

Morphological data for 69 specimens collected on Mt. Washington were evaluated with a discriminant model developed from provenance plantation material located in Stewartstown, NH (72 red spruce individuals from a range-wide sample) and Alfred, Me (60 black spruce individuals from the eastern Picea mariana complex). Known hybrid material (15 individuals) supplied by A. Gordon, was included in the morphological discriminant analysis. A discriminant model which correctly reclassified red and black spruce more than 95% of the time, and which reclassified hybrids more than 64% of the time, was used to sort the samples according to species and hybrids.

DNA Extraction

DNA was extracted from 3 to 10 grams of needles using the method described by Greenwood, Hopper and Hutchinson (1989) with slight modifications: Intact needles were added directly to 150 ml ice cold extraction buffer [10 mM Tris (Ph 8.0), 5 mM EDTA, 350 mM D-sorbitol, 0.1% bovine serum albumin, 14μM 2-mercaptoethanol, and 10% polyethylene glycol (average molecular weight=8000)]. The needles were homogenized in extraction buffer for 20 to 30 seconds using a Brinkman Polytron. Ethanol was substituted for isopropanol for the first DNA precipitation. The RNase digestion step was found to be unnecessary. This extraction procedure normally yielded 1 mg DNA per 6 grams of needles.

Absorbance readings at 260 nm were used to quantitate DNA concentration. Electrophoresis was used to determine RNA contamination of the DNA preparations and the overall quality of the extracted DNA. DNA's have been stable at 4°C for at least one year. DNA samples used for the quantitative analysis of rDNA copy number were further purified by an RNase digestion step and an isopropanol precipitation.

Restriction Digests and Gel Electrophoresis

DNA samples were digested with restriction endonuclease using the conditions specified by the manufacturer (Boehringer Mannheim Biochemical, BRL Life Technologies Inc., New England Biolabs Inc., Promega Corporation, or United States Biochemical Corporation).

Restricted DNA fragments and radiolabeled lambda DNA size markers were separated by size on 0.5 % to 0.9 % agarose gels as described by Sambrook, Fritsch and Maniatis (1989; section 6.3-6.15, book 1). For fragment isolation, restricted DNAs were

separated on Low melt, SeaPlaque GTG or NuSieve GTG agarose purchased from FMC Bioproducts. Direct labeling of DNA was accomplished by separating DNAs on 0.8% SeaPlaque agarose gel in N-TBE buffer (100 mM Tris-HCl (pH 8.0), 100 mM Boric acid, and 0.2 mM EDTA) as described by Feinberg and Vogelstein (1984).

Southern Blots

Southern blots were made by transferring DNA fragment from agarose gels, using capillary action, onto nylon membrane (Bio-Rad Corporation Zeta-probe blotting membrane) as described by Sambrook, Fritsch, and Maniatis (1989; section 9.38, book 2). DNAs were fixed onto the membrane by UV irradiation, using a modified procedure of Khandjian (1986) as described by Dowe, Roman and Klein (1990). Southern blots were stored at 4°C.

Quantitative analysis of rDNA copy number

Slot blots were used to determine the nuclear rDNA copy concentration. Slot blots were produced using a Schleicher & Schuell Minifold II Slot Blot System. Blots were produced following the directions provided by the manufacturer. The plasmid pXBr 1 was used as a standard. One μg of genomic DNA was loaded into each slot. The 18S fragment cloned into pXBr 1 was used as a probe for the slot-blots.

Probes

Probes consisted of an 18S and 26S rDNA fragment from soybean (Glycine max) rDNA, a PCR amplified ITS 1 fragment (internal transcribed spacer one) from spruce, a

maize mitochondrial clone and a petunia chloroplast clone. The 18S and 26S rDNA fragments were obtained from Dr. Elizabeth Zimmer in the form of plasmids pXBr 1 and pGmr 1 (Zimmer et al 1988). These coding regions of the rDNA are highly conserved throughout the plant kingdom and hybridized to the spruce DNA at high stringency. Soybean ITS and IGS (inter-genic-spacer) regions did not hybridize to the Picea DNA. The red spruce ITS 1 region was amplified using PCR techniques and primers described by Smith and Klein (in preparation). The mitochondrial clone was obtained from Dr. C. S. Levings III and consisted of a 6.0 kbp fragment containing the maize mitochondrial 18S-5S region (Chao, et al 1984). The chloroplast fragment, plasmid pST6, was obtained from Dr. David Neale (Palmer, et al 1983).

DNA fragments were labeled using a United States Biochemical or a Boehringer Mannheim Random Primed DNA labeling kit and alpha ^{32}P labeled dATP and TTP (New England Nuclear (Dupont)).

Hybridizations and Autoradiograms

Pre-hybridizations and hybridizations were performed as described by Sambrook, Fritsch and Maniatis (1989; section 9.47-9.51, book 2) with minor modifications. Pre-hybridization solutions contained 50% deionized formamide, 4% SSPE (1% SSPE = 0.15M sodium chloride, 0.01M sodium phosphate, and 0.001 M EDTA, pH = 7.4), 1% SDS, 0.25% BLOTTO (5% non-fat dried milk dissolved in water containing 0.02% sodium azide) and 0.3 mg per ml blocking DNA (herring sperm or salmon sperm DNA). Hybridization solutions were similar to the pre-hybridization solution except that dextran sulfate, final concentration 2.5%, and probe were added to the hybridization reaction. Both pre-hybridizations and hybridizations were performed in a batch mode, in Nalgene

plastic boxes. Blots were pre-hybridized for a least 4 hrs at 42°C and were hybridized for at least 20 hrs at 42°C. Blots were then washed as follows: 1X SSC (0.15M sodium chloride and 0.015 M sodium citrate, pH = 7.0) at room temperature for 10 min, 0.5X SSC, 0.5% SDS at 65°C for 15 min, 0.2X SSC, 0.2% SDS at 65°C for 30 min, and 0.15X SSC at 65°C for 20 min.

Blots were wrapped in saran wrap and exposed to Cronex film (Dupont) with a Cronex Quanta III intensifying screen (Dupont) at -70°C for variable times to produce autoradiograms. Autoradiogram exposures were limited to between 15 minutes to 2 hours for slot blot procedures to insure that exposures were within the linear range of the X-ray film.

Densitometry

The relative amounts of rDNA repeats polymorphic for HinD III, Sst I (or Sac I), and EcoR I fragments were measured for each individual. In most cases two bands appeared on the autoradiograms for a specific probe-restriction digest combination for an individual. The intensity of a band on an autoradiogram is linearly related to the amount of radioactive probe which hybridized to the fragment on the Southern blot; therefore the relative intensities of the restriction fragment bands are directly proportional to the frequency of the polymorphic rDNA types within the genome.

The distribution of polymorphic restriction fragments within each individual, were determined by reading down the autoradiogram using a GS 300 Transmittance/Reflectance Scanning Densitometer (Hoefer Scientific Instruments). The relative intensities of these bands were calculated using a GS 350 Data System Electrophoresis Data Reduction System (Hoefer Scientific Instruments) on a Zenith Data System computer connected to

the densitometer.

A number of controls were run to determine the accuracy of the RFLP frequency analysis. Random needle samples were divided and extracted separately to determine variation due to the DNA extraction procedure. Random DNA samples were selected to run in duplicate, and sometimes triplicate, to determine the variation due to incomplete endonuclease digestion and variation among different Southern blots. A number of Southern blots were reprobed and analyzed to determine the variation due to hybridizations and washes. The distribution between polymorphic restriction fragments was accurate to within 5% (0.05) among replicates with the exception of frequencies reported as 0.0 or 1.0. Samples reported to have a frequency of either 0.0 or 1.0 are much more accurate. Overexposures and duplicate samples were used to confirm the presence or absence of these fragments.

Analysis of Data

The frequencies of the upper and lower molecular weight fragments sum to 1.00 where the upper molecular weight fragment represented the absence of a specific site and the lower molecular weight fragment represented its presence. The frequency of one fragment can be calculated directly from the frequency of the other; therefore, the lower molecular weight fragment frequencies were used as the variable for the analyses. Within some black spruce individuals there were three over-lapping Hind III fragments, the frequencies of the 1.6 molecular weight and 3.4 molecular weight HinD III fragments were used as variables.

The organelle genomes were screened for polymorphisms that may be useful in species identification. A mitochondrial polymorphism was identified by Dirk E. Smith and

a chloroplast polymorphism was identified by Denis Guenette. Mitochondria and chloroplast haplotypes were scored and entered as transformed variables: 1 for haplotype A and 0 for other haplotypes. Mitochondria and chloroplast data were combined with the nuclear rDNA allelic frequencies and analyzed using discriminate analysis. The discriminant function, used to classify each individual into a specific species, is based on a measure of the generalized squared distance (Rao 1973).

The frequencies of polymorphic restriction fragments for individuals were analyzed using a number of statistical methods including: Multiple regression analysis, canonical correlation analysis and discriminant analysis. Canonical correlation analyses were used to examine the relationship between the different molecular markers and the geographic origin of the provenances (Appendix 3). The discriminant analysis was used to classify each individual according to species based on a measure of the generalized squared distance (Rao 1973) (Appendix 4). Ten-fold verification of the discriminant model was accomplished using randomly selected sub samples of the original data set. Canonical variates generated in discriminant analysis were used to assess dimensionality in the data and for plots of canonical scores of individual trees in each species. A more complete description of canonical analysis is given by Gittins (1979).

The statistical calculations were accomplished on a Digital VAX 5800 computer. SAS version 6.1 program was used for all statistical analyses (SAS Institute Inc.).

Chapter 1

Characterization of the nuclear rDNA repeat unit of *Picea rubens* and *Picea mariana*.[1]

[1] Results from the following chapter will be being published as: Bobola M.S., Smith D.E., and Klein A.S. 1992. Five major nuclear ribosomal repeats represent a large and variable fraction of the genomic DNA of Picea rubens and P. mariana. Molecular Biology and Evolution. In press.

Results and Discussion

Structure of the Nuclear rDNA Repeat Unit in Picea

A composite restriction map of the nuclear rDNA repeat unit (figure 1) in Picea was generated using heterologous soybean rDNA probes and a PCR amplified spruce ITS 1 DNA fragment. Coding regions from the soybean rDNA repeat and the 18S and 26S rDNAs hybridized at high stringency (0.15 X SSC, 30 min at 65°C). Hybridizations, using coding regions as probes, produced autoradiograms with a high signal to noise ratio (figure 2). Non-coding sequences, the soybean IGS and ITS 1 and 2 regions, were not sufficiently conserved between angiosperm and gymnosperm to allow the cloned soybean IGS, or ITS regions to be used as probes for mapping the gymnosperm rDNA repeat (data not shown). Chloroplast and mitochondrial rDNA clones (from petunia and maize respectively) annealed to restriction fragments of different sizes than that of the nuclear rDNAs (data not shown).

Eleven restriction endonucleases were used to map the Picea rDNA repeat. The positions and distances among the restriction sites within the 18S and 26S regions were identical between P. rubens and P. mariana. Several restriction fragment length polymorphisms (RFLPs) were detected in the IGS and ITS regions; these are denoted with lower case symbols in figure 1. There are several rDNA alleles within a given individual. For example, a given restriction enzyme-probe combination sometimes revealed both shorter and longer forms of the restriction fragment indicating that the rDNA repeat units within an individual are polymorphic with respect to the specific restriction site (Figure 2 A, B, and C). Figure 2 (A) represents a Hind III digest of genomic DNAs hybridized with the 18S probe. In sample 4986-1 three bands appear (1.6,

Figure 1: A composite restriction map of the nuclear rDNA repeat of Picea rubens and P. mariana. Total genomic DNAs were restricted, separated by size using agarose gel, electrophoresis and transferred via capillary action to nylon membranes. Radiolabeled, restricted lambda DNA were included on all the gels as internal size standards. Blots were hybridized with soybean rDNA or spruce ITS 1 probes. Polymorphic restriction sites are shown as lower case letters. Restriction fragments corresponding to rDNA alleles are indicated by bars. B = BamH I, Bg = Bgl II, E = Bste III, H = Hind III, K = Kpn I, P = Pst I, RI = EcoR I, RV = EcoR V, S = Sac I, X = Xba I, Y = Bsty I.

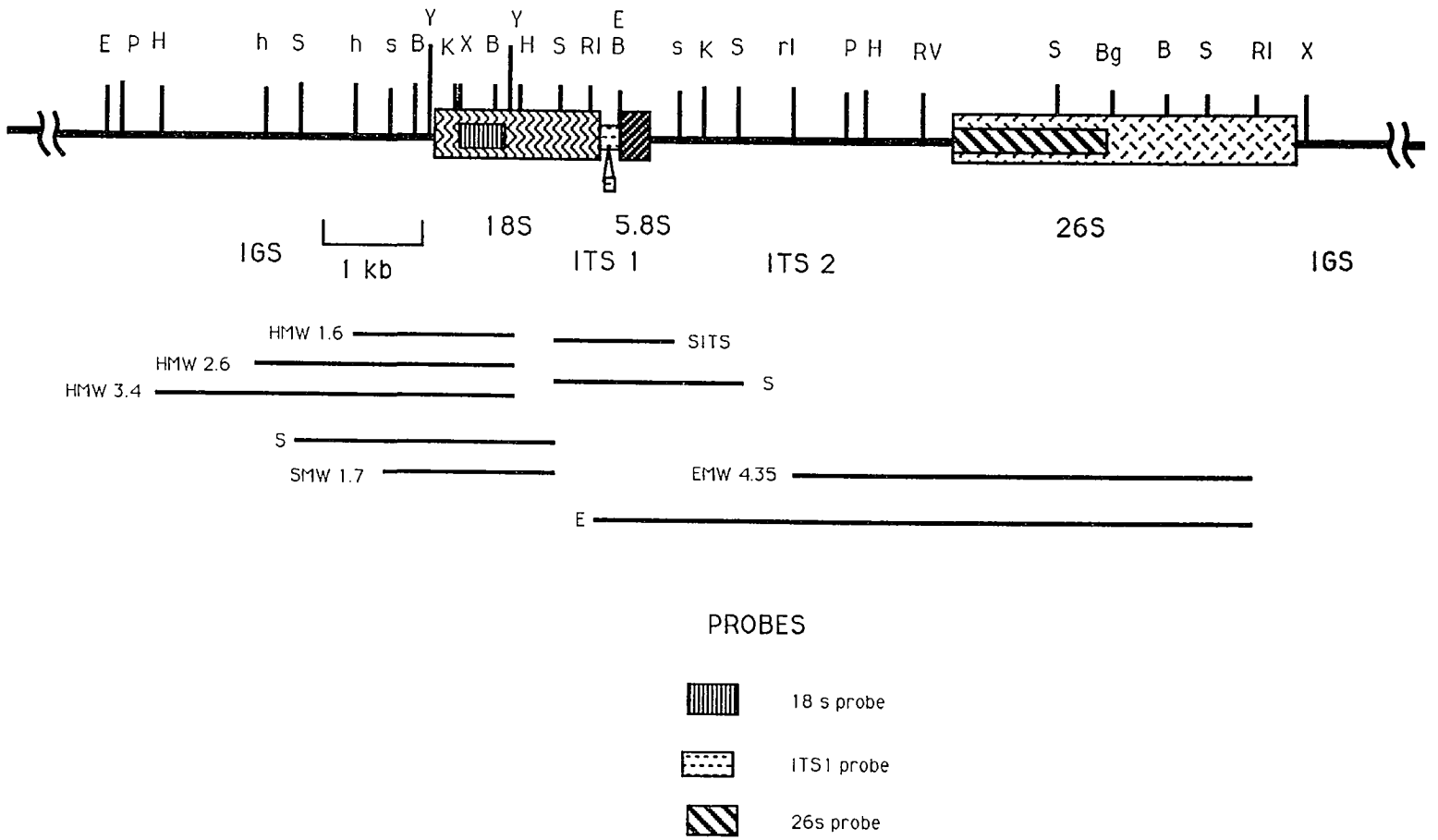


Figure 2: Polymorphic restriction fragment in the Picea rubens and P. mariana rDNA. DNAs cut with A: HinD III B: EcoR I and C: Sac I, separated electrophoretically and blotted onto nylon membrane (Southern blotting). The Southern blots were then probed with the appropriate DNA fragment: HinD III with the 18S probe (A), EcoR I with the 26S probe (B), and Sac I with the ITS 1 probe (C). Note the relative intensities of the bands differ between the samples. The sample identification number, corresponding to tree number in appendix 1, is located below each lane. Identification numbers followed by an A or B represent duplicate samples.

2.6, and 3.4 kbp respectively). The 3.4 kbp fragment represents rDNA alleles lacking both of the polymorphic Hind III sites.

Repeat Size

The size of the spruce rDNA repeat unit was estimated from Southern blots of EcoR V or Bgl II digested DNA using the 18S probe. The fragments observed using this combination of probe and either enzyme varied between 32 kbp and greater than 40 kbp in size (data not shown). Reprobing blots with the 26S probe generated bands the same size, suggesting that only one EcoR V and only one Bgl II recognition site is located within a repeat unit. The sizes of the Picea rDNA repeat units are more than twice that observed in angiosperms but are similar to the 27 kbp rDNA repeat size reported in Pinus radiata D. Don (Cullis et al. 1988).

Within individual variation in the rDNA repeat unit

Restriction analysis was extended to provenances of red and black spruce. Red spruce was sampled from a range wide provenance test. The black spruce provenance test represents the Eastern complex of Picea mariana. No single polymorphic restriction site was exclusive to one species, but the relative proportions of restriction sites varied between individuals (Figure 2).

Fragments generated using the 18S probe on EcoR I digested DNA varied between 22 and 30 (+/- 2) kbp in size. Five major rDNA fragments, representing different rDNA repeat types (alleles), could be distinguished: 23, 24, 25, 27, and 28 kbp. Additional experiments indicated that the multiple fragments were not due to incomplete digestion. EcoR I digests probed with the 26S probe yielded distinct fragments of 4.35

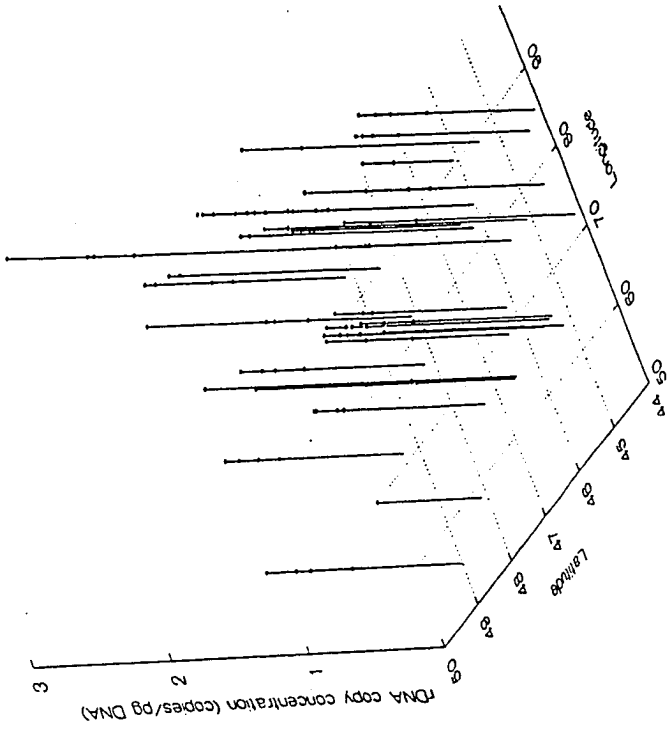
and 6.5 kbp. Taken together, these data suggest that as many as five different nuclear rDNA repeats exist within the Picea genome. In Pinus radiata, ten major rDNA arrays were detected by in situ hybridization along with a number of minor arrays (Cullis et al. 1988). The different major polymorphic repeat units detected in Picea may correspond to different major arrays in the genome.

Slot-blot analysis was performed on all red and black spruce individuals from the provenance test sites to determine the copy number of the rDNA repeat unit, (Figure 3). The standard slots contained pXBr 1 corresponding to 770 to 3850 copies per pg spruce DNA. The mean rDNA copy concentration observed in red spruce was 1320 (+/- 290) copies per pg with nearly a three fold range (750 to 2220 copies per pg) and 1250 (+/- 370) copies per pg for black spruce with a six fold range (430 to 2650 copies per pg). Strauss and Tsai (1988) observed as much as twelve fold variation in the relative rDNA content of Douglas Fir (Pseudotsuga menziesii (Mirb.) Franco) provenance plantation samples. The rDNA copy concentration was not significantly different between red and black spruce based on a F-ratio test ($F=3.35$, $p=0.069$); therefore, the rDNA copy concentration cannot be used to accurately discriminate between the two species.

Relationship between geographic origins and nuclear rDNA concentration

No significant statistical relationship were detected between rDNA copy concentration and geographic origin of the Picea mariana provenances. In Picea rubens, however, a significant ($p=0.002$) relationship was detected between rDNA copy concentration and geographic origins (latitude, longitude, latitude X longitude and latitude squared) (table 1) using a standard multiple regression analysis. Longitude squared, elevation and elevation squared did not display a significant relationship with copy

Figure 3: Three dimensional plot depicting rDNA copy concentration for (a) Picea rubens and (b) P. mariana over geographic origins of provenance seed sources.



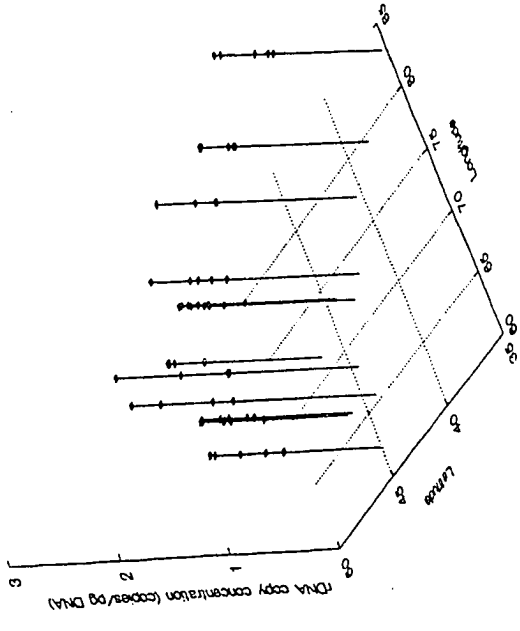


Table 1 Regression analysis of copy number on geographic origins of red spruce provenances.

Parameter	DF[1]	Standard Error for	T test Hypothesis[2]	Prob>|T|
Intercept	1	71.53	2.65	0.0106
Latitude	1	2.33	-2.49	0.0158
Longitude	1	0.54	-2.86	0.0060
Lat*Lon	1	0.01	2.84	0.0063
Latitude Sq.	1	0.02	2.21	0.0318

R^2= 0.27[3] Adj-R^2= 0.21[4] p = 0.002

[1]Degrees of freedom

[2] The T test statistic measures the strength of the linear relationship between the dependent and independent variables.

[3] R^2 is a measure of the proportion of the variation in the dependent variable explained by the model.

[4] Adj-R^2 corrects R^2 to more closely reflect the goodness of fit of the model in a natural population. Adj-R^2 takes both the sample size and the number of independent variables into consideration.

concentration and were therefore eliminated from the analysis. The red spruce provenance plantation used in this study was not set up for examining variation over elevational gradients; consequently elevational effects were not examined. Geographic origins of red spruce provenance samples explained 21% of the variance in rDNA copy concentration. This implies that within P. rubens, rDNA copy number differentiates across the geographic range of the species. Since only a fraction of the variations in rDNA content was accounted for by geographic origin, there must be other factors influencing the rDNA content of the spruce genome.

A weak relationship was observed between the latitude and elevation of Douglas fir provenances and relative rDNA copy number (Strauss and Tsai 1988), with relative rDNA copy number increasing with latitude. Grime and Mowforth (1982) and Price (1988) suggest that increases in genome size, in plant species, correlate with the climate at which the organism originates, with colder climates giving rise to larger genomes. Grime and Mowforth (1982) also suggest that the increases in genome size may be due to amplification of genes involved in protein synthesis.

The total genomic DNA content for several species of conifers has been shown to vary with geographic location and these differences are not thought to be due to simple changes in ploidy (Mergen and Thielges 1967; Miksche 1967, 1968, 1971; Dhir and Miksche 1974). The relatively high rDNA content of the Picea genomes and substantial range in rDNA content could contribute to the overall within-species variation in conifer genome sizes.

Chapter 2

Restriction fragment variation in the nuclear rDNA repeat unit within and between *Picea rubens* and *Picea mariana*.[1]

[1] Results from the following chapter will be published as: Bobola M.S., Eckert R.T., and Klein A.S. 1992. Restriction fragment variation in the nuclear rDNA repeat unit within and between Picea rubens and Picea mariana. Canadian Journal of Forest Research. In press.

Results

Polymorphic fragment within the rDNA repeat unit

Restriction fragment length polymorphisms (representing different allelic forms of the rDNA repeat unit) were identified in the inter-genic spacer,IGS, and internal transcribed spacer, ITS, regions of the nuclear rDNA repeat unit of Picea (Figure 1). Restriction fragment variation was evident between rDNA repeat units within an individual. The rDNA allelic frequencies varied between individuals of both P. rubens and P. mariana. Based on the frequency of the different polymorphic restriction fragments there must be at least five distinct repeat units in the Picea rDNA multi-gene family to account for the minimum possible combination of sites (Appendix 2). This result is in agreement with the previous observation that there were at least five size class variants of the nuclear rDNA repeat (Chapter 1).

The relative amounts of the polymorphic HinD III fragments (HMW 3.4, HMW 2.6, and HMW 1.6)(Figure 2A) and Sac I fragments (SMW 2.7 and SMW 1.7) located in the IGS region and the polymorphic EcoR I fragments (E and EMW 4.35)(Figure 2B) and Sac I fragments (S and SITS)(Figure 2C) located in the ITS region were measured for each individual, 55 red spruce individuals and 85 black spruce individuals (Appendix 1) (Figure 1). Duplicate samples taken from the same individual revealed no significant variation in rDNA types. The allelic frequencies of the rDNA alleles were analyzed to reveal correlations between individual alleles, between-species variation and variation over geographic origins of the provenance samples.

Correlations between individual rDNA alleles

The red and black spruce data sets were examined for correlations between alleles. The Pearson product-moment correlation matrix (Table 2), derived using the red spruce restriction data, revealed significant ($p \leq 0.05$) positive correlations between three of the markers (Figure 1): HMW 1.6 and SITS (r=0.51), HMW 1.6 and EMW 4.35 (r=0.41), and SITS and EMW 4.35 (r=0.28), and a significant ($p \leq 0.01$) negative correlation between HMW 1.6 and SMW 1.7 (r=-0.31). Correlations between variables using the black spruce restriction data (Appendix 1) revealed a significant ($p \leq 0.01$) positive correlation between HMW 1.6 and SITS (r=0.34) and a significant ($p \leq 0.01$) negative correlation between EMW 4.35 and SMW 1.7 (r=-0.31).

In the previous chapter considerable variation was detected in the concentration of nuclear rDNA among individuals in each species. The rDNA concentration data were compared to the polymorphic restriction data to determine if any correlation existed between allelic frequencies and concentration of nuclear rDNA repeat units (Table 2). A negative correlation ($p \leq 0.05$) occurred with SITS and rDNA concentration in red spruce (r=-0.28). No significant correlations occurred between the black spruce molecular markers and rDNA concentration.

Between-species statistical analysis

The means and standard deviations of each marker were calculated for red spruce and black spruce (Table 3). One marker (HMW 3.4) appeared in only 15 (16%) of the black spruce tested and was not present in the red spruce or the hybrids. Univariate analysis of variance showed significant ($p \leq 0.05$) between-species variation for all the molecular markers: HMW 3.4, HMW 1.6, EMW 4.35, SMW 1.7 and SITS (F = 6.0, 167.7,

Table 2. Pearson Product-Moment Correlation matrix. Correlations and significance levels among molecular markers and rDNA copy concentration for red spruce and black spruce.

	HMW 1.6	EMW 4.35	SMW 1.7	SITS
EMW 4.35	r[1] 0.406/0.002 b[2] -0.065/0.56			
SMW 1.7	r -0.306/0.023 b -0.077/0.49	r 0.066/0.63 b -0.309/0.005		
SITS	r 0.515/0.0001 b 0.335/0.002	r 0.283/0.037 b 0.069/0.53	r -0.083/0.55 b -0.118/0.288	
COPY	r -0.149/0.28 b 0.179/0.11	r 0.103/0.46 b -0.006/0.96	r 0.171/0.21 b -0.064/0.57	r -0.289/0.032 b -0.144/0.20

[1] r: Red spruce correlation coefficient/p.

[2] b: Black spruce correlation coefficient/p.

Table 3. Means, and standard deviations (S.D.) of the polymorphic restriction fragment frequencies for red and black spruce.

Site	Mean (S.D.)
HMW 3.4	Black 0.023 (0.07)
	Red 0.000 (0.0)
HMW 1.6	Black 0.512 (0.153)
	Red 0.182 (0.136)
EMW 4.35	Black 0.671 (0.256)
	Red 0.323 (0.189)
SMW 1.7	Black 0.582 (0.085)
	Red 0.527 (0.073)
SITS	Black 0.997 (0.016)
	Red 0.768 (0.106)

75.1, 15.9, and 384.5 respectively). There was no significant between-species variation in rDNA copy concentration (F=3.7).

Discriminant analysis of Picea species.

A discriminant function performed on the allelic frequencies, using a measure of generalized squared distance (Rao 1977) based on the pooled covariance matrix, distinguished red from black spruce (Appendix 4). Homogeneity of the variance-covariance matrices were violated in this analysis based on the allelic frequency. Violations were due to inequality of the red spruce, black spruce and pooled within group variance-covariance matrixes. Discriminant analysis, the multivariate extension of one-way ANOVA, is a robust procedure. For my purpose here in generating a statistical model which distinguishes red from black spruce, I am satisfied that the violations of the variance equality criterion does not strongly influence the effectiveness of the model in discriminating red from black spruce. The violations were due in part to the differences in sample size (black=85; red=55). These violations may result in misclassification of red spruce individuals as black spruce.

Reclassification of the trees used to produce the discriminant function was very accurate for red spruce (91%) and black spruce (99%). If the percentage of correct classification was low, there would be cause for concern about inequalities of the variance-covariance matrices. Greater dispersion of red spruce canonical scores may have resulted in classification of four red spruce as black spruce. Alternatively, these individuals may be hybrid or introgressed. Ten-fold verification of the discriminant model was based on ten populations, each consisting of independent random sampling of half the total data set. Each population was used to produce a discriminant model that was applied to the

samples not included in producing that model. The average correct classification, using the ten separate discriminant models, was 99% for black spruce and 96% for red spruce.

RFLP frequencies were also analyzed using canonical discriminant analysis to provide dimensional analysis and aid in spatial interpretation of the data. Canonical discriminant analysis uses a dimension-reduction technique related to principal component analysis and canonical correlation.

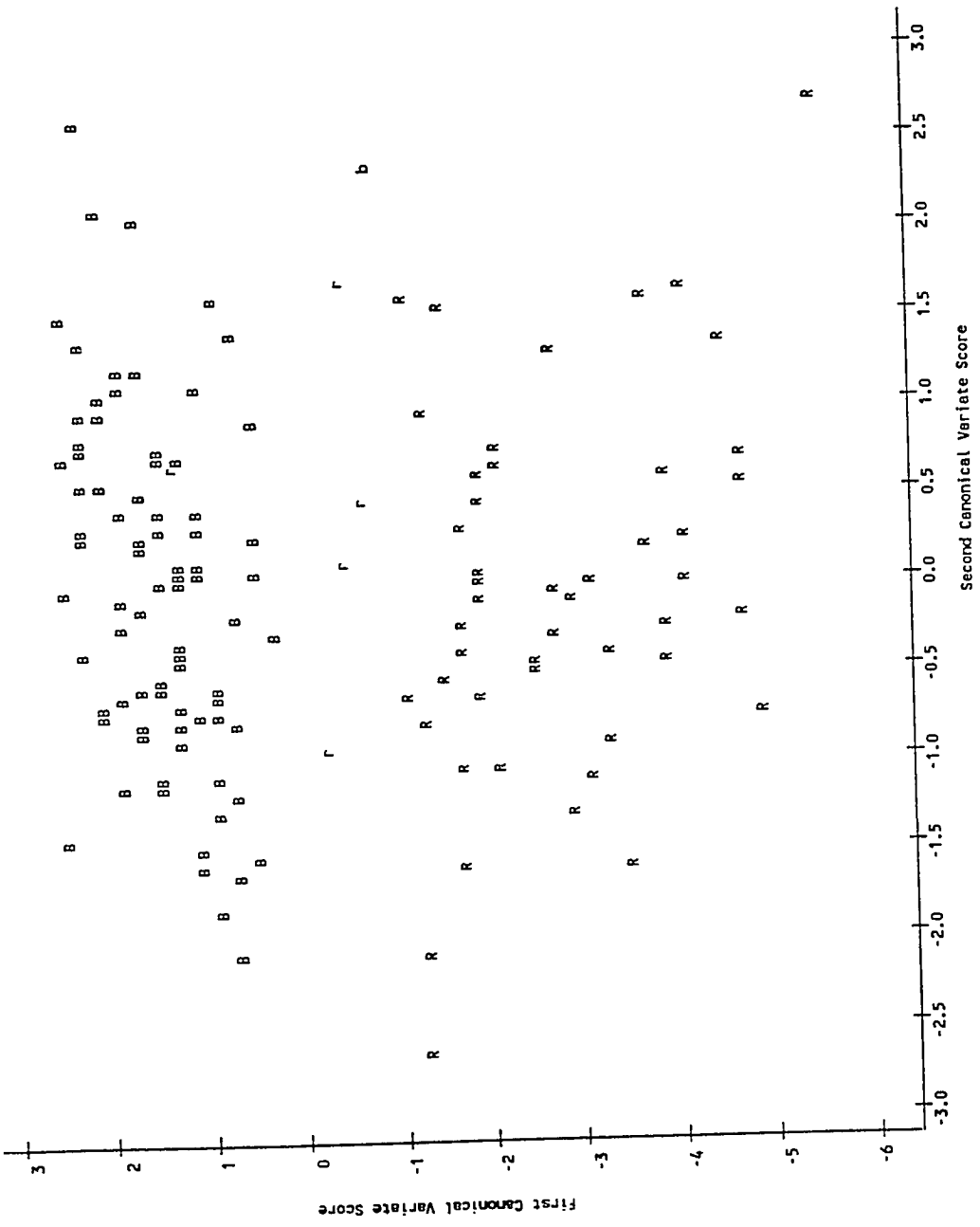
Applying the canonical discriminant analysis procedure to the allelic frequencies yielded one significant vector useful in separating the species ($p \leq 0.0001$, Wilks' Lambda=0.204). A second vector, unrelated to the first did not display any between-species variation but is useful for plotting the results. The first canonical vector accounts for 100% of the explained variation in RFLP data.

Canonical scores were calculated for each individual. Figure 4 was produced by plotting canonical scores for individual trees for sub spaces defined by the first two canonical vectors. Variation in scores along the first canonical vector gave good separation between red and black spruce.

Relationship between geographic origins of red and black spruce and rDNA RFLP frequencies

Canonical correlation analysis was used to identify relationships between the frequency of the rDNA allelic forms and geographic origins of the provenance samples. Statistically significant ($p \leq 0.05$) correlations occurred between allelic frequencies and geographic origins, controlling for correlations among the alleles and for correlations among the geographic variables. The marker identified as HMW 3.4 was deleted from the analysis because no red spruce individuals had this RFLP and only 15 black spruce

Figure 4. Plot of discriminant canonical variate scores for red and black spruce. The first canonical axis, can 1, is statistically significant (p=0.0001). The second canonical axis, can 2, is not statistically significant but assists in plotting the canonical scores. Plots of canonical scores for red (R) and black (B) spruce illustrate the clustering of the two species found when canonical analysis is applied to nuclear rDNA restriction fragment frequencies. The mean score for the first canonical variate is -2.44 for red spruce and 1.58 for black spruce. The mean score for the second canonical variate is 0 for both species. Misclassified individuals are shown in lower case letters. Four black spruce and three red spruce data points are hidden.



individuals (16%) had this RFLP. The marker identified as SITS was deleted from the analysis of black spruce variation because only five black spruce individuals displayed variation at this site. One red spruce individual was deleted from the analysis of red spruce variation because it was identified as a multivariate outlier.

Canonical correlation analysis revealed one significant (p=0.035) canonical correlation between the molecular markers (HMW 1.6 EMW 4.35, SMW 1.7,and SITS) and geographic origin (latitude, longitude, latitude squared, longitude squared, and latitude*longitude) of red spruce (appendix 3). Geographic origin accounted for 31.8% of the variation in the red spruce rDNA variables with latitude and latitude squared accounting for a major portion of this variation. The molecular marker EMW 4.35 accounted for a major proportion of the correlation between allelic frequencies and geographic origin. Examining the correlations between variables and the red spruce variate (Table 4) reveal that the frequency of the smaller EcoR I fragment, labeled EMW 4.35, increases with latitude.

Canonical correlation analysis also revealed one significant (p=0.0015) canonical correlation between the molecular markers for black spruce and geographic origin (latitude, longitude, latitude squared, longitude squared, and latitude*longitude). For black spruce, 24.7% of the variation in rDNA variables was accounted for by the geographic variables, with latitude and latitude squared contributing a major portion of this variation. The molecular marker HMW 1.6 accounted for a major proportion of the correlation between allelic frequencies and geographic origin. The sign of the correlations with the black spruce variate (Table 4) reveal that the frequency of the smaller HinD III fragment, labeled HMW 1.6, increases with latitude.

Table 4. Correlations of the first canonical variates for red and black spruce with molecular variables and geographic variables. Magnitude and sign of the correlation displays the contribution of each variable to the overall relationship between geographic origins and restriction site frequencies.[1]

Molecular Marker	Red Spruce Canonical Variate	Black Spruce Canonical Variate
HMW1.6	0.4759	0.8793
EMW4.35	0.7398	-0.2038
SMW 1.7	-0.4472	-0.3466
SITS	0.1852	
Geographic variable		
Latitude	0.6991	0.2170
Latitude squared	0.6855	0.2155
Longitude	-0.3793	0.0649
Longitude squared	-0.4099	0.0574
latitude*longitude	0.6109	0.1173

[1] Red spruce canonical R^2=0.318 (p=0.035)
Black spruce canonical R^2=0.247 (p=0.0015)

Discussion

In this chapter, the combination of restriction fragment frequency in the rDNA repeat (representing different rDNA alleles) and multivariate statistical procedures were used to discriminate between species that may not always be distinguished using morphological traits. Examining variation at the DNA level avoids some of the inherent problems of morphological and isozyme analysis. DNA analysis is not influenced by gene regulation, post-transcriptional or post-translational modifications.

Red and black spruce from two provenance tests were examined for variation in the nuclear rDNA repeat. The frequencies of different polymorphic rDNA alleles for each individual were obtained via Southern blot analysis and densitometry. Variation in the rDNA repeat was analyzed for within-species and between-species differences using a number of statistical techniques.

Environmental effects and geographic location may have some influence on genome size (Grime and Mowforth 1982, Miksche 1971), nuclear concentration of rDNA repeat units (Strauss and Tsai 1988) and on rDNA allelic frequencies (Saghai Maroof et al. 1990). The polymorphic rDNA alleles in these Picea were first analyzed for within individual variation. Duplicate samples were taken from a number of individuals and tested for within tree variation in rDNA allelic frequencies. Duplicate samples were taken from current year foliage from both lower branches of the tree and from the upper crown. No differences in the frequencies of the nuclear rDNA alleles were detected within the

limits of our methods (+/- 0.05[1]). These results suggest that there are no large changes in rDNA allelic frequencies over the several decades of growth separating the cell lineages from the crown to the lower branches of the trees.

The polymorphic restriction fragments, alleles, were then examined for relationships with rDNA copy concentration and for relationships among alleles. The relative copy number of rDNA genes was previously found to vary several-fold within Picea species (Chapter 1). Only one significant ($p \leq 0.05$) correlation was revealed between allelic frequencies and nuclear rDNA concentration (relative copy number), r=-0.2891 with SITS, in red spruce and no significant correlations occurred in black spruce (Table 2). The absence of strong correlation between rDNA concentration and allelic frequencies indicates that the mechanism(s) responsible for altering the nuclear rDNA concentration within Picea preserves the ratio of rDNA allelic types. Significant correlations did occur between certain allelic forms (Table 2) for both red and black spruce. The statistically significant correlations which did occur between the different rDNA markers were not large. The strongest correlation, between SITS and HMW 1.6 in red spruce, explained only 26% of that relationship (R^2=0.26). This suggests that the frequencies of the different rDNA alleles are changing independently of one another.

The allelic frequencies were analyzed for quantitative differences between species using discriminant analysis. Discriminant analysis, using the rDNA molecular marker frequencies produced a classification model which very accurately reclassified red (91%) and black spruce (99%). I therefore reject the null hypothesis that there is no difference between rDNA molecular markers for red and black spruce. Ten-fold verification of the

---

[1] Samples reported to have allelic frequencies of either 0.0 or 1.0 are much more accurate. Overexposures and duplicate samples were used to confirm the presence or absence of these fragments.

model yielded classifications of 96% for red spruce and 99% for black spruce. These classifications are my best estimates of the accuracy of this model for classification of field samples.

The allelic frequencies were also analyzed using canonical discriminant analysis. The canonical discriminant analysis uses a dimension-reduction technique related to principal component analysis (PCA) and canonical correlation (and is less sensitive to error variance than PCA). One significant discriminant vector was produced ($p=0.0001$) useful in discriminating between the species. Figure 4 illustrates the clustering of the two species observed when canonical discriminant analysis is applied to nuclear rDNA allelic frequencies. Five red spruce individuals are classified as black spruce. One of the five red spruce individuals had RFLP frequencies (HMW 3.4=0.00, HMW 1.6=0.55, EMW 4.35=0.72, SMW 1.7=64 and SITS=0.95) that were very similar to the black spruce frequencies, and this individual plots within the black spruce cluster (Figure 4). The four other red spruce individuals had RFLP frequencies that ranged between the mean frequencies for red spruce and black spruce (Table 3, Appendix 1) and plot between the red and black spruce clusters (Figure 4). The black spruce individual misclassified in the discriminant analysis plots within the red spruce cluster (figure 4) and had RFLP frequencies (HMW 3.4= 0.00, HMW 1.6=0.05, EMW 4.35=0.64 SMW 1.7=0.68 and SITS=0.87) that were similar to red spruce, particularly for the two strongest univariate predictors, SITS and HMW 1.6. The individuals, both red and black spruce, which plotted within the other species' cluster may result from the violations of the variance equality criterion or to prior between-species hybridization events.

Morphological and isozyme analysis have been used in the past to distinguish between red and black spruce. Gordon (1976) presented models, based on 14 to 24

morphological traits, observed on over 10,000 spruce samples, useful for distinguishing red from black spruce. Reclassification of parental species based on isozyme analysis, 97% for black spruce and 92% for red spruce (Eckert 1989), was comparable to the results presented in this study of rDNA allelic frequencies. The red spruce and black spruce provenance samples used in the isozyme study are the same as those presented in this study.

Discriminant modeling based on rDNA alleles produced a powerful analytical technique for distinguishing Picea species. Necessary morphological data may not always be available for all individuals during field surveys; whereas with DNA analysis, all that is needed are small samples of needles for DNA extraction. In this study fresh growth needles were used because they give the highest yields; however, DNA may be extracted from older needles (data not shown). Classification of Picea species using the discriminant procedure applied to rDNA RFLP frequencies was at least as good as the isozyme model presented by Eckert (1989). Nine loci and 28 alleles were needed to produce the model based on isozyme variation. Using DNA markers, accurate classification was obtained using just five quantitative variables.

Allelic frequencies were also examined for within-species differentiation over geographic origins. Canonical correlation analysis of the molecular rDNA markers and the geographic origin of provenances revealed significant variation over geographic origins. Analysis of the black and red spruce samples revealed one significant canonical correlation for each species ($p=0.0015$ and $p=0.035$ respectively); the null hypothesis, that no relationship exists between the rDNA molecular markers and geographic origins, is therefore rejected. The correlation between geographic origins and nuclear rDNA allele frequencies indicates that the rDNA repeat units have diverged significantly among sub-

populations. Both species displayed significant variation in rDNA allelic frequencies with the variation occurring mainly over latitude; however, different allelic frequencies were differentiating for red and black spruce, EMW 4.35 and HMW 1.6 respectively (Table 4), suggesting that this variation is not due to introgression between species.

Differentiation of rDNA alleles appears to be occurring over latitude among populations of red spruce and among populations of black spruce. This differentiation may be due to genetic drift or caused by some selective forces acting on the rDNA repeat unit. Increased latitude corresponds to shorter growing season, lower temperatures, and harsher winters. Certain rDNA types may have some selective advantage over other rDNA alleles, for instance increased transcription resulting in increases in the ribosomal RNA pool and increased protein synthesis. My results suggests that there may be some sort of ecogeographic selection for rDNA types in Picea, a result similar to that postulated to occur in wild barley (Saghai Maroof et al. 1990).

Chapter 3

Identification of hybridization events between red and black spruce using nuclear and organelle DNA markers.

Results and Discussion

Organelle markers and distribution.

The Picea organelle genomes were screened for RFLPs that would be useful in classifying species and identifying hybridization events (Appendix 1). Polymorphic restriction fragments were identified for both the chloroplast and mitochondrial genomes. The distribution of these haplotypes was examined for within-species and between-species variation using provenance seed source samples.

Two chloroplast haplotypes were identified using the petunia clone (pST 6). One chloroplast haplotype consisted of a 20 kbp, 7.0 kbp, 3.0 kbp and 0.9 kbp RFLP pattern (cp-haplotype A, Figure 5). In another haplotype, the 7.0 kbp fragment was replaced by a 5.8 kbp fragment (cp-haplotype B, Figure 5). The chloroplast haplotypes were strongly associated with species but were not strictly species specific (Table 5). Within the black spruce, 95% of the individuals displayed cp-haplotype A (89 out of 94), whereas, cp-haplotype B was observed in 95% of the red spruce sampled (54 out of 57).

The controlled-cross hybrids were all formed using red spruce as a paternal parent. The red spruce-associated chloroplast haplotype (cp-haplotype B) was observed in 12 out of 13 of the hybrids (Figure 5, Table 5), suggesting that the chloroplast is paternally inherited. This result is in agreement with observations made for other gymnosperms (Neale, Marshall and Sederoff 1989) including Engelmann spruce (Picea engelmannii Parry ex Engelm.) and blue spruce (Picea pungens Engelm.)(Stine and Keathley 1990).

Five black spruce individuals had the red spruce-associated chloroplast haplotype (cp-haplotype B). Four of these individuals were within the range of red spruce

Figure 5. Southern blot analysis revealed a Sma I polymorphism in the chloroplast genome of red and black spruce. Two distinct banding patterns were evident: One pattern, cp-haplotype A included a 7.0 kbp restriction fragment which was replaced by a 5.8 kbp restriction fragment in the second pattern. The sample identification numbers (Appendix 1) were used to label each lane.

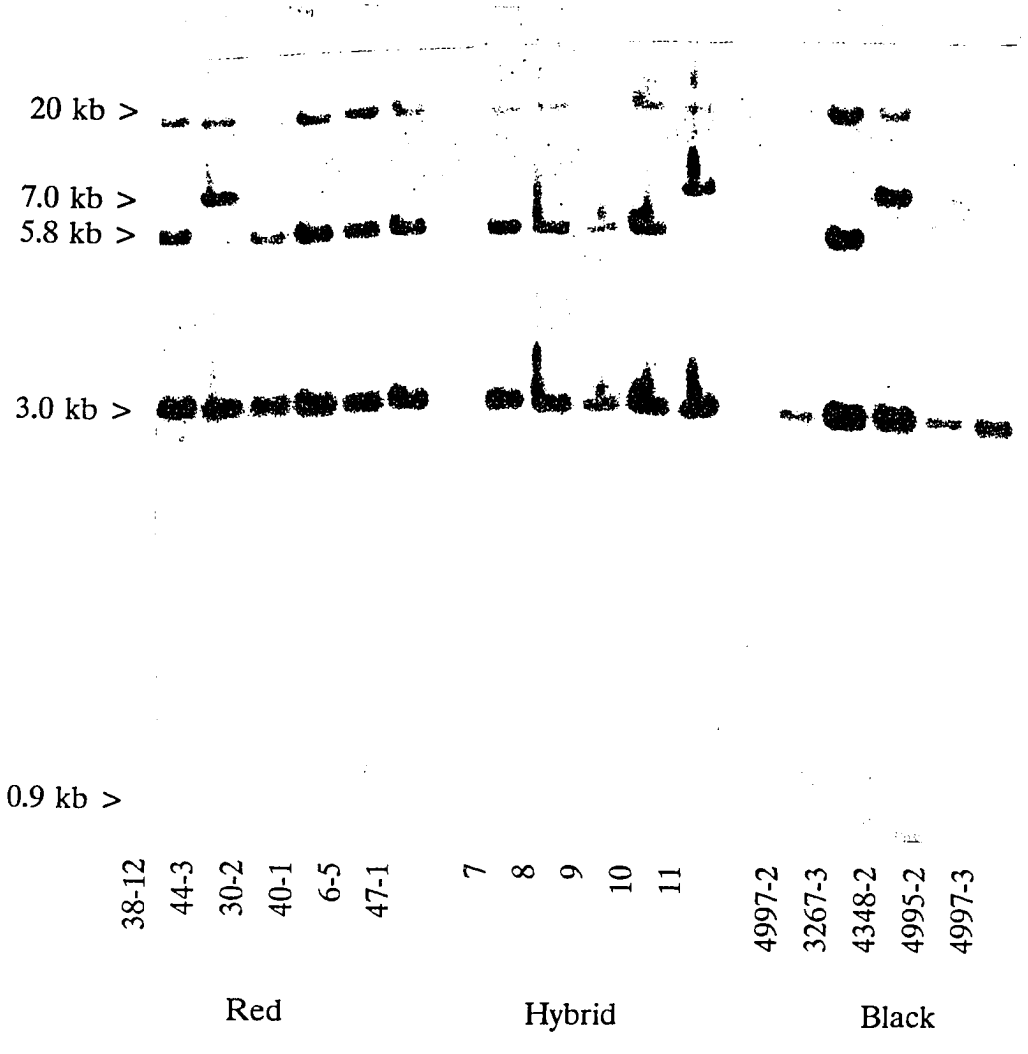


Table 5. Distribution of chloroplast and mitochondrial molecular markers over red spruce, black spruce and hybrid samples.

	Chloroplast haplotype		Mitochondria haplotype		
Species	A	B	A	B	C
Black	89	5	28	64	2
Hybrid	1	12	0	13	0
Red	3	54	51	6	0

(Figure 6), the fifth individual was located in a more northern provenance (latitude 49.01 longitude 55.26) outside of the red spruce range. Three red spruce samples had the black spruce-associated chloroplast haplotype (cp-haplotype A). All three individuals were within the range of black spruce (Figure 6).

Three mitochondrial haplotypes were identified using the maize mitochondrial clone. One haplotype was identified by a 9.6 kbp and 7.2 kbp RFLP pattern, (mt-haplotype A, Figure 7) and was observed in 89% (51 out of 57) of the red spruce samples (Table 5, Figure 6). A second haplotype (mt-haplotype B, Figure 7), observed in 67% (64 out of 94) of the black spruce (Table 5, Figure 6), was identified by a 9.6 kbp, 5.6 kbp and 5.0 kbp RFLP pattern. A third haplotype (mt-haplotype C) was observed in only two black spruce individuals (Table 5) and displayed a 9.6 kbp, 6.7 kbp, and 3.95 kbp RFLP pattern. A control, using the petunia chloroplast rDNA region as a probe (pST 6), revealed that the 9.6 Kbp fragment is due to cross-hybridization between the maize mitochondrial rDNA and the spruce chloroplast rDNA.

The black spruce-associated mitochondrial haplotype (mt-haplotype B) was observed in all of the hybrids (Figure 7, Table 5). This suggests that the mitochondria is maternally inherited in red and black spruce. This is the predominant inheritance pattern observed in most higher eukaryotes including gymnosperms (Neale, Marshall and Sederoff 1989).

All six red spruce individuals displaying the black spruce-associated haplotype (mt-haplotype B) originated from within the black spruce range (Figure 6). Four of the six individuals were from the same southern red spruce provenance (latitude 40.45 longitude 79.50). The mt-haplotype A pattern was observed in 30% of the black spruce samples and occurred throughout the black spruce provenances (Figure 6). Using linear regression

Figure 6. Distribution of mitochondrial haplotypes over geographic origin of provenance samples of red spruce and eastern range of black spruce. The ranges of red and black spruce are indicated on the map. Red spruce provenances are represented by squares and black spruce provenances by circles. Empty symbols represent provenances displaying only mt-haplotype A, solid symbols represent provenances displaying only mt-haplotype B, and checkered symbols represent provenances which displayed both mitochondrial haplotypes. One black spruce provenance displayed a third polymorphism (mt-haplotype C) and is represented by a stripped circle.

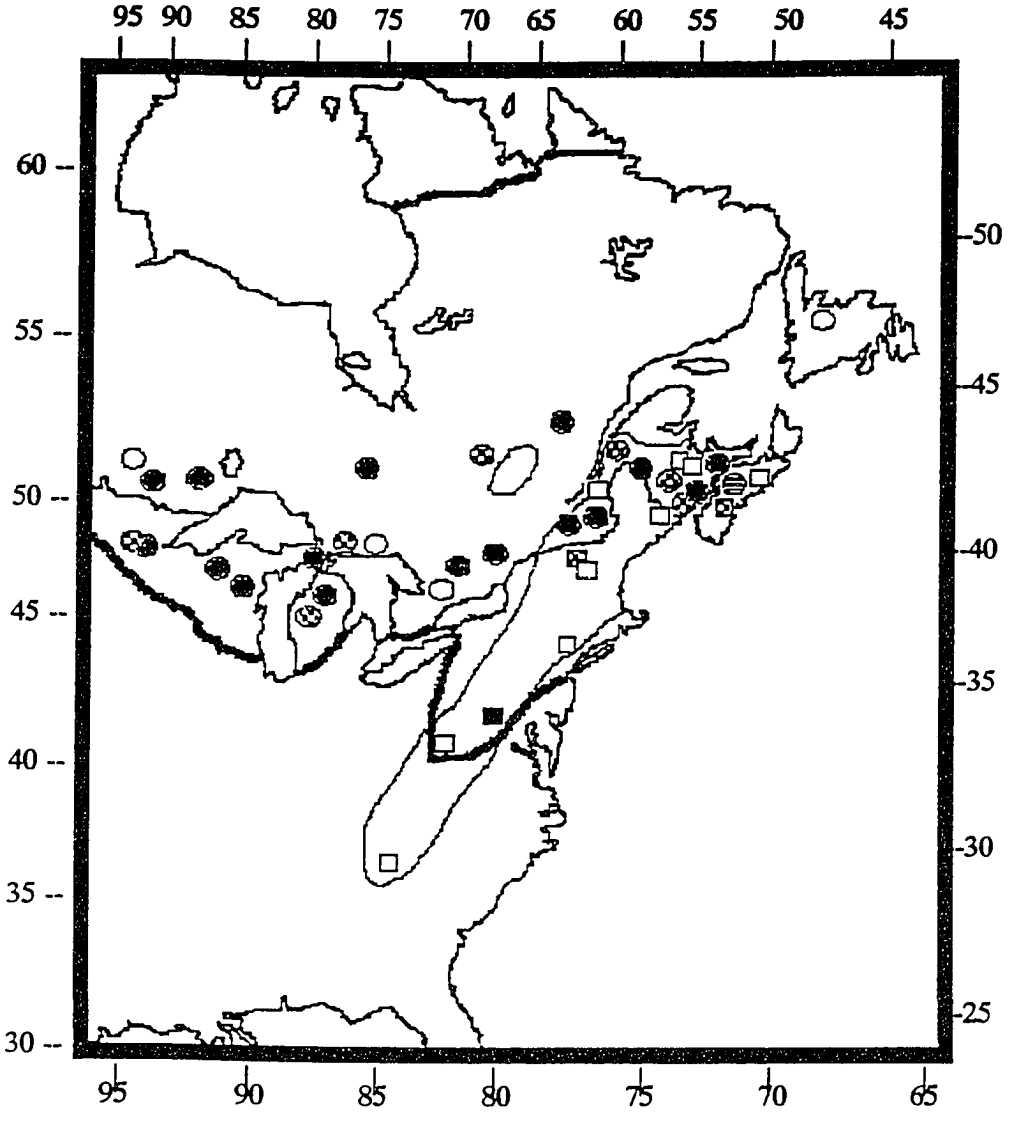
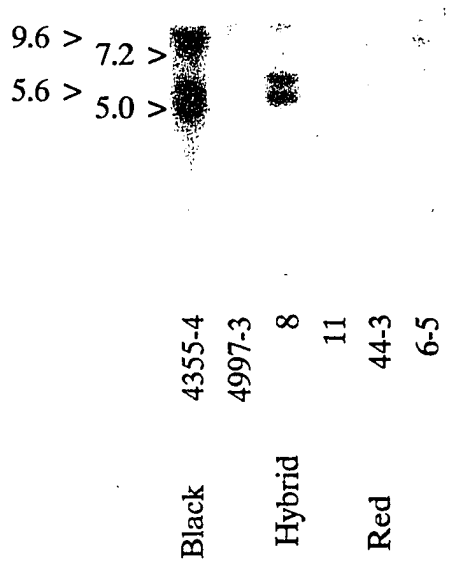


Figure 7. Southern blot analysis revealed three polymorphic Hind III RFLP patterns in the mitochondria genome for Picea. The first polymorphic pattern included a restriction fragment of 7.2 kbp in length (mt-haplotype A). A second pattern (mt-haplotype B) lacked the 7.2 kbp restriction fragment but included a 5.5 kbp and a 4.5 kbp restriction fragment. The third banding pattern (mt-haplotype C, not shown in this figure) consisted of a 9.6 kbp a 6.7 kbp and a 3.95 kbp restriction fragment pattern. A control, using the petunia chloroplast rDNA region as a probe (pST 6), revealed that the 9.6 Kbp fragment is due to cross-hybridization between the maize mitochondrial rDNA and the spruce chloroplast rDNA.



analysis no statistically significant relationship between mitochondrial haplotypes and geographic origins of provenances was detected. The occurrence of this haplotype in provenances outside of the red spruce range may represent past introgression events occurring during glaciation events or may represent within population variation, in black spruce, rather than hybridization and/or introgression.

Classification of Picea

The internal frequency of five nuclear rDNA restriction sites was measured for controlled cross red-on-black hybrids and for red and black spruce provenance samples. Figure 2 represents sample autoradiograms displaying variation in the nuclear rDNA markers. The relative intensities of the autoradiograph bands correspond to the frequency of different rDNA alleles within the genome of each sample (Chapter 2). A discriminant function based on the internal frequencies of the five nuclear rDNA RFLPs generated an index capable of discriminating red and black spruce (Chapter 2). This index could not, however, distinguish the hybrids from black spruce (data not shown).

Addition of the organelle markers allowed for accurate identification of hybrids. A discriminant function was derived by combining the chloroplast, mitochondria, and nuclear rDNA allelic data which could differentiate red spruce, black spruce and red-on-black hybrids (Wilks' Lambda=0.080, $p \leq 0.0001$). The discriminant function accurately reclassified red spruce (93%), black spruce (94%) and red-on-black hybrids (92%).

Examination of the univariate between-species variation in the nuclear rDNA allelic frequencies (Table 6) suggested that a number of the nuclear markers could be eliminated without compromising the discriminant model. With the organelle haplotype data added to the model, the accuracy of the discriminant model was preserved after

Table 6. Means, standard deviation and univariate between group F statistic for the nuclear rDNA restriction site frequencies for red spruce (R), black spruce (B), and red-on-black hybrids (H).

Site	Mean (S.D.)	Univariate between group F-ratio (P)
HMW 3.4	B 0.023 (0.070)	3.7 (0.027)
	H 0.000 (0.000)	
	R 0.000 (0.000)	
HMW 1.6	B 0.512 (0.153)	91.3 (0.0001)
	H 0.461 (0.068)	
	R 0.182 (0.136)	
EMW 4.35	B 0.671 (0.256)	40.0 (0.0001)
	H 0.498 (0.119)	
	R 0.323 (18.9)	
SMW 1.7	B 0.582 (0.085)	9.0 (0.0002)
	H 0.589 (0.0689)	
	R 0.526 (0.073)	
SITS	B 0.997 (0.016)	196.6 (0.0001)
	H 0.942 (0.061)	
	R 0.768 (0.106)	

removing four of the nuclear markers (HMW 3.4, HMW 1.6, SMW 1.7, and EMW 4.35). Reclassification of the spruce samples using one nuclear marker (SITS) and the chloroplast and mitochondria markers was very accurate for red spruce, black spruce and hybrids (Wilks' Lambda = 0.096[1], $p \leq 0.0001$) (Table 7). Misclassifications may be due to limitations in our statistical model or due to the presence of hybrids and/or introgressed individuals within the provenance seed sources, as has been suggested by Morgenstern and Farrar (1964), Fowler (1969), Morgenstern (1969), Roche (1969), Manley and Ledig (1978), Morgenstern et al. (1981).

Correct classification of species and hybrids can be accomplished without the use of complicated statistical procedures. A three character index was derived using only one nuclear marker and the organelle markers (Figure 8). Among the nuclear markers, SITS displays the highest univariate between-species variation (Table 6). Hybrids tend to have intermediate frequencies of the nuclear rDNA markers, particularly the marker SITS (Figure 1, Figure 2 C). The hybrids used in this study were all red-on-black hybrids and 12 out of 13 had a chloroplast RFLP patterns closely associated with red spruce and all of the hybrids have the mitochondrial pattern associated with black spruce (Table 5).

Accurate classification of pure black spruce can be accomplished through examination of the nuclear Sac I marker, SITS, and the organelle markers (Figure 8). An internal frequency of less than or equal to 0.99 for the marker SITS suggest that the individual is not a pure black spruce. The internal frequency of the polymorphic Sac I site, labeled SITS in figure 1, was 1.00 in 95% (92 out of 97) of the black spruce surveyed. One of the black spruce individuals (#4962-3) had RFLP patterns more closely associated

---

[1] Because of the non-normal distribution of molecular markers within species, particularly SITS within black spruce, the accuracy of statistical values, such as Wilks' Lambda may be called into question. This would be of concern if the frequency of misclassification was high.

Table 7. Reclassification of provenance plantation individuals using both the three character discriminant model (D)[1] and three character index (I).

SP		Pure black spruce	Hybrid/ introgressed	Pure red spruce
Black spruce	D	95%	4%	1%
plantation	I	93%	6%	1%
Controlled-cross	D	8%	92%	0%
hybrids	I	0%	100%	0 %
Red spruce	D	5%	2%	93%
plantation	I	0%	14%	86%

[1] The discriminant model presented here is only sensitive to red on black F1 hybrids and not sensitive to introgression and black on red hybrids.

Figure 8. Three character classification index. The nuclear marker SITS is not polymorphic within black spruce and is reliable for distinguishing between red and black spruce; however, it is not dependable when identifying hybridization/introgression events. The chloroplast polymorphism is a dependable indicator of paternal lineage. An individual displaying the black spruce haplotype (cp-haplotype A) in combination with a polymorphic SITS site is classified as introgressed. The mitochondrial polymorphism is useful when identifying red-on-black hybrids, an individual displaying the red spruce-associated chloroplast haplotype (cp-haplotype B) with either mt-haplotype B or C is classified as hybrid/introgressed. Because mt-haplotype A was found in 30% of the black spruce surveyed, some red-on-black hybrids may be misclassified as red spruce. Examining within stand variation in nuclear and organelle markers enhances the accuracy of this index.

Marker	Black spruce	Hybrid/Introgressed		Red spruce
		Black-on-Red	Red-on-Black	
Chloroplast RFLP pattern	cp-haplotype A	cp-haplotype A	cp-haplotype B	cp-haplotype B
Mitochondrial RFLP pattern	mt-haplotype A, B or C	mt-haplotype A (for F1 hybrids)	mt-haplotype A, B or C	mt-haplotype A
Internal frequency of the nuclear SITS rDNA marker	SITS > 0.99	Intermediate levels of SITS (as compared to natural population) 0.99≤SITS≤0.80 for F1 hybrids		SITS ≤ 0.95

Table 8. Provenance plantation and hybrid individuals misclassified by the three character discriminant model and/or the three character characterization model. The three character discriminant function is only sensitive to red-on-black hybrids. The three character model is sensitive to hybridization and introgression in both directions. ID is sample origins (controlled-cross hybrid (H), or provenance plantation at Coleman (C) Massabesic (M)). Allelic frequencies are given as percent of total. Individuals displaying RFLP patterns suggestive of a past hybridization event but do not appear to be an F1 hybrid are referred to as introgressed (I).

Provenance seed source	ID	HMW 1.6	SITS	mtRFLP	cpRFLP	Discriminant function	Three character index
4348(1)[1]	M	13	100	B	B	R-on-B	I
4355(1)	M	52	100	B	B	R-on-B	I
4959(2)	M	75	73	B	A	B	I
4962(3)	M	5	87	A	B	R	R
4986(1)	M	35	95	B	B	R-on-B	R-on-B
4993(2)	M	45	94	B	A	B	I
5004(3)	M	44	100	A	B	R-on-B	I
(11)	H	50	93	B	A	B	I
2019(16-11)	C	55	95	A	B	R-on-B	R
2021(6-1)	C	20	55	B	B	R	I
2021(6-6)	C	31	82	B	B	R	I
2021(7-9)	C	28	74	B	B	R	I
2021(44-33)	C	7	87	B	A	B	I
2024(31-20)	C	10	74	B	B	R	I
2101(47-10)	C	16	85	B	B	R	R-on-B
2103(2-1)	C	12	79	A	A	B	B-on-R
2103(40-2)	C	19	87	A	A	B	B-on-R

[1] Provenance seed source-(sample number)

with red spruce (Table 8). This individual was recollected in the summer of 1991 (2 years after the first collection) and reanalyzed. The nuclear rDNA frequencies and the organelle haplotypes were verified as correct. Analysis of the genotype of this sample suggests that it is either introgressed with a strong red spruce nuclear influence, a red spruce that was misidentified, or a genetic outlier. One sample had RFLP patterns associated with a red-on-black hybrid (Table 8). Six other individuals had RFLP patterns suggestive of past introgression event(s) (Table 8).

Red spruce individuals can also be identified by examining both the organelle and nuclear markers (Figure 8). All of the red spruce surveyed had an internal frequency of less than or equal to 0.95 for SITS, suggesting that none of the individuals were pure black spruce. The mt-haplotype A and cp-haplotype B patterns were observed in 86% of the red spruce individuals. One southern provenance (provenance 2021, latitude 40.45 longitude 77.45, Table 8, Figure 6) lacked the mt-haplotype A pattern but had SITS frequencies similar to those observed in red spruce and was previously reported to be introgressed (Wilkinson 1990). Three of the four red spruce individuals from this provenance were reclassified as red spruce using the four character discriminant analyses; therefore, it seems probable that at least these three individuals are not F1 hybrids. They may, however, be introgressed. The fourth individual displayed mt-haplotype B and cp-haplotype A, both associated with black spruce; however, it had an internal frequency of 0.87 for the nuclear rDNA marker SITS suggesting that this individual is also introgressed. One red spruce sample displayed RFLP patterns suggestive of a red-on-black hybridization event. Two other spruce from the Coleman plantation displayed RFLP patterns suggestive of black-on-red hybridization events (Table 8).

Using the three character index (Figure 8) to distinguish red spruce, black spruce,

and hybrids provides a classification technique comparable with more complicated statistical procedures (Table 7 and 8). The increase in the number of hybrid/introgressed identified with the three character index is not surprising because it is sensitive to introgression and hybridization events in both directions. The discriminant model is only sensitive to red-on-black F1 hybrids.

Six of the seven black spruce individuals that were identified as hybrid/introgressed (Table 8), using the three character index (Figure 8), originated from regions where introgression has been previously reported (Morgenstern and Farrar 1964; Wilkinson 1990). One of the individuals (4962(3)) from the black spruce provenance plantation was classified as red spruce and originated from a provenance outside of the red spruce range. This individual may have resulted from a past introgression event or may be a genetic outlier. Seven of the red spruce individuals that were identified as hybrid/introgressed (Table 8) originated from provenances previously reported to contain introgressed individuals (Morgenstern and Farrar 1964; Wilkinson 1990). One red spruce individual (2024(31-20), Table 8) classified as hybrid/introgressed with the three character index, originated from a provenance reported to be pure red spruce. This individual displayed a strong red spruce nuclear influence; however, had a black spruce mitochondrial haplotype and was consequently identified as introgressed.

Because of the widespread presence of mt-haplotype A in the black spruce provenance samples it is advisable to examine the RFLP patterns within natural populations when evaluating hybridization and introgression. The possibility of extensive hybridization and gene flow between red and black spruce during glaciation events may account for the presence of mt-haplotype A in the black spruce provenances; however, within-species variation cannot be ruled out by the results presented here. Because the

second possibility cannot be ruled out the index presented here relies on the nuclear rDNA marker, SITS, for identifying red-on-black hybrids in populations which do not display species specific mitochondrial haplotypes. Examining RFLP patterns within natural populations will enhance the sensitivity of the molecular indexes described here.

Chapter 4

Red spruce population variation along elevational transects on Mount Washington and Mount Lafayette.

Results

High levels of hybridization/introgression were reported between red and black spruce on Mt. Washington (Berlyn et al. 1990). Using a controversial hybrid index, modified from Manley (Manley 1971, as described in Berlyn et al. 1990), and nuclear genome content, introgression was reported up two elevational transects on Mt. Washington. Berlyn et al. (1990) concluded that hybridization between red and black spruce was frequent and that levels of introgression were related to elevation.

In this study, samples collected from Mts. Washington and Lafayette were examined for within-species and between-species variation. The internal frequencies of three nuclear rDNA markers, HMW 1.6, EMW 4.35 and SITS (Figure 1), were quantitated for all individuals sampled from Mount Washington and Mount Lafayette (Appendix 1). Organelle haplotypes, haplotype A and B for the chloroplast and haplotype A, B, and C for the mitochondria (Chapter 3), were also scored for all individuals. RFLP variation was examined to classify trees, determine the extent of hybridization on these two mountains, to evaluate elevation effects on red spruce rDNA allelic frequencies, and to compare red spruce populations on these two mountains.

Distribution of organelle markers.

Organelle haplotypes were scored for all individuals on Mount Washington and Mount Lafayette. The chloroplast haplotype (cp-haplotype B), associated with black spruce, was observed in only one individual on Mount Lafayette at an elevation of 708 m. On Mount Washington 12 individuals displayed the cp-haplotype B: one individual

collected at 790 m, three individuals at 1220 m, and eight individuals at 1433 m. One sample on Mount Lafayette, elevation 1430 m, had mt-haplotype B pattern, all other individuals displayed mt-haplotype A associated with red spruce. On Mount Washington all individuals had mt-haplotype A.

Classification of Picea individuals

Samples from both Mts. Washington and Lafayette were classified into species using the molecular RFLP models developed in chapter 3. The three character discriminant model, based on the organelle haplotypes and the internal frequencies of one nuclear rDNA markers (SITS), is useful when identifying species and red-on-black F1 hybrids; however, is not sensitive to black-on-red hybrids. The three character molecular index, based on the internal frequency of SITS and the organelle haplotypes, is useful for classifying individuals and identifying hybridization events in both directions (Chapter 3).

Black spruce were not detected among the samples surveyed from Mt. Lafayette. Discriminant classification analysis, based on the three molecular markers, classified only one individual as hybrid on Mt. Lafayette (sample 1103 Table 9). Applying the three character molecular index to these samples, three trees displayed RFLP patterns that are suggestive of a past hybridization event (Table 9). Because of the strong red spruce nuclear influence in two of these samples (1150 and 1123 Table 9), they are probably introgressed and not F1 hybrids.

Red spruce, black spruce, and hybrids were identified on Mount Washington. Discriminant classification analysis based on morphology of the spruce samples collected along the elevation transect on Mt Washington indicated that 64 (93%) were red spruce, 5 (7%) were hybrids, and 1 (1%) was black spruce (K. Stapelfeldt, R. Eckert personal

Table 9. Spruce individuals from Mt. Washington (W) or Mt. Lafayette (L) classified as hybrid or introgressed with the three character index.

Mts	ID#	Elevation	SITS	cp-haplotype	mt-haplotype	Type of cross
L	1150	710	53	A	A	Black-on-Red
L	1103	1190	99	B	A	Red-on-Black
L	1123	1430	65	B	B	Red-on-Black
W	31	790	100	B	A	Red-on-Black
W	46	790	98	A	A	Black-on-Red
W	77	1220	98	B	A	Red-on-Black
W	79	1220	98	B	A	Red-on-Black
W	80	1220	100	B	A	Red-on-Black
W	89	1220	84	A	A	Black-on-Red
W	95	1435	75	A	A	Black-on-Red

communication). Morphology was damaged by severe weather conditions at high elevations, increasing the probability of error in recording color and twig-ridge shape variables. A number of the black spruce individuals collected at the elevation of 1430 m were not classified using the morphological model due to damage to the samples. Most morphological samples which classified as hybrids, and the black spruce, were located above 1220 m in elevation. One sample which classified as a hybrid was located at 790 m. There was no strong relationship between elevation of sample and species identity, in contrast to observations made by Berlyn et al. (1990).

The three character molecular index generated species classifications similar to the morphological model: 64 (80%) were red spruce, 7 (9%) were hybrid/introgressed, and 9 (11%) were black spruce. Table 9 displays the molecular data obtained for the hybrids identified on the two mountains using the three character molecular index. The three character discriminant model provided results similar to the other indexes except the molecular discriminant model is insensitive to black-on-red hybrids classifying them as black spruce: 65 (81%) were identified as red spruce, 4 (5%) were identified as hybrid, and 11 (14%) were identified as black spruce. Table 10 compares the results from the different indexes used on the Mt. Washington samples.

Elevational trends in the nuclear rDNA allelic frequencies.

Regression analysis was used to examine the relationship between elevation and rDNA allelic frequencies in red spruce (Figure 9). No significant variation was observed when the data collected from both mountains were combined (data not shown); however, when samples from theses two mountains were analyzed separately, significant variation was detected on Mount Washington but not on Mount Lafayette.

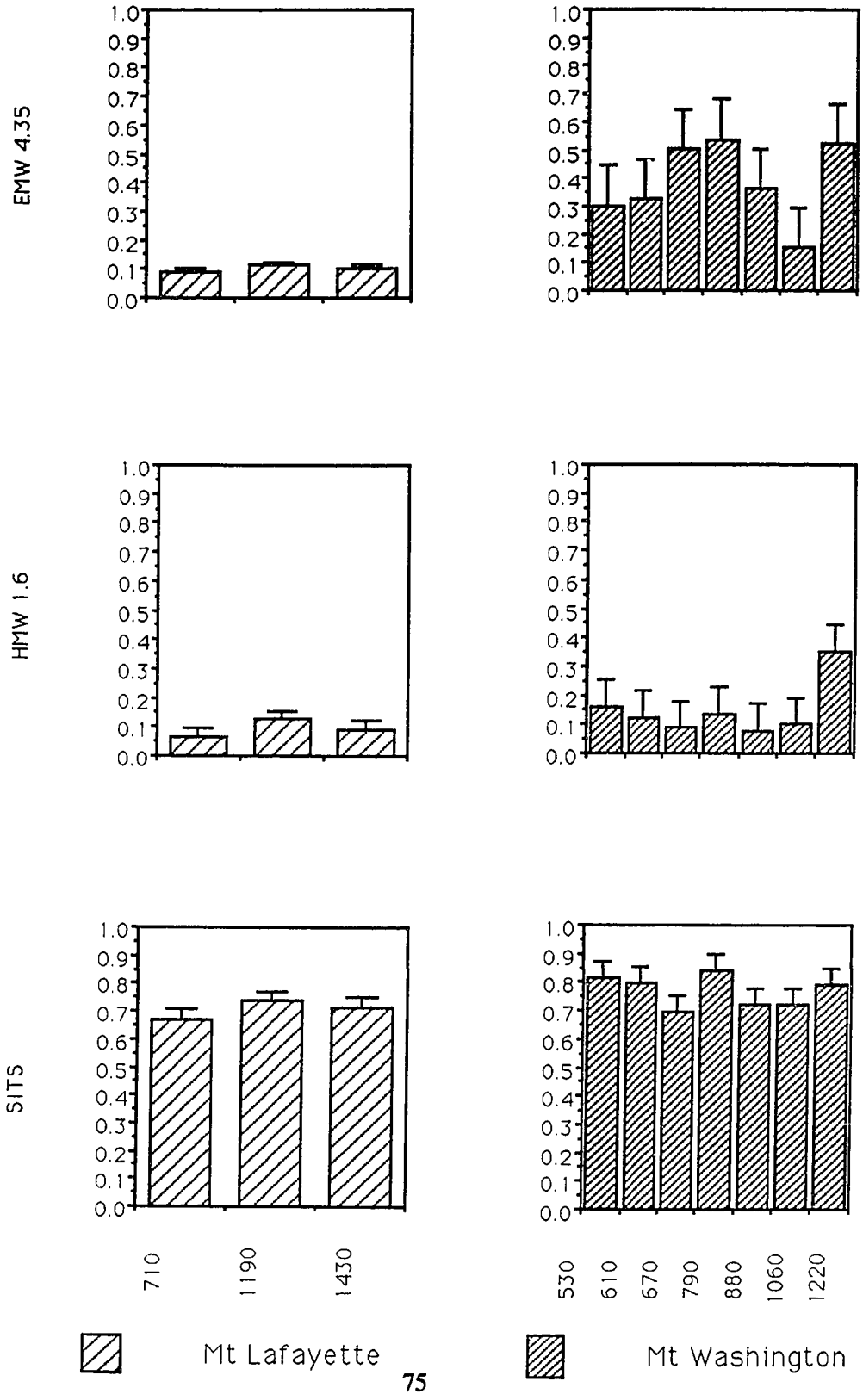
Table 10. Comparison among the morphological and molecular indexes applied to the Mt. Washington samples for trees where disagreements occurred among the analyses. The three character discriminant model is insensitive to black on red hybridization events, classifying black on red hybrids as black spruce.

	Three Character molecular index	Morphological discriminant index Posterior Probabilities[1]			Molecular discriminant index Posterior Probabilities[1]		
ID#	Direction of cross	Black	Hybrid	Red	Black	Hybrid	Red
46	Black on Red	-[2]	-	-	1.000	0.000	0.000
77	Red on Black	0.0000	0.0002	0.9998	0.0080	0.5669	0.4251
81	Black	0.0000	1.0000	0.0000	1.0000	0.0000	0.0000
87	Black	0.0000	1.0000	0.0000	1.0000	0.0000	0.0000
89	Black on Red	0.0000	0.9972	0.0028	0.9922	0.0000	0.0077
95	Black on Red	-	-	-	0.7451	0.0000	0.2549

[1] Posterior probabilities give the probability that an individual belongs to a species group based on the discriminant function (Klecka 1980).

[2] A dash indicates missing observation due to poor morphological condition of sample.

Figure 9. **Variation in the internal frequency of nuclear rDNA allelic markers up elevational transects of Mts. Washington and Lafayette.**



On Mount Washington significant ($p=0.02$) variation was observed modeling elevation squared with the nuclear rDNA allelic frequencies and their squared values. Examining each of the restriction sites separately revealed that variation in EMW 4.35 displayed a significant ($p=0.009$, adjusted $R^2=0.13$) relationship with elevation squared (Table 11), while the other alleles did not. Variation in EMW 4.35 was not linear but more a function of elevation squared (Figure 9). The nuclear rDNA allelic marker SITS is used to classify species and displays the strongest between-species variation ($F=197$, $p\leq 0.0001$ (Chapter 2)). I could not detect any significant relationship between allelic frequencies and elevation that could be owed to between-species introgression.

Variation between red spruce populations on Mts Washington and Lafayette.

Nuclear rDNA restriction site frequencies in red spruce were examined for population difference between Mount Washington and Mount Lafayette. The Mount Washington population displayed more overall variance in nuclear rDNA restriction site frequencies (Figure 9), particularly for the EcoR I polymorphism, EMW 4.35.

Canonical discriminant analysis was used to examine between mountain differences in red spruce populations (Figure 10). Applying canonical discriminant analysis to the red spruce nuclear rDNA allelic frequencies yielded one significant vector useful in separating the populations ($p\leq 0.0001$, Wilks' Lambda=0.622). The markers EMW 4.35 and SITS displayed significant between population variation (univariate $F=72.0$ and 14.0 respectively) while the marker HMW 1.6 did not ($F=0.4$). Using the canonical model, scores were calculated for each individual. Figure 10 displays the variation over the first canonical variate for Mount Washington and Mount Lafayette.

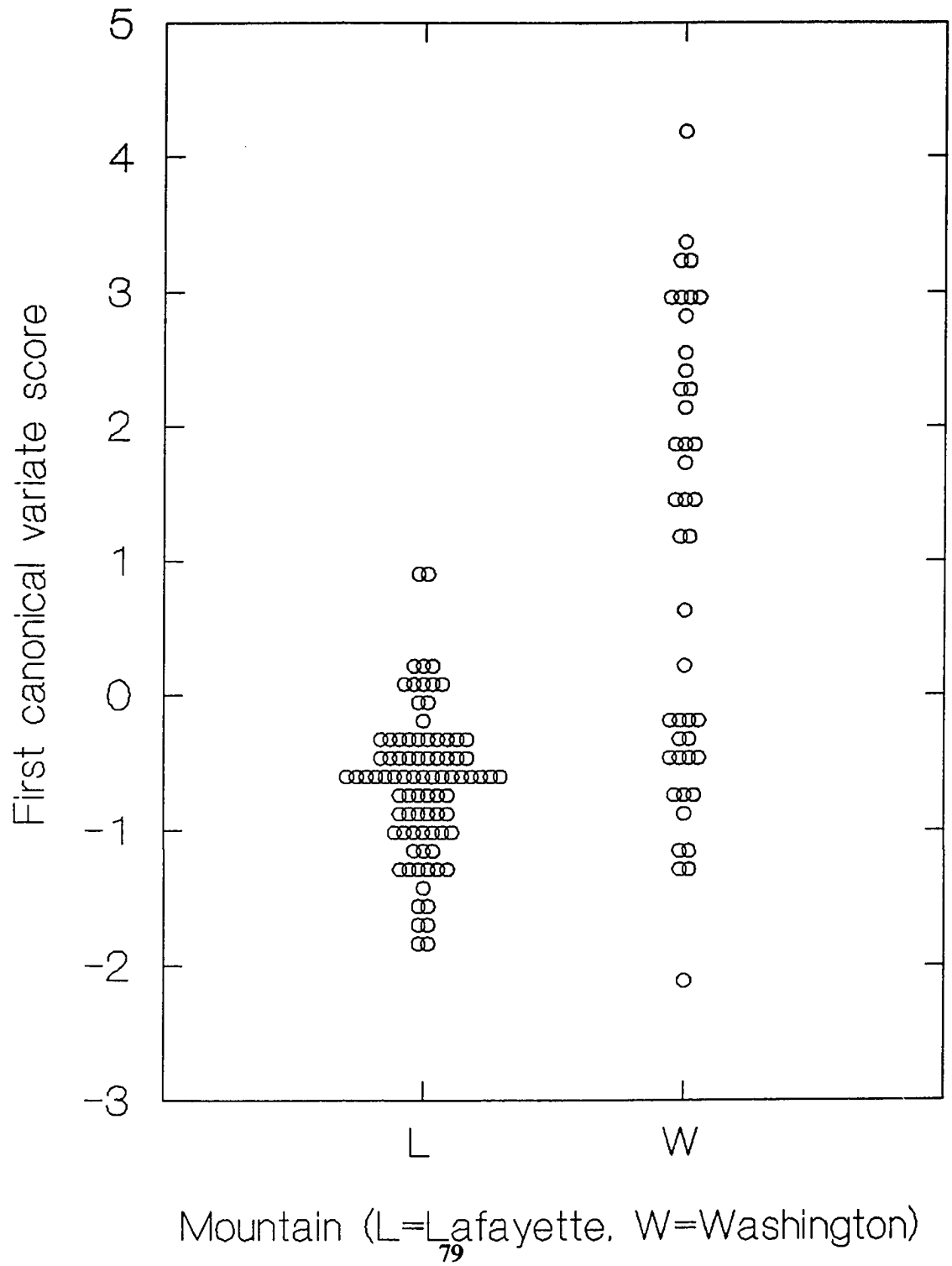
Table 11. Regression analysis of elevation squared on the nuclear rDNA allelic marker EMW 4.35 and EMW 4.35 squared on Mt. Washington.

Parameter	Degrees of Freedom	Standard Error	T test for Hypothesis[1]	Prob $\geq$ |T|
Intercept	1	877077	5.90	0.0001
EMW 4.35^2	1	676	-3.18	0.002
EMW 4.35	1	58016	2.98	0.004

R^2=0.16 Adj-R^2=0.13 p=0.009

[1] The T-test statistic measures the strength of the linear relationship between the dependent and independent variables.

Figure 10. Between population variation in nuclear rDNA alleles. Plot of discriminant canonical variate scores for red spruce individuals from Mts. Washington and Lafayette. The first canonical variate is statistically significant (p=0.0001, Wilk's Lambda=0.622). Plots of canonical scores illustrate the distribution of the two populations found when canonical analysis is applied to nuclear rDNA restriction fragment frequencies.



Discussion

Classification and analysis of the spruce populations on Mts. Washington and Lafayette are based on molecular and morphological techniques using models developed from provenance tests of red and black spruce. This approach provides a basis for classification of species identities and assessment of hybridization detectable through DNA analyses.

Three trees were identified as introgressed on Mt. Lafayette using the molecular index (Table 9). One of the three was classified as hybrid using the molecular discriminant classification model. The other two samples appeared to have a strong red spruce nuclear influence and were consequently classified as red spruce using the discriminant classification model. These results suggest that on Mt. Lafayette hybridization and introgression has occurred but at a low frequency; only 4% displayed RFLP patterns suggestive of hybridization/introgression.

On Mount Washington, red spruce, black spruce, and hybrids were identified using the three character molecular index and the two discriminant classification analyses, one based on morphology and one based on molecular markers. Using the three character index seven individuals were identified as hybrid/introgressed, two at 790 m, four at 1220 m and one at 1430 m. The molecular discriminant classification analysis was in agreement with the three character molecular index in identifying red-on-black hybrids; however, the molecular discriminant analysis is not sensitive to black-on-red hybrids and, as expected, misclassified these hybrids as black spruce. The molecular and morphological classification indexes were in close agreement for the majority of the samples (Table 10).

The mitochondrial haplotype (mt-haplotype B) associated with black spruce from

the provenance plantation samples, was lacking in the spruce populations from the two mountains. The chloroplast haplotypes, on the other hand, appeared to be closely associated with species within the two mountain populations. Thirty percent of the black spruce provenance samples surveyed in my previous study (Chapter 3) displayed the mitochondrial haplotype associated with red spruce (mt-haplotype A). Therefore, the presence of mt-haplotype A in the black spruce from Mount Washington is not surprising and probably represents genetic drift rather than introgression.

Berlyn et al. (1990) observed variation in nuclear DNA content up elevational transects of a number of mountains including Mount Washington. They reported that red spruce had twice the DNA content of the high elevation black spruce and that the DNA content decreased as elevation increased on Mount Washington (R^2=0.828). They concluded that the decrease in nuclear DNA content was due to hybridization and introgression and that the degree of introgression was related to elevation. I did observe hybridization on Mt Washington: at 1220 m 50% of the samples surveyed were identified as hybrid; however the other samples at this elevation identified as either red or black spruce. I did not observe the high degree of introgression up the eastern slope of Mount Washington or the strong relationship (R^2=0.828) between elevation and introgression that was previously reported on Mount Washington (Berlyn et al. 1990). It is possible that the observed variation in nuclear genome content was due to some elevational effect rather than a true species difference, similar to the variation I have observed in the nuclear rDNA allelic marker, EMW 4.35.

Nuclear rDNA restriction site frequencies were examined for elevational trends. On Mount Lafayette, no significant relationships were detected between rDNA alleles and

elevation. There was, however, a small (Adj-R^2=0.13) but significant (p=0.009) relationship between rDNA alleles and elevation (Table 11) on Mount Washington. The observed variation in red spruce nuclear rDNA markers does not appear to be due to hybridization and/or introgression with black spruce. The allele displaying the variation (EMW 4.35) was the same marker that showed the strongest relationship with geographic origins of red spruce provenance plantation samples (Chapter 2). The variation in the marker EMW 4.35 may be due to some ecogeographical selection similar to the selection for rDNA alleles observed in barley (Saghai Maroof et al. 1990). The tendency for this spruce rDNA allele to display the strongest relationship (among the different alleles) with geographic parameters, both within population and among provenances, suggests rDNA allele frequencies are subject to some type of selection. On Mount Washington the observed relationship is small (Adj-R^2=0.13), suggesting that any selective force is probably weak. The allelic marker, EMW 4.35, is displaying selection whereas the other nuclear rDNA markers (SITS and HMW 1.6) do not, suggesting that the frequencies of the different rDNA allelic markers are varying independent of one another.

The Mount Washington population displayed greater variation compared to the Mount Lafayette population (Figure 9 and 10). Canonical discriminant analysis was utilized to examine differences in the genetic makeup between these two populations. One significant vector was derived which could distinguish the populations ($p \leq 0.0001$, Wilks' Lambda=0.622), with the markers SITS and EMW 4.35 displaying a significant contribution to the model. The low genetic diversity observed in samples collected from a northwestern transect of Mount Lafayette could be due to genetic drift caused by a severe population decline or to founder effects. Alternatively, the differences may be due to different selection pressures on the two populations.

The genetic structure of these two mountain populations were dissimilar. Black spruce were observed on the upper elevations of Mt Washington but not Mt. Lafayette. Hybridization between red and black spruce was identified on Mt. Washington at 1220 m however no relationship between elevation and introgression were evident. Elevational trends in rDNA allelic variation, in red spruce, were observed on Mt. Washington but not on Mt. Lafayette. Finally the red spruce population on Mt. Washington displayed a higher level of variability, in rDNA alleles, in comparison with red spruce population sampled from Mt. Lafayette.

Chapter 5

High frequency hybridization and introgression between red and black spruce on Isle au Haut, Maine.

Results

Hybridization and introgression was previously reported to occur between black spruce, localized to boggy sites, and red spruce located upland of the black spruce (Osawa 1986; Thorpe 1986; Osawa, Spies, Dimond 1986). I applied my molecular classification index to samples obtained from similar habitats to examine interactions between the two species and to evaluate relationships between habitat and hybridization frequency.

Picea individuals were sampled from Head Harbor, Isle au Haut, ME, to examine between-species hybridization/introgression. A number of hybrids were previously identified, based on morphology, in a area between a black spruce stand, located within a bog, and an upland red spruce stand (Smoot Major, personal communication).

Nuclear rDNA allelic frequencies, EMW 4.35, HMW 1.6, and SITS, were measured and organelle haplotypes were scored for all individuals collected from Isle au Haut (Appendix 1). A total of 87 spruce trees were sampled: 26 individuals from within the bog (site 2), 18 individuals from a stand located between the bog and the ocean (site 3), 17 individuals upland and inland from the bog (site 4), and 26 individuals from the putative hybrid habitat (site 1). Molecular markers were examined to determine the degree of between-species hybridization/introgression, to determine if habitat influences hybridization, and to examine variation in nuclear rDNA alleles.

Distribution of organelle haplotypes.

Organelle haplotypes were scored for all individuals sampled from Isle au Haut. The chloroplast haplotypes, cp-haplotype A (associated with black spruce) and cp-haplotype B (associated with red spruce) were both observed in the samples. The

distribution of the chloroplast haplotypes over the different test sites is displayed in table 12. All individuals sampled from Isle au Haut displayed the red spruce associated mitochondrial haplotype, mt-haplotype A.

Classification of Picea individuals from Isle au Haut

Individuals were classified into species using the molecular index previously described (Chapter 3). Using the molecular index 37 individuals (43%) were classified as hybrid, 28 individuals (32%) were classified as black spruce, and 23 individuals (25%) were classified as red spruce. Thirteen of the hybrids (15% of the total) displayed RFLP patterns suggestive of black-on-red hybrid and twenty three (28% of the total) of the hybrids displayed RFLP patterns suggestive of red-on-black hybrids. There was a disproportionate degree of hybridization among the sites with enhanced levels of hybridization occurring between the bog and upland site and between the bog and the ocean (Table 13).

Variation in nuclear rDNA alleles between species.

Three nuclear rDNA markers were examined for between-species variation; EMW 4.35, HMW 1.6 and SITS (Figure 1). Within the Isle au Haut samples, the marker HMW 1.6 displayed higher internal frequencies within the red spruce and lower internal frequencies within the black spruce in comparison to observations made for provenance plantation samples (Figure 11). The frequency of the rDNA allelic marker, EMW 4.35, displayed low frequencies within the Isle au Haut samples (Figure 12). The allelic marker, SITS, used for classification of species displayed a distribution similar to that observed in the provenance plantation samples (Appendix 1).

Table 12. Distribution of chloroplast haplotypes over test sites on Isle au Hault.

Site	cp-haplotype A	cp-haplotype B
1	17 (65%)	9 (35%)
2	19 (73%)	7 (27%)
3	1 (6%)	17 (94%)
4	4 (24%)	13 (76%)

Table 13. Classification table for Isle au Haut Samples.

Test site	Pure black spruce	Black-on-red hybrid/ introgressed	Red-on-black hybrid/ introgressed	Pure red spruce
1	9 (35%)	8 (31%)	5 (19%)	4 (5%)
2	17 (61%)	2 (8%)	3 (12%)	4 (5%)
3	1 (5%)	0 (0%)	12 (67%)	5 (28%)
4	1 (6%)	3 (18%)	3 (18%)	10 (56%)

Figure 11. Within class distribution of the nuclear rDNA allelic marker HMW 1.6 on Isle au Haut: b=black, h=hybrid and r=red.

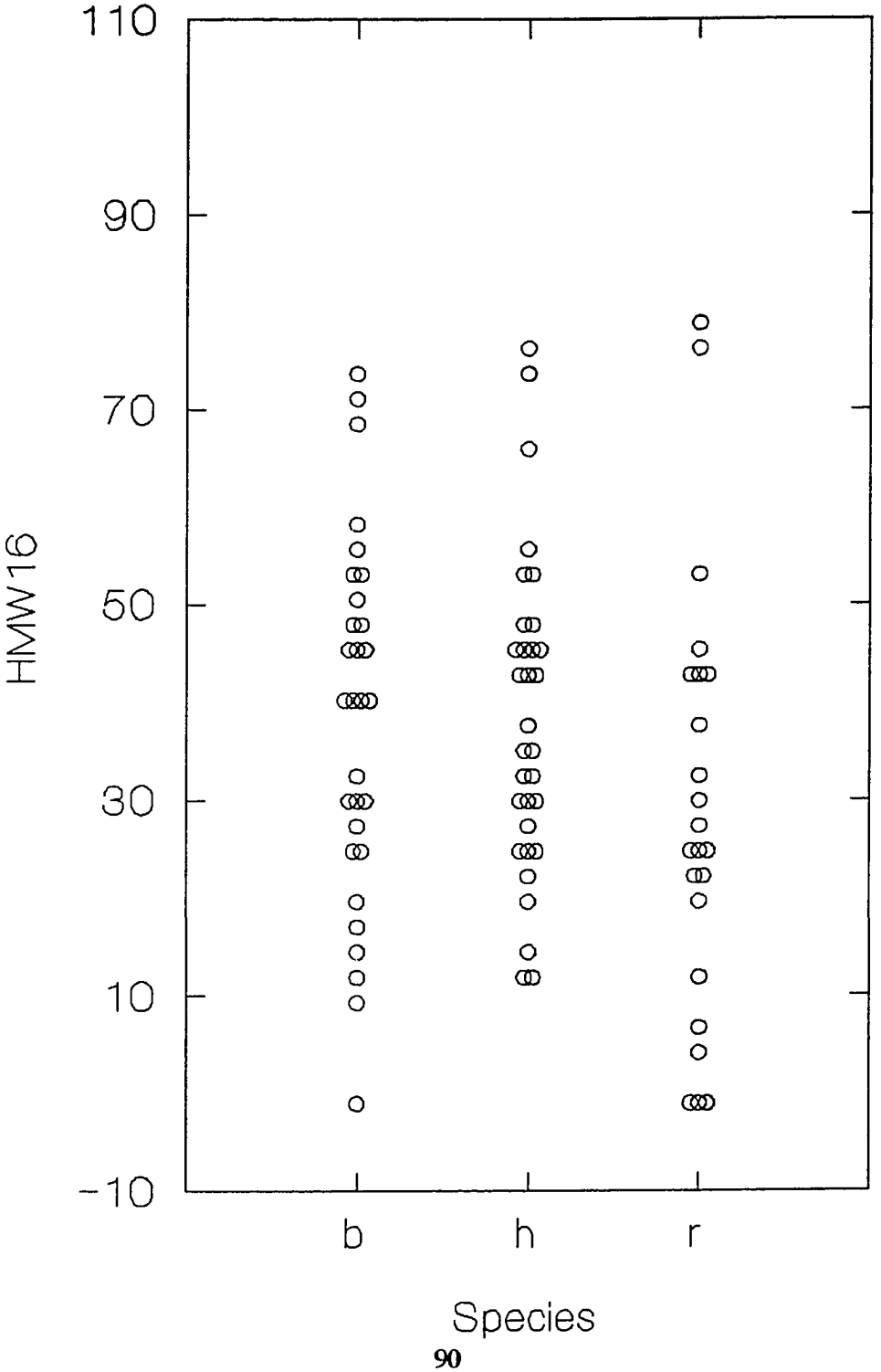
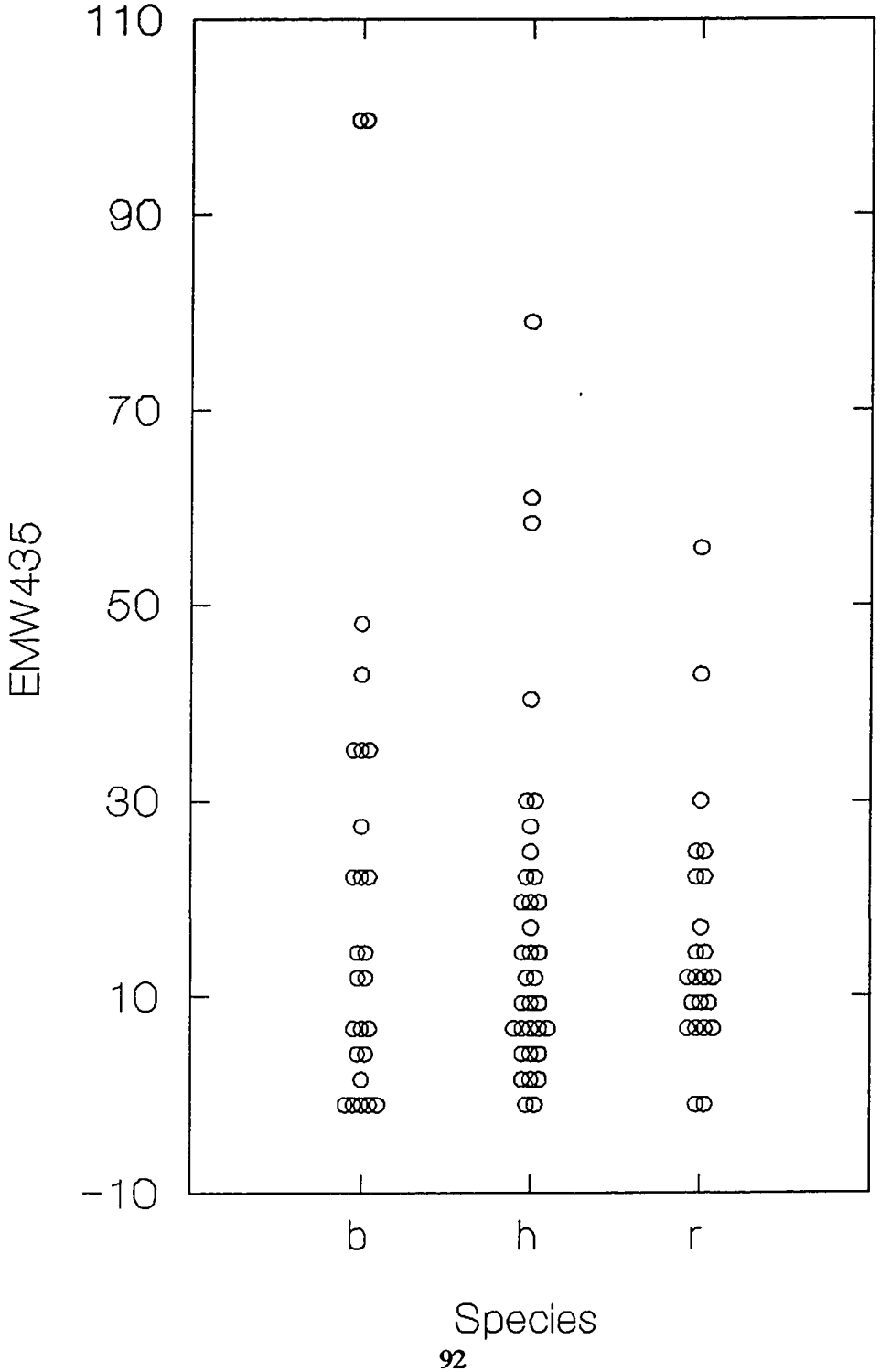


Figure 12. **Within class distribution of the nuclear rDNA allelic marker EMW 4.35 on Isle au Haut: b=black, h=hybrid and r=red.**



Discussion

The molecular classification index applied to samples collected from Isle au Haut, Maine, is based on characters identified using provenance plantation samples of red spruce, black spruce and controlled-cross red-on-black hybrids. Molecular markers used to identify hybridization events have not been shown to display any relationships with geographic origin or environmental type (Chapters 2, 3 and 4). The molecular classification techniques provide an accurate basis for species identification.

The frequency of hybridization and introgression among the samples collected on Isle au Haut was greater than expected (43%) as compared to my other studies. Only 3% of the individuals on Mount Lafayette and 9% of the individuals on Mount Washington were identified as hybrids (Chapter 4). In the black spruce provenance test population 6% of the individuals displayed RFLP patterns suggestive of past hybridization event(s), and 14% of the red spruce provenance test population displayed RFLPs suggestive of past hybridization (Chapter 3).

The black spruce sampled from Isle au Haut tended to be localized in the bogs whereas the red spruce was observed upland and inland of the black spruce. Habitats favorable to hybridization and introgression exist between these two sites. Habitats similar to those observed in the Isle Au Haut were previously reported to favor hybridization and introgression (Osawa 1986; Thorpe 1986; Osawa, Spies, Dimond 1986). Thorpe (1986) reported that the composition of lowland spruce-fir stands in Maine shifted from black spruce to hybrids to red spruce as soil fertility and drainage improved. Spruce budworm sensitivity followed this same trend (Osawa 1986; Osawa, Spies, Dimond 1986) with black spruce in the boggy sites being the most resistant to infestation and the red spruce located

upland being the most sensitive to damage. In these studies trees were identified based on sensitivity to spruce budworm outbreaks and morphological traits; there does exist some question as to whether they were observing within-species variation and/or environmental effects as oppose to true between-species variation.

The degree of hybridization/introgression on Isle au Haut varied in the different test sites from 21% to 67% (Table 13). The mitochondria was not polymorphic in the samples obtained from Isle au Haut, which may result in misclassification of hybrids as red spruce in these populations (Chapters 3 and 4). Therefore the degree of hybridization may be greater than reported.

In zone 1 (the bog site), four trees classified as red spruce using the three character molecular index (Table 13). These spruce were characterized, using morphological indices, as red spruce (Eckert personal communication) or hybrid (Major personal communication). The fact that these individuals were growing in boggy soil may have influenced the morphology of these sample making them appear introgressed; alternatively, they may be introgressed with a strong red spruce nuclear influence.

Variation in two of the nuclear rDNA allelic frequencies were dissimilar compared to observations made for provenance plantation samples (Chapter 2). The frequency of EMW 4.35 was lower then that observed among provenance plantation samples (Figure 12) suggesting ecogeographical selection on this allele, as has been observed in provenance plantation samples (Chapter 2) and on Mount Washington (Chapter 4). The tendency for this allele to display significant variation over geographic origins and habitat type suggests selection rather than genetic drift. The frequency of HMW 1.6 appears to be high for red spruce and low for black spruce in comparison with observations made

from provenance plantation samples (Figure 11). These results indicate that extensive introgression and gene flow has occurred between Picea rubens and P. mariana populations on Isle au Haut.

On Isle Au Haut, hybridization was most frequently observed in the regions between black spruce located within the bog and upland red spruce. From these findings I conclude that this habitat enhances hybridization and introgression which provides a route for substantial gene flow between red and black spruce.

Summary

Restriction fragment length polymorphism (RFLP) variation was examined at the specific and interspecific level among provenance plantation samples from the entire range of red spruce (Picea rubens Sarg.), the eastern complex of black spruce (Picea mariana (Mill.) B.S.P.), control-cross red-on-black hybrids, and natural populations of red and black spruce. RFLP variation was used to examine within-species and population variation and to develop an accurate species index capable of identifying red spruce, black spruce and hybridization between the two species. The repetitive nature of the rDNA repeat unit made it attractive for use in examining hybridization and introgression. The inheritance patterns of organelle haplotypes made them desirable for examining the direction of hybridization.

The first chapter presents a characterization of the nuclear ribosomal repeats for the two closely related Picea species. RFLPs were identified in the IGS and ITS sequences (Figure 1); however, no polymorphism was species specific. The rDNA repeat units were found to be polymorphic within an individual genome with up to five distinct rDNA repeat units types (alleles) evident in some individuals. The nuclear rDNA repeat unit size ranged from a minimum of 32 kbp to greater than 40 kbp, two to three fold larger than the typical angiosperm rDNA unit.

The copy concentration of the rDNA repeat unit (in copies/pg genomic DNA) was also examined for within-species and between-species variation. Extensive within-species variation in rDNA copy concentration was observed. Three fold variation was observed within P. rubens individuals and as much as six fold variation among P. mariana individuals (Figure 3); however, between-species variation was not significant (F=3.35, p=0.069)

making copy concentration unusable as a species identifier. Regression analysis revealed a significant relationship between copy concentration of the rDNA repeat unit in P. rubens and geographic origins (Table 1). At a size greater than 32 kbp and a concentration averaging 1.2-1.3 X 10^4 copies per pg, the rDNA repeat constitutes approximately 4% of the total genome. Differences in the rDNA content in Picea could contribute to the variation in overall genome size that has been observed within conifer species.

The RFLPs observed in the rDNA repeat were not species specific; however, trends in internal allelic frequencies were noticed which were useful for between-species differentiation (Table 3, Table 6). In chapter two the rDNA allelic frequencies were further investigated for within and between-species variation using statistical techniques. The frequencies of polymorphic restriction fragments for the nuclear rDNA repeat were compared for 12 provenances of Picea rubens and 34 provenances of Picea mariana.

Significant correlations were observed between individual allelic markers and between individual markers and copy concentration (Table 2). Correlations between alleles were not strong suggesting that the rDNA alleles are varying independently of one another. Correlations between copy concentration and allelic frequencies were weak implying that the mechanism(s) responsible for altering the rDNA copy concentration preserve the ratio of rDNA alleles.

Discriminant analysis, using the restriction fragment frequencies for the rDNAs, was used to develop a classification model for the two species. Ten fold verification of the model produced an average correct classification of 99% for black spruce and 96% for red spruce. Canonical correlation analysis revealed significant variation of restriction fragment frequencies with a geographic variate comprised of latitude and longitude of provenances

(Table 4).

Variation in the nuclear rDNA repeat could not accurately differentiate hybrids from black spruce. In addition, one marker (EMW 4.35) displayed a strong relationship with latitude (Table 4) suggesting that the observed between-species variation for this marker may be due to selection rather than a true species difference. It was therefore necessary to identify additional markers useful for species classification and to eliminate markers which displayed a significant relationship with geographic origins.

In the third chapter RFLPs were identified for the organelle genomes of red spruce and black spruce. Organelle haplotypes were closely associated with species but were not species specific. The red spruce associated mitochondrial haplotype was observed in black spruce provenances outside of the red spruce range (Figure 6). This may represent within-species variation in black spruce or may have resulted from hybridization during glacial retreats. Northeastern spruce populations were centered in the southern Appalachian Mountains prior to the last glacial retreat (Davis et al. 1980) which may have allowed extensive hybridization and gene flow between red and black spruce.

Data collected from controlled-cross hybrids suggest that the mitochondria is maternally inherited whereas the chloroplast is paternally inherited in these spruce species. Organelle markers were combined with allelic data from the nuclear rDNA repeat to derive a simple three character index capable of identifying red spruce, black spruce and hybridization between the two species (Figure 8).

In the fourth and fifth chapters natural populations of red and black spruce were examined for hybridization/introgression and for variation related to elevation and habitat.

Spruce samples were collected from an elevational transect on Mount Lafayette

(three elevations) and Mount Washington (eight elevations) to examine within and between population variation. The frequencies of nuclear rDNA alleles and mitochondria and chloroplast haplotypes were scored for all individuals. Trees were classified into species using a molecular index and two discriminant classification functions, one based on morphology and another based on molecular markers. Among the red spruce on Mount Washington a correlation was detected between one of the rDNA alleles (EMW 4.35 and its squared value) and elevation squared (R^2=0.13 p=0.009) suggesting that variation in this marker is due to selection rather than between-species differentiation. Using canonical discriminant analysis, significant variation (Wilk's Lambda=0.622, p=0.0001) was detected between red spruce populations on these two mountains. Hybridization and introgression between red and black spruce did not appear to be a major force in the observed patterns of variation on these two mountains.

Extensive hybridization (50%) was detected on Mount Washington at an elevation of 1220 m where red and black spruce overlapped. I did not, however, observe the strong relationship (R^2=0.828) between elevation and introgression previously reported by Berlyn et al. (1990). The observed variation in nuclear genome content (Berlyn et al. 1990) may have been due to elevational effects rather than true between-species variation, similar to that observed in the rDNA allelic marker EMW 4.35. The study on nuclear genome content (Berlyn et al. 1990) lacked the appropriate controls to differentiate between elevational effects and species differences. One control, missing in Berlyn et al. (1991), was examination of red and black spruce nuclear genome content in trees from the same elevation. All of the black spruce surveyed by Berlyn et al. (1991) were high elevation (>1220 m). Alternatively, a more complete study could be preformed on provenance test populations of both red and black spruce.

Samples collected from Isle au Haut, ME were examined for hybridization and introgression occurring between black spruce located in a bog and red spruce located upland and inland of the bog. The Isle au Haut samples displayed a high degree of hybridization and introgression (42% of the trees sampled were identified as hybrid/introgressed) as compared to the provenance plantation samples (9%) and mountain samples (6%). The degree of hybridization was also related to habitat type (Table 12). Examination of additional nuclear markers suggest that significant gene flow has occurred between red and black spruce populations located at Head Harbor, Isle au Haut, ME. These findings suggest that hybridization and introgression between red and black spruce is influenced by not only proximity of the two species but also by habitat type.

References

Allard, R. W., M. A. Saghai Maroof, Qifa Zhang and R. A. Jorgensen 1990. Genetic and molecular organization of ribosomal DNA (rDNA) variants in wild and cultivated barley. Genetics 126: p 743-751.

Appels, R., and R. L. Honeycutt 1986. rDNA: Evolution Over a Billion Years. In, S.K. Dutta (Ed), DNA Systematics Vol. II: Plants. CRC Press, Inc. Boca Raton, Florida. p 81-135.

Bennett, M.D. 1976. DNA amount, latitude, and crop plant distribution, Environ. Exp. Bot. 16: p 93-108.

Berlyn, G.P., J. L. Royte and A. O. Anoruo 1990. Cytophotometric differentiation of high elevation spruces: physiological and ecological implications. Stain Tech. 65: p 1-14.

Chao, S., R. Sederoff, and C. S. Levings III 1984. Nucleotide sequence and evolution of the 18S ribosomal RNA gene in maize mitochondria. Nucleic Acids Research 12: p 6629-6644.

Critchfield, W.B. 1984. Impact of the Pleistocene on the genetic structure of North American conifers. In: Proc. of the 8th North American Forest Biology Workshop, Logan Utah, July 30-Aug 1 1984. p 70-118.

Cullis, C. A., G. P. Creissen, S. W. Gorman, and R. D. Tiasdale 1988. The 25S, 18S, and 5S ribosomal RNA genes from Pinus radiata D.Don. p. 34-40 in W.M. Cheliak and A.A. Yapa, eds. Molecular genetics of forest trees. Can For. Serv. Petawawa Natl. For. Inst. Inf. Rep PI-X-80.

Dallimore, W. and A. B. Jackson 1967. A handbook of Coniferae and Ginkgoaceae. Ed. 4, rev by S.G. Harrison, St. Martin's Press, New York. p. 729.

Davis, M.B., R.W. Spear, and L.C.K. Shane 1980. Holocene climate of New England. Quaternary Research 14 (2): p 240-250.

Dhir, N.K. and J.P. Miksche 1974. Intraspecific variation of nuclear DNA content in Pinus resinosa Slt. Can. J. Genet. Cytol. 16: 77-83.

Dowe, M. F., G. W. Roman, and A. S. Klein 1990. Excision and transposition of two Ds transposons from the bronze mutable 4 Derivative 6856 allele of Zea mays L. Mol. Gen. Genet. 221: p 475-485.

Dower, W. J., J. F. Miller and C. W. Ragsdale 1988. High efficiency transformation of E. coli by high voltage electroporation. Nucleic Acids Res. 16 (3): p6127-6145.

Doyle J. J., and R.N. Beachy 1985. Ribosomal gene variation in soybean (Glycine) and its relatives. Theor. Appl. Genet. 70: p 369-376

Eckert, R. T. 1990. Genetic variation in red spruce pollution response. Final Report. Spruce-Fir Research Cooperative, Northeast Forest Experiment Station, 5 Radnor Corporate Center, 100 Matsonford Rd., Radnor, PA 19087.

Eckert, R. T. 1989. Genetic variation in red spruce and its relation to forest decline in the Northeastern United States. In, J.B. Bucher and I. Bucher-Wallin (eds.), Air Pollution and Forest Decline. Proc. 14th Int. Meeting for Specialists in Air Pollution Effects on Forest Ecosystems IUFRO P2.05, Interlaken, Switzerland, Oct 2-8, 1988. Birmensodorf. 1989, p. 319-324.

El-Lakany, M. H. and O. Sziklai 1971. Intraspecific variation in nuclear characteristics of Douglas-fir. Adv. Front. Plant Sci. 28: p 363-378.

Feinberg, A.P., and B. Vogelstein 1984. "Addendum: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity." Anal. Biochem. 137: p 266-267.

Fisher, R. A. 1936. The use of multiple measurements in taxonomic problems. Annals of Eugenics 7: p 179-188.

Friedland, A. J., R. A. Gregory, L. Karenlampi and A. H. Johnson 1984. Winter damage to foliage as a factor in red spruce decline. Can. J. For. Res. 16: p 963-965.

Fowells, H. A. 1965. Silvics of forest trees of the United States. U.S. Forest Service Ag. Handbook No. 271: p 762.

Fowler, D. P., H. E. MacGillivray, S. A. M. Manley, and J.M. Bonga 1970. Tree breeding in the Maritimes region. 1968-1969. Proceedings of the Twelfth Meeting of the Committee of Forest Tree Breeding in Canada. Univ. Laval, Quebec. p 33-38.

Gittins, R. 1979. Ecological applications of canonical analysis. In, L. Orloci, C.R. Rao and W.M. Stiteler (eds.), Multivariate Methods in Ecological Work. Int. Co-operative Pub House, Fairland, Md. P 309-535.

Gordon, A. G. 1976 The taxonomy and genetics of P. rubens and its relationship to P. mariana. Can J. Bot. 54: p 781-813.

Greenwood, M. S., C. A. Hopper, and K. W. Hutchison, 1989. Maturation in Larch 1. Effect of age on shoot growth, foliar characteristics, and DNA methylation. Plant Physiol. 90: p 406-412.

Grime, J.P. and M.A. Mowforth 1982. Variation in genome size-an ecological interpretation. Nature 299: 151-153.

Johnson, A. H., E. R. Cook, and T. G. Siccama 1988. Climate and red spruce growth and decline in the northern Appalachians. Proc. Nat. Acad. Sci. USA. 85: p 5369-5373.

Jorgensen, R. A. and P. D. Cluster 1988. Modes and tempos in the evolution of nuclear ribosomal DNA: new characters for evolutionary studies and new markers for genetic and

population studies. Ann. MO Bot. Gard. 75: 1238-1247.

Khandjian E. W. 1986. UV crosslinking of RNA to nylon membrane enhances hybridization signals. Mol. Biol. Rep. 11: p 107-117.

King, L. M. and B. A. Schaal 1989. Ribosomal-DNA variation and distribution in Rudbeckia missouriensis. Evolution 43: p 1117-1119.

Levin, D. A., and S. W. Funderburg 1979. Genome size in angiosperms: Temperate versus tropical species. Am. Nat. 114: p 784-798.

Manley, S.A.M. 1971. Identification of red, black and introgressed spruce. Can. Dep. Environ., Can. For. Serv. Publ 1301.

-----1972. The occurrence of hybrid swarms of red and black spruce in central New Brunswick. Can. J. For. Res. 2: p. 381-391.

Manley, S.A.M. and Ledig, F.T. 1979 Photosynthesis in black and red spruce and their hybrid derivatives: Ecological isolation and hybrid adaptive inferiority. Can. J. Bot. 57: p 305-314.

Mergen, F. and B. A. Thielges 1967. Intraspecific variation in nuclear volume in four conifers. Evolution 21: 720-724.

Miksche, J. P. 1967. Variation in DNA content of several gymnosperms. Can. J. Genet. Cytol. 9: 717-722.

Miksche, J. P. 1968. Quantitative study of intraspecific variation of DNA per cell in Picea glauca and Pinus banksiana. Can. J. Genet. Cytol. 10: 590-600.

Miksche, J. P. 1971. Intraspecific variation of DNA per cell between Picea sichensis (Bong.) Carr. provenances. Chromosoma 32: 343-352.

Morgenstern, E.K. and J. L. Farrar 1964. Introgressive hybridization in red spruce and black spruce. Technical Report, Faculty of Forestry. University of Toronto. No. 4.

Morgenstern, E.K. 1969. Winter drying of red spruce provenances related to introgressive hybridization with black spruce. Bi-Mon. Res. Notes, Can. For. Serv. 25 (4): p 34-36.

Morgenstern E. K., A. G. Corriveau and D. P. Fowler 1981. A provenance test of red spruce in nine environments in eastern Canada. Can. J. For. Res. 11: p. 124-131.

Neale, D. B., and R. R. Sederoff 1988. Inheritance and evolution of conifer organelle genomes. Genetic Manipulation of Woody Plants. J.W. Hanover Publisher. Plenum Press, New York p 251-264.

Neale, D. B., K. A. Marshall and R. R. Sederoff 1989. Chloroplast and mitochondrial DNA are paternally inherited in Sequoia sempervirens D. Don Endl. Proc Natl Acad Sci USA 86: p 9347-9349.

Osawa, A. 1986. Patch dynamics of spruce-fir forests during a spruce budworm outbreak in Maine. Doctoral dissertation. Yale University Library, New Haven CT.

Osawa, A., C. J. Spies III, J. B. Dimond 1986. Patterns of tree mortality during an uncontrolled spruce budworm outbreak in Baxter State Park, 1983. Maine Agricultural Experiment Station. Technical Bulletin 121.

Palmer, J. D., C. R. Shields, D. B. Cohen and T. J. Orton 1983. Chloroplast DNA evolutions and the origin of amphidiploid Brassica species. Theor. Appl. Genet. 65: 181-189.

Rao, C.R. 1973. Linear Statistical Inference. New York: John Wiley & Sons.

Roche. L. 1969. Variation in growth behavior of fifteen red spruce (Picea rubens Sarg.) provenances at three sites in Quebec. Can. Dep. Fish. For.,Can For. Serv. Inf. Rep. Q-X-

Saghai Maroof, M. A., K. M. Soliman, R. A. Jorgensen and R. W. Allard 1984. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. USA 81: p 8014-8018.

Saghai Maroof, M. A., R. W. Allard and Qifa Zhang 1990. Genetic diversity and ecogeographical differentiation among ribosomal DNA alleles in wild and cultivated barley. Proc. Natl. Acad. Sci. USA 87: p. 8486-8490.

Sambrook, J., E. F. Fritch and T. Maniatis 1989. Molecular Cloning. A Laboratory Manual 2nd ed. Cold Spring Harbor Laboratory Press.

SAS 1982. SAS User's Guide:Statistics. SAS Inst., Cary, North Carolina.

Schier, G.A. 1985. Response of red spruce and balsam fir seedlings to aluminum toxicity in nutrient solutions. Can. J. For. Res. 15: p. 29-33.

Scott, J. T., T. G. Siccama, A. H. Johnson and A. R. Breisch 1984. Decline of red spruce in the Adirondacks, New York. Bull Torrey Bot Club 111: p 438-444.

Smith, D.E. and A. S. Klein (1992). A ribosomal DNA-based phylogeny of North American and European Picea species. In preparation.

Stern, K. and L. Roche 1974. Genetics of forest ecosystems. Springer-Verlag. New Yor, Heidelberg, Berlin.

Stine M. and D. E. Keathley 1990. Paternal inheritance of plastids in Engelmann spruce X blue spruce hybrids. Heredity 81: p 443-446.

Strauss S.H. and C.H. Tsai 1988. Ribosomal gene number variability in Douglas-Fir. Heredity 79: 453-458.

Sytsma K. J., and B. A. Schaal 1985. Phylogenetics of the Lisianthius skinneri (Gentianaceae) species complex in Panama utilizing DNA restriction fragment analysis. Evolution 39: p 594-608.

Thorpe, J. P. 1986. An analysis of stand composition in relation to site in an area of spruce-fir-northern hardwoods in Maine. Yale University Library, New Haven CT.

Vogelmann, H.W., G. J. Badger, M. Bliss and R. M. Klein 1985. Forest decline on Camels Hump, Vermont. Bull Torrey Bot Club 112: p 274-287.

Wagner, D.B., G. R. Furnier, M. A. Saghai Maroof, B. P. Dancik, and R. W. Allard 1987. Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. Proc. Natl. Acad. Sci. USA 84: p 2097-2100.

Wilkinson, R.C. 1990. Effects of winter injury on basal area and height growth of 30-year-old red spruce from 12 provenances growing in northern New Hampshire. Can. J. For. Res. 20: p 1616-1622.

Wright, J. W., 1955: Species crossability in spruce in relation to distribution and taxonomy. Fors. Sci. 1(4): 319-349.

Zimmer, E.A., E. R. Jupe and V. Walbot 1988. Ribosomal gene structure, variation and inheritance in maize and its ancestors. Genetics 120: p 1125-1136.

Appendixes

Appendix 1: Organelle haplotypes and rDNA allelic frequencies for A) red spruce provenance plantation samples collected at Coleman State Park Stewartstown, NH (C), black spruce provenance plantation located at Massabesic Experiment forest, Alfred ME (M), red on black controlled-cross hybrids (H) and B) samples collected from Isle au Haut, ME (1-4), spruce from Mts. Washington, NH (W) and Lafayette NH (L). Trees were classified into species using the three character molecular index described in chapter 3 (SP). For elevational surveys elevation was substituted for stand number. Missing data are indicated by a dashed line (-). Restriction fragment frequencies are given in percentages. Elevation is given in meters.

A)

ID	SP	STAND	TREE #	HMW 3.4	HMW 1.6	SMW 1.7	EMW 4.35	SITS	MITO	CHLORO	LAT	LON	ELE
C	r	2019	16.110	0	55	64	72	95	1	0	35.36	83.27	1740
C	r	2019	21.100	0	0	61	8	87	1	0	35.36	83.27	1740
C	r	2019	25.200	0	0	44	10	83	1	0	35.36	83.27	1740
C	r	2019	38.120	0	22	61	16	81	1	0	35.36	83.27	1740
C	r	2019	56.110	0	20	63	12	81	1	0	35.36	83.27	1740
C	r	2020	16.100	0	10	36	20	79	1	0	38.38	79.50	1250
C	r	2020	16.200	0	8	57	33	84	1	0	38.38	79.50	1250
C	r	2020	17.500	0	6	60	27	83	1	0	38.38	79.50	1250
C	r	2020	18.400	0	17	47	29	67	1	0	38.38	79.50	1250
C	r	2020	20.200	0	17	50	31	88	1	0	38.38	79.50	1250
C	h	2021	6.101	0	20	53	53	55	0	0	40.45	77.45	550
C	h	2021	6.600	0	31	39	61	82	0	0	40.45	77.45	550
C	h	2021	7.900	0	28	55	55	74	0	0	40.45	77.45	550
C	h	2021	44.300	0	7	55	35	87	0	1	40.45	77.45	550
C	r	2022	36.700	0	19	47	8	59	1	0	42.22	73.15	550
C	r	2022	37.600	0	27	48	13	88	1	0	42.22	73.15	550

ID	SP	STAND	TREE #	HMW 3.4	HMW 1.6	SMW 1.7	EMW 4.35	SITS	MITO	CHLORO	LAT	LON	ELE
C	r	2022	51.300	0	8	51	8	63	1	0	42.22	73.15	550
C	r	2022	52.200	0	18	47	26	89	1	0	42.22	73.15	550
C	r	2022	54.400	0	22	54	18	77	1	0	42.22	73.15	550
C	r	2024	15.100	0	0	59	38	62	1	0	44.25	73.40	610
C	r	2024	16.800	0	14	54	13	64	1	0	44.25	73.40	610
C	h	2024	31.200	0	10	51	14	74	0	0	44.25	73.40	610
C	r	2024	33.100	0	47	43	67	83	1	0	44.25	73.40	610
C	r	2024	34.300	0	0	51	28	62	1	0	44.25	73.40	610
C	r	2027	26.000	0	18	58	14	72	1	0	43.12	72.05	470
C	h	2027	27.300	0	0	-	-	73	-	-	43.12	72.05	470
C	r	2027	29.400	0	13	50	34	58	1	0	43.12	72.05	470
C	r	2027	30.200	0	9	50	11	70	1	0	43.12	72.05	470
C	r	2027	36.500	0	17	63	60	83	1	0	43.12	72.05	470
C	r	2030	56.100	0	13	58	66	82	1	0	44.54	68.23	50
C	r	2030	56.200	0	21	58	42	77	1	0	44.54	68.23	50
C	r	2030	57.100	0	25	59	47	89	1	0	44.54	68.23	50
C	r	2030	59.200	0	10	57	14	77	1	0	44.54	68.23	50
C	r	2030	60.300	0	35	53	33	94	1	0	44.54	68.23	50
C	r	2032	6.100	0	35	41	33	74	1	0	46.55	71.33	275
C	r	2032	6.500	0	20	57	43	77	1	0	46.55	71.33	275
C	r	2032	8.100	0	24	45	25	73	1	0	46.55	71.33	275
C	r	2032	46.300	0	34	53	37	81	1	0	46.55	71.33	275
C	r	2032	49.130	0	0	53	25	68	1	0	46.55	71.33	275

ID	SP	STAND	TREE #	HMW 3.4	HMW 1.6	SMW 1.7	EMW 4.35	SITS	MITO	CHLORO	LAT	LON	ELE
C	r	2100	13.500	0	51	34	37	88	1	0	45.12	62.44	75
C	r	2100	14.400	0	28	58	12	75	1	0	45.12	62.44	75
C	r	2100	15.300	0	0	60	0	71	1	0	45.12	62.44	75
C	r	2100	50.600	0	19	57	16	73	1	0	45.12	62.44	75
C	r	2100	50.800	0	0	55	10	68	1	0	45.12	62.44	75
C	r	2101	46.100	0	13	61	43	72	1	0	44.10	65.54	50
C	h	2101	47.100	0	16	58	54	85	0	0	44.10	65.54	50
C	r	2101	47.200	0	0	45	18	65	1	0	44.10	65.54	50
C	r	2101	47.500	0	0	67	35	46	1	0	44.10	65.54	50
C	h	2103	2.100	0	19	61	73	87	1	1	46.00	66.20	70
C	r	2103	3.300	0	44	41	30	88	1	0	46.00	66.20	70
C	r	2103	4.500	0	0	52	57	66	1	0	46.00	66.20	70
C	r	2103	37.140	0	20	49	43	85	1	0	46.00	66.20	70
C	h	2103	40.200	0	12	-	43	79	1	1	46.00	66.20	70
C	r	2505	36.300	0	30	49	67	90	1	0	46.37	66.40	70
C	r	2505	38.500	0	35	47	32	80	1	0	46.37	66.40	70
C	r	2505	40.100	0	5	-	43	76	1	0	46.37	66.40	70
C	r	2505	57.6	0	30	50	34	87	1	0	46.37	66.40	70
C	r	2505	58.6	0	18	47	37	77	1	0	46.37	66.4	70
M	b	3267	2	0	42	74	100	100	1	1	45.44	89.03	
M	b	3267	3	24	42	50	97	100	1	1	45.44	89.03	
M	b	3293	2	0	60	73	86	100	0	1	46.09	90.47	
M	b	3293	3	0	32	-	95	100	0	1	46.09	90.47	

ID	SP	STAND	TREE #	HMW 3.4	HMW 1.6	SMW 1.7	EMW 4.35	SITS	MITO	CHLORO	LAT	LON
M	b	3293	4	0	55	70	84	100	0	1	46.09	90.47
M	b	4274	1	16	41	61	73	100	0	1	44.15	72.16
M	b	4274	4	0	62	47	100	100	1	1	44.15	72.16
M	b	4346	1	0	56	42	90	100	0	1	44.38	84.20
M	b	4346	2	0	42	-	58	100	0	1	44.38	84.20
M	b	4346	3	0	41	57	66	100	0	1	44.38	84.20
M	b	4348	2	0	50	55	88	100	1	1	44.12	85.35
M	b	4348	3	0	18	58	89	100	1	1	44.12	85.35
M	h	4348	4	0	13	63	66	100	0	0	44.12	85.35
M	b	4351	1	0	48	57	90	100	0	1	46.03	84.47
M	b	4351	2	0	61	68	19	100	0	1	46.03	84.47
M	h	4355	1	0	52	63	46	100	0	0	47.42	92.28
M	b	4355	2	0	59	59	56	100	0	1	47.42	92.28
M	b	4355	3	0	45	66	57	100	0	1	47.42	92.28
M	b	4355	4	0	40	56	86	100	0	1	47.42	92.28
M	b	4355	4	0	40	56	86	100	0	1	47.42	92.28
M	b	4356	1	0	58	54	63	100	0	1	47.42	92.28
M	b	4356	2	0	52	56	42	100	0	1	47.42	92.28
M	b	4356	3	0	41	64	69	100	1	1	47.42	92.28
M	b	4356	4	0	52	57	72	100	0	1	47.42	92.28
M	b	4359	1	0	39	54	81	100	0	1	46.32	69.55
M	b	4359	2	0	51	55	78	100	0	1	46.32	69.55
M	b	4359	3	0	46	50	85	100	0	1	46.32	69.55

ID	SP	STAND	TREE #	HMW 3.4	HMW 1.6	SMW 1.7	EMW 4.35	SITS	MITO	CHLORO	LAT	LON
M	b	4359	4	0	48	65	82	100	0	1	46.32	69.55
M	b	4899	2	0	46	64	59	100	0	1	45.10	77.10
M	b	4899	3	65	21	-	36	100	0	1	45.10	77.10
M	b	4899	4	0	44	50	36	100	0	1	45.10	77.10
M	b	4905	1	6	27	64	25	100	1	1	46.20	82.50
M	b	4905	2	4	60	67	20	100	1	1	46.20	82.50
M	b	4905	3	14	60	-	58	100	1	1	46.20	82.50
M	b	4915	1	0	63	54	72	100	0	1	48.12	82.23
M	b	4915	2	15	61	57	50	100	0	1	48.12	82.23
M	b	4915	3	0	83	59	91	100	0	1	48.12	82.23
M	b	4915	4	0	69	49	81	100	0	1	48.12	82.23
M	b	4916	1	0	55	64	30	100	0	1	48.40	90.10
M	b	4916	2	0	63	65	86	100	0	1	48.40	90.10
M	b	4918	1	0	68	38	84	100	0	1	48.48	93.40
M	b	4918	2	0	100	50	87	100	0	1	48.48	93.40
M	b	4918	3	0	68	64	77	100	0	1	48.48	93.40
M	b	4921	1	26	38	44	94	100	1	1	49.20	93.55
M	b	4921	2	26	38	64	5	100	1	1	49.20	93.55
M	b	4921	3	0	43	67	4	100	1	1	49.20	93.55
M	b	4921	4	0	43	-	6	100	1	1	49.20	93.55
M	b	4942	3	0	57	54	92	100	1	1	44.50	78.05
M	b	4942	4	0	48	67	80	100	1	1	44.50	78.05
M	b	4946	1	0	34	69	90	100	0	1	45.58	77.25

ID	SP	STAND	TREE #	HMW 3.4	HMW 1.6	SMW 1.7	EMW 4.35	SITS	MITO	CHLORO	LAT	LON
M	b	4946	3	0	19	49	90	100	0	1	45.58	77.25
M	b	4946	4	0	72	61	89	100	0	1	45.58	77.25
M	b	4959	1	0	61	54	100	100	0	1	46.43	84.23
M	h	4959	2	0	75	-	0	73	0	1	46.43	84.23
M	b	4959	3	0	65	54	88	100	0	1	46.43	84.23
M	b	4959	4	0	67	56	69	100	1	1	46.43	84.23
M	b	4962	1	41	33	55	16	100	1	1	49.01	55.26
M	b	4962	2	17	63	-	22	100	1	1	49.01	55.26
M	r	4962	3	0	5	68	64	87	1	0	49.01	55.26
M	b	4962	4	6	31	63	45	100	1	1	49.01	55.26
M	b	4981	1	12	68	59	63	100	1	1	48.22	76.57
M	b	4981	2	15	44	44	75	100	0	1	48.22	76.57
M	b	4981	4	0	30	42	98	100	1	1	48.22	76.57
M	b	4985	1	0	58	55	89	100	0	1	49.36	71.18
M	b	4985	2	0	78	56	62	100	0	1	49.36	71.18
M	b	4985	3	0	56	62	0	100	0	1	49.36	71.18
M	b	4985	4	0	58	-	42	100	0	1	49.36	71.18
M	h	4986	1	6	35	57	41	95	0	0	46.16	72.02
M	b	4986	2	0	46	61	61	100	0	1	46.16	72.02
M	b	4986	3	0	58	64	42	100	0	1	46.16	72.02
M	b	4988	1	0	69	62	55	100	0	1	46.52	71.39
M	h	4993	2	0	45	60	55	94	0	1	46.49	65.09
M	b	4993	3	0	55	59	85	100	0	1	46.49	65.09

ID	SP	STAND	TREE #	HMW 3.4	HMW 1.6	SMW 1.7	EMW 4.35	SITS	MITO	CHLORO	LAT	LON
M	b	4995	1	0	30	56	89	100	0	1	45.58	66.20
M	b	4995	2	0	50	55	25	100	0	1	45.58	66.20
M	b	4995	3	0	75	59	66	100	0	1	45.58	66.20
M	b	4995	4	0	46	55	53	100	0	1	45.58	66.20
M	b	4997	1	0	46	46	76	100	1	1	45.42	66.31
M	b	4997	2	0	67	54	85	100	0	1	45.42	66.31
M	b	4997	3	0	58	40	79	100	0	1	45.42	66.31
M	b	4997	4	0	69	67	84	100	0	1	45.42	66.31
M	b	4999	1	0	62	47	50	100	0	1	45.20	64.28
M	b	4999	2	0	52	46	76	100	0	1	45.20	64.28
M	b	4999	3	0	60	49	75	100	0	1	45.20	64.28
M	b	5002	1	0	62	73	59	100	0	1	45.35	66.29
M	b	5002	2	0	43	60	93	100	0	1	45.35	66.29
M	b	5002	4	0	63	70	72	100	0	1	45.35	66.29
M	b	5003	1	0	44	71	95	100	1	1	47.33	66.24
M	b	5003	2	0	56	66	85	100	0	1	47.33	66.24
M	b	5003	3	0	45	75	22	100	1	1	47.33	66.24
M	b	5003	4	0	60	57	34	100	1	1	47.33	66.24
M	b	5004	1	0	46	45	72	100	1	1	46.53	65.05
M	b	5004	2	0	51	-	78	100	1	1	46.53	65.05
M	h	5004	3	0	44	75	11	100	1	0	46.53	65.05

ID	SP	STAND	TREE #	HMW 3.4	HMW 1.6	SMW 1.7	EMW 4.35	SITS	MITO	CHLORO	LAT	LON
H	h	1	1	0	52	57	65	99	0	0		
H	h	2	4	0	48	68	19	88	0	0		
H	h	3	4	0	60	55	48	99	0	0		
H	h	4	1	0	38	65	54	99	0	0		
H	h	5	1	0	37	71	50	100	0	0		
H	h	6	4	0	42	52	57	99	0	0		
H	h	7	4	0	50	57	43	99	0	0		
H	h	8	1	0	43	66	59	84	0	0		
H	h	9	1	0	36	48	55	89	0	0		
H	h	10	4	0	49	59	38	86	0	0		
H	h	11	4	0	50	54	53	93	0	1		
H	h	12	1	0	48	57	61	89	0	0		
H	h	13	4	0	46	57	45	100	0	0		

B)

ID	SP	STAND	TREE	HMW 1.6	EMW 4.35	SITS	MITO	CHLORO
I	r	1	1	38	7	79	1	0
I	h	1	2	52	13	95	1	1
I	b	1	3	46	14	100	1	1
I	r	1	4	44	42	87	1	0
I	b	1	5	9	23	100	1	1
I	h	1	6	42	19	92	1	1
I	h	1	7	26	6	94	1	1
I	h	1	8	28	0	92	1	1
I	b	1	9	31	21	100	1	1
I	h	1	10	-	-	64	1	1
I	h	1	11	46	14	93	1	1
I	b	1	12	58	2	100	1	1
I	h	1	13	29	23	80	1	1
I	b	1	14	11	0	100	1	1
I	b	1	15	53	5	100	1	1
I	h	1	16	46	15	100	1	0
I	b	1	17	55	0	100	1	1
I	h	1	18	35	7	82	1	1
I	b	1	19	39	-	100	1	1
I	h	1	20	74	14	100	1	0
I	b	1	21	27	0	100	1	1
I	h	1	22	23	40	100	1	0
I	h	1	23	-	26	100	1	0
I	r	1	25	25	8	91	1	0
I	h	1	26	29	19	100	1	0
I	r	1	27	24	12	87	1	0
I	h	2	29	20	8	84	1	1
I	h	2	30	47	-	90	1	1
I	h	2	32	37	11	100	1	0
I	b	2	33	72	-	100	1	1
I	b	2	34	48	12	100	1	1
I	r	2	35	27	9	90	1	0
I	r	2	36	33	9	82	1	0

ID	SP	STAND	TREE	HMW 1.6	EMW 4.35	SITS	MITO	CHLORO
I	b	2	37	46	11	100	1	1
I	b	2	38	69	8	100	1	1
I	b	2	39	48	7	100	1	1
I	b	2	40	19	0	100	1	1
I	b	2	42	51	8	100	1	1
I	h	2	43	43	1	100	1	0
I	h	2	44	45	21	100	1	0
I	b	2	46	29	34	100	1	1
I	b	2	47	46	49	100	1	1
I	b	2	48	31	42	100	1	1
I	r	2	49	53	24	84	1	0
I	b	2	50	18	35	100	1	1
I	r	2	52	0	13	76	1	0
I	b	2	53	39	100	100	1	1
I	b	2	54	73	100	100	1	1
I	b	2	55	54	14	100	1	1
I	b	2	56	25	34	100	1	1
I	b	2	58	41	-	100	1	1
I	b	2	59	33	-	100	1	1
I	h	3	61	-	2	100	1	0
I	h	3	62	55	4	100	1	0
I	h	3	63	-	2	100	1	0
I	r	3	64	21	18	92	1	0
I	r	3	65	76	23	72	1	0
I	h	3	66	33	19	100	1	0
I	h	3	67	-	0	100	1	0
I	h	3	68	30	9	100	1	0
I	r	3	69	20	6	63	1	0
I	h	3	70	46	30	100	1	0
I	h	3	71	26	30	100	1	0
I	r	3	72	-	-	-	1	0
I	h	3	73	66	61	100	1	0
I	h	3	74	44	80	100	1	0
I	r	3	75	-	6	-	1	0

ID	SP	STAND	TREE	HMW 1.6	EMW 4.35	SITS	MITO	CHLORO
I	r	3	76	23	14	71	1	0
I	h	3	77	32	58	100	1	0
I	r	3	78	44	6	90	1	0
I	h	3	79	36	27	100	1	0
I	b	3	80	40	21	100	1	1
I	r	4	81	7	13	62	1	0
I	b	4	82	15	5	100	1	1
I	h	4	83	13	5	89	1	1
I	h	4	84	11	7	85	1	1
I	r	4	85	26	12	79	1	0
I	r	4	86	13	56	62	1	0
I	h	4	87	52	9	100	1	0
I	h	4	89	47	3	89	1	1
I	h	4	90	14	16	100	1	0
I	r	4	91	24	-	82	1	0
I	r	4	92	30	23	66	1	0
I	r	4	93	78	0	85	1	0
I	h	4	94	0	0	-	1	1
I	h	4	95	75	10	100	1	0
I	r	4	97	3	29	59	1	0
I	r	4	99	0	0	45	1	0
I	r	4	100	43	25	76	1	0
I	r	4	101	0	14	55	1	0
L	r	710	1140	18	10	79	1	0
L	r	710	1141	20	2	68	1	0
L	r	710	1142	0	0	82	1	0
L	r	710	1143	0	0	44	1	0
L	r	710	1144	13	16	64	1	0
L	r	710	1145	0	13	56	1	0
L	r	710	1146	0	11	65	1	0
L	r	710	1147	0	17	35	1	0
L	r	710	1148	8	7	70	1	0
L	r	710	1149	7	10	60	1	0
L	h	710	1150	0	12	53	1	1

ID	SP	STAND	TREE	HMW 1.6	EMW 4.35	SITS	MITO	CHLORO
L	r	710	1151	0	-	54	1	0
L	r	710	1152	0	4	79	1	0
L	r	710	1153	8	13	90	1	0
L	r	710	1154	41	14	79	1	0
L	r	710	1155	0	4	67	1	0
L	r	710	1156	10	21	74	1	0
L	r	710	1157	11	14	88	1	0
L	r	710	1159	12	16	74	1	0
L	r	710	1161	0	11	66	1	0
L	r	710	1162	0	9	58	1	0
L	r	710	1163	18	10	75	1	0
L	r	710	1164	0	0	50	1	0
L	r	710	1165	0	5	79	1	0
L	r	710	1167	20	0	62	1	0
L	r	710	1163	39	15	77	1	0
L	r	1190	1065	0	7	74	1	0
L	r	1190	1066	0	15	45	1	0
L	r	1190	1067	7	20	57	1	0
L	r	1190	1068	0	11	76	1	0
L	r	1190	1069	38	14	85	1	0
L	r	1190	1070	0	0	68	1	0
L	r	1190	1071	0	-10	78	1	0
L	r	1190	1072	20	13	83	1	0
L	r	1190	1073	0	13	73	1	0
L	r	1190	1074	11	14	92	1	0
L	r	1190	1075	23	7	81	1	0
L	r	1190	1076	10	15	67	1	0
L	r	1190	1077	18	12	78	1	0
L	r	1190	1078	0	26	69	1	0
L	r	1190	1079	18	15	67	1	0
L	r	1190	1080	0	10	75	1	0
L	r	1190	1081	16	0	54	1	0
L	r	1190	1082	0	20	61	1	0
L	r	1190	1083	0	0	65	1	0

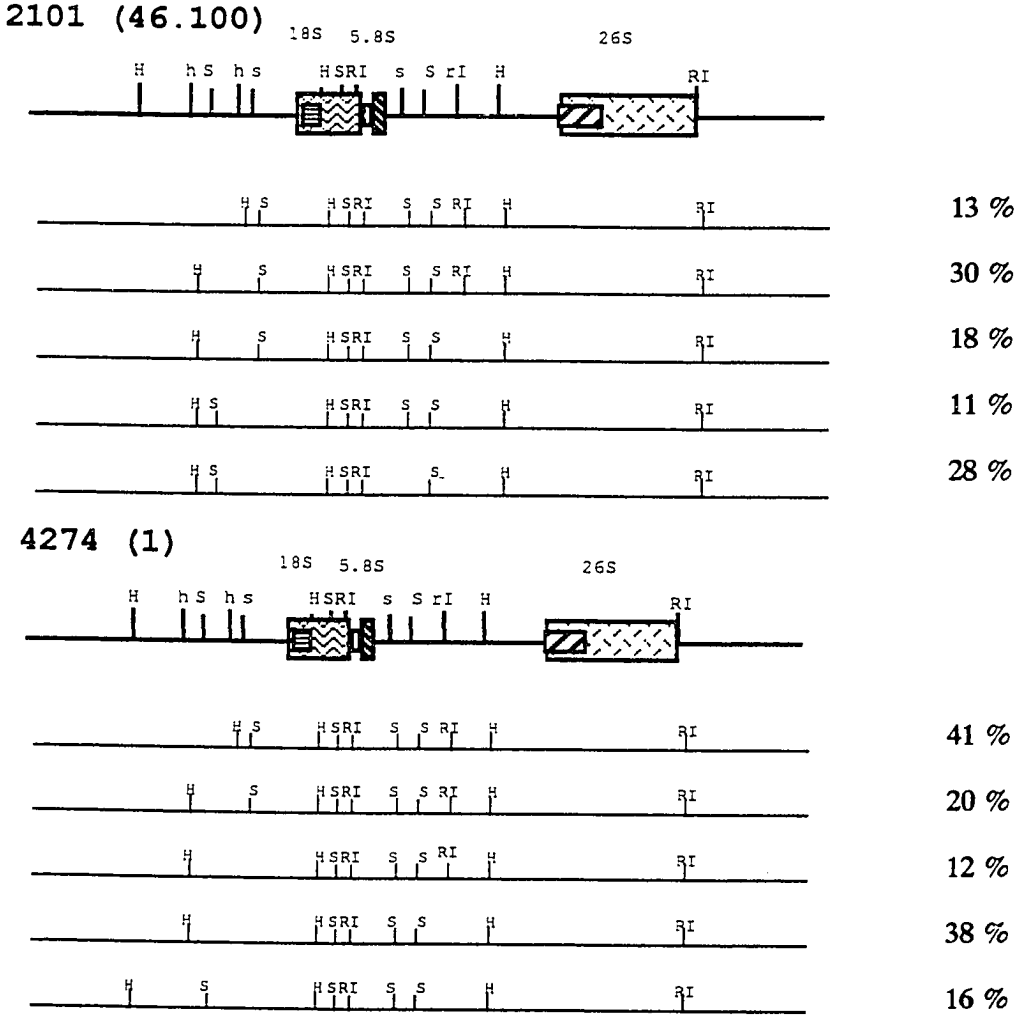
ID	SP	STAND	TREE	HMW 1.6	EMW 4.35	SITS	MITO	CHLORO
L	r	1190	1084	7	10	64	1	0
L	r	1190	1085	18	4	64	1	0
L	r	1190	1086	15	10	68	1	0
L	r	1190	1087	34	12	75	1	0
L	r	1190	1088	34	8	62	1	0
L	r	1190	1089	0	9	78	1	0
L	r	1190	1090	0	10	81	1	0
L	r	1190	1091	32	11	65	1	0
L	r	1190	1092	0	11	91	1	0
L	r	1190	1093	38	10	68	1	0
L	r	1190	1094	14	10	79	1	0
L	r	1190	1095	13	11	83	1	0
L	r	1190	1096	44	-	78	1	0
L	r	1190	1097	14	-	73	1	0
L	r	1190	1098	21	12	79	1	0
L	r	1190	1099	19	7	82	1	0
L	r	1190	1100	23	10	64	1	0
L	r	1190	1101	0	10	86	1	0
L	h	1190	1102	0	-	99	1	0
L	r	1190	1103	0	28	89	1	0
L	r	1430	1104	35	19	90	1	0
L	r	1430	1105	0	38	70	1	0
L	r	1430	1106	0	5	76	1	0
L	r	1430	1107	0	-	91	1	0
L	r	1430	1108	21	6	63	1	0
L	r	1430	1109	0	11	59	1	0
L	r	1430	1110	0	0	69	1	0
L	r	1430	1111	0	11	65	1	0
L	r	1430	1112	23	10	53	1	0
L	r	1430	1113	0	14	91	1	0
L	r	1430	1114	30	8	53	1	0
L	r	1430	1115	0	21	73	1	0
L	r	1430	1116	0	3	54	1	0
L	r	1430	1117	0	15	66	1	0

ID	SP	STAND	TREE	HMW 1.6	EMW 4.35	SITS	MITO	CHLORO
L	r	1430	1119	45	9	82	1	0
L	r	1430	1120	0	14	67	1	0
L	r	1430	1121	0	11	74	1	0
L	r	1430	1122	0	2	75	1	0
L	h	1430	1123	12	8	65	0	0
L	r	1430	1125	3	8	72	1	0
L	r	1430	1126	15	5	72	1	0
L	r	1430	1127	18	2	75	1	0
L	r	1430	1128	14	5	80	1	0
L	r	1430	1129	-	-	73	1	0
L	r	1430	1130	12	12	71	1	0
L	r	1430	1132	0	5	83	1	0
W	r	530	1	30	8	81	1	0
W	r	530	2	34	0	80	1	0
W	r	530	5	16	94	93	1	0
W	r	530	6	0	83	78	1	0
W	r	530	7	7	11	75	1	0
W	r	530	8	18	66	82	1	0
W	r	530	9	14	0	82	1	0
W	r	610	11	19	40	92	1	0
W	r	610	12	12	11	83	1	0
W	r	610	13	7	3	95	1	0
W	r	610	14	1	0	61	1	0
W	r	610	16	20	51	90	1	0
W	r	610	17	0	56	58	1	0
W	r	670	21	8	74	83	1	0
W	r	670	22	12	75	91	1	0
W	r	670	23	20	13	87	1	0
W	r	670	24	0	52	56	1	0
W	r	670	25	0	77	74	1	0
W	r	670	26	9	85	40	1	0
W	r	670	27	5	11	76	1	0
W	r	670	28	8	0	71	1	0
W	r	670	29	18	-	47	1	0

ID	SP	STAND	TREE	HMW 1.6	EMW 4.35	SITS	MITO	CHLORO
W	r	670	30	7	65	72	1	0
W	h	790	31	39	89	100	1	0
W	r	790	42	21	-	58	1	0
W	r	790	43	19	44	86	1	0
W	r	790	44	10	75	79	1	0
W	r	790	45	0	12	94	1	0
W	h	790	46	0	60	98	1	1
W	r	790	47	18	75	92	1	0
W	r	790	50	16	-	66	1	0
W	r	880	51	1	13	78	1	0
W	r	880	53	1	82	65	1	0
W	r	880	54	5	71	82	1	0
W	r	880	55	11	69	77	1	0
W	r	880	57	13	7	75	1	0
W	r	880	58	0	60	76	1	0
W	r	880	59	38	69	66	1	0
W	r	880	61	10	28	81	1	0
W	r	880	64	0	7	76	1	0
W	r	880	65	0	18	52	1	0
W	r	880	60	0	0	64	1	0
W	r	1060	66	24	14	87	1	0
W	r	1060	68	11	3	88	1	0
W	r	1060	69	17	10	78	1	0
W	r	1060	71	1	57	70	1	0
W	r	1060	74	0	-	67	1	0
W	r	1060	75	10	0	41	1	0
W	r	1220	76	-	66	65	1	0
W	h	1220	77	15	41	98	1	0
W	r	1220	78	35	38	93	1	0
W	h	1220	79	30	38	98	1	0
W	h	1220	80	13	37	100	1	0
W	b	1220	81	11	9	100	1	1
W	b	1220	87	-	44	100	1	1
W	h	1220	89	10	30	84	1	1

ID	SP	STAND	TREE	HMW 1.6	EMW 4.35	SITS	MITO	CHLORO
W	b	1435	92	58	5	100	1	1
W	b	1435	93	32	0	100	1	1
W	b	1435	94	50	0	100	1	1
W	h	1435	95	-	-	75	1	1
W	b	1435	96	44	10	100	1	1
W	b	1435	97	56	0	100	1	1
W	b	1435	104	50	0	100	1	1
W	b	1435	105	47	1	100	1	1

Appendix 2. A number of spruce individuals displayed nuclear rDNA RFLP patterns that could only be explained by a minimum of five rDNA alleles. The following displays the minimum number of alleles possible in one red spruce individual and one black spruce individual based on rDNA RFLP frequencies:



Appendix 3. Canonical correlation analysis is a multivariate technique used to examine relationships between two sets of data where each set consist of two or more variables. For my purposes, I have employed canonical correlation analysis to examine relationships between molecular markers and geographic origins of provenances.

Canonical correlation analysis maximizes the relationship between the two data sets while controlling for correlations among the variable within each data set (R_{11} and R_{22}). Below are the correlation matrixes for molecular markers and geographic origins of red spruce[1] provenances in the following format:

R_{11}	R_{12}
R_{21}	R_{22}

where R_{11} displays the correlations between the molecular markers, R_{22} displays the correlations between the geographic variables and R_{12} and R_{21} displays the correlations between molecular markers and geographic variables.

	HMW 1.6	EMW 4.35	SMW 1.7	SITS	LAT	LON	LAT^2	LON^2	LAT*LON
HMW 1.6	1.00	0.268	-0.443	0.509	0.250	-0.146	0.250	-0.159	0.179
EMW 4.35	0.268	1.00	0.059	0.274	0.420	-0.305	0.425	-0.310	0.156
SMW 1.7	-0.443	0.059	1.00	-0.128	-0.069	-0.070	-0.063	-0.052	-0.283
SITS	0.509	0.274	-0.128	1.00	-0.111	0.074	-0.097	0.088	-0.108
LAT	0.250	0.420	-0.069	-0.111	1.00	-0.865	0.999	-0.883	0.166
LON	-0.146	-0.305	-0.070	0.074	-0.865	1.00	-0.865	0.998	0.346
LAT^2	0.250	0.425	-0.063	-0.097	0.999	-0.865	1.00	-0.881	0.159
LON^2	-0.159	-0.310	-0.052	0.089	-0.883	0.998	-0.881	1.00	0.304
LAT*LON	0.179	0.156	-0.283	-0.108	0.166	0.346	0.159	0.304	1.00

[1] The matrixes presented here were derived using only those red spruce individuals that were classified as red spruce using the three character molecular index. The analysis presented in Chapter 2 was derived using all the red spruce from the provenance test plantation with the exception of the individual identified as a multivariate outlier.

Canonical variates are calculated from the four matrixes as described by Gittins (1979), the number of canonical variates possible for a given analysis is equal to the number of variable in the smaller data set, four in the above example. The different canonical variates are orthogonal (uncorrelated). A summery of the canonical variates calculated when canonical correlation analysis was applied to red spruce provence samples is displayed below.

Variate	Canonical correlation (r)	r^2	Eigenvalue	F	Degrees of freedom	p
1	0.6256	0.39	0.6432	2.29	20	0.003
2	0.5722	0.33	0.4870	1.86	12	0.05
3	0.2489	0.06	0.066	0.75	6	0.6
4	0.2060	0.04	0.044	0.93	2	0.4

Appendix 4. Discriminant analysis is a multivariate technique used to classify individuals into a specific group. The data matrix is examined to determine which variables, or combination of variables, contribute to between group differences. Weighted linear combinations of the variables are derived for each group which maximizes between group variation while minimizing within group variation (Fisher 1936). The data in the test populations are then applied to each derived equation. Each individual is classified into the species according to the equation which the individual scored highest. Below are the weighted linear functions used to classify red and black spruce based on the nuclear rDNA allelic data.

B= -147.88 + 0.308*(HMW 3.4) - 0.038*(HMW 1.6) + 0.122*(EMW 4.35) + 1.011*(SMW 1.7) + 2.240*(SITS)

R= -95.234 + 0.117*(HMW 3.4) - 0.137*(HMW 1.6) + 0.066*(EMW 4.35) + 0.923*(SMW 1.7) + 1.827*(SITS)

The posterior probabilities are calculated as a ratio of scores (Klecka 1980) as follows:

$Prob_B = B / (B+R)$

Where $Prob_B$ is the posterior probability for an individual being classified as black spruce, B is the individual's score in the black spruce equation and R is the individual's score in the red spruce equation.

Appendix 5. Isolation of a black spruce rDNA clone.

A black spruce rDNA repeat unit was cloned in the vector pUC 18. The clone consisted of a Xba I fragment of the rDNA containing the entire coding region (Figure 1). Black spruce DNA, sample 4915 (2), was cut using Xba I and fragments were separated by size on a 0.6 % Seaplaque GTG (FMC Corporation) low melting temperature agarose gel. DNA fragments greater than 6.0 Kbp were isolated from the gel using the techniques described by Feinberg and Vogelstein (1984). The plasmid pUC 18 was cut with Xba I and dephosporalated using bacterial alkaline phosphatase (BAP) (BRL Life Technologies Inc.). Plasmid and genomic DNA fragments were purified by phenol/chloroform extractions and an ethanol precipitation. DNA fragments were ligated into pUC 18 with T4 DNA ligase (BRL Life Technologies Inc.) using the conditions specified by the manufacturer.

Ligated DNAs were transformed into Escherichia coli (E. coli) strain ER 1648 (New England Biolabs Inc.). E. coli strain ER 1648 is not sensitive to high methylation. Transformations were performed using high voltage electroporation as described by Dower et al. (1988). Transformed colonies were screened using colony lifts onto nylon membrane hybridized with the cloned soybean 18 S rDNA fragment (pXBR 1) as described by Sambrook, Fritsch and Maniatis (1989; section 1.96-1.104, book 1).

A black spruce clone was obtained. The clone was confirmed by restriction site analysis and by hybridizations to genomic southern blots. Hybridizing specific fragments of this clone (corresponding to the probes listed in Figure 1) yielded the same rDNA allelic frequencies observed with the soybean clones and spruce ITS 1 fragment.