Diss. ETH No. 12 156

Multivariate patterns of genetic variation in ponderosa pine (*Pinus ponderosa Dougl. ex Laws.*), based on allozyme scores and seedling quantitative traits

A dissertation submitted to the

SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZUERICH

for the degree of

Doctor of Natural Sciences

presented by

Peter Rotach

dipl. Forsting. ETH

born September 1, 1953

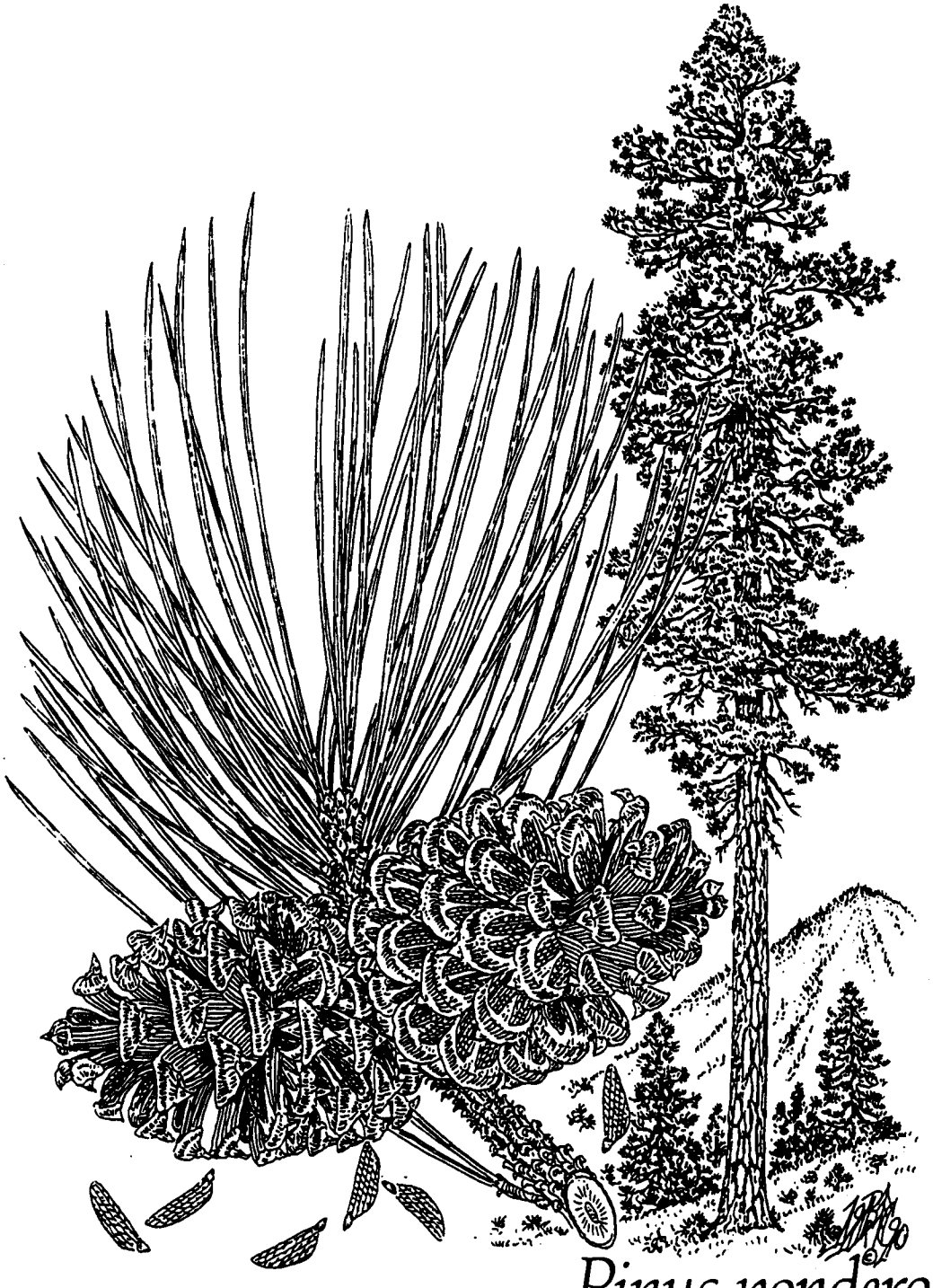
citizen of Schwellbrunn, AR

accepted on the recommendation of

Prof. Dr. J.-Ph. Schütz, examiner

Prof. Dr. W.T. Adams, co-examiner

1997



©1990 - Bruce Lyndon Cunningham *Pinus ponderosa*

with permission for reprint by the artist, 26 April 1991

Table of contents

List of figures

2. Ponderosa pine (*Pinus ponderosa Dougl. ex. Laws.*)

3. Materials and Methods

4. Genetic structure - measures of diversity and differentiation

List of tables

3. Materials and Methods

4. Genetic structure - measures of diversity and differentiation

5. Patterns of genetic variation - spatial structures and environmental associations of single alleles

6. Multivariate patterns of variation based on allozyme scores and associations with climate conditions

Abstract

The study describes and interprets patterns of genetic variation in ponderosa pine (*Pinus ponderosa Dougl. Ex. Laws.*) from two contrasting regions of the state of Oregon, USA, namely the east slope of the Cascade Range with parts of the high desert (Warner and Ochoco mountains) and the Klamath and Siskiou mountains in Southwest Oregon. Isozyme data as well as seedling quantitative traits of the same families were analyzed primarily by multivariate statistical techniques. Trend surface analysis was used to describe relationships of allozymes and metric traits with climate conditions at source location. The resulting multilocus allozyme surfaces were compared with those of metric traits and the utility of allozymes for describing patterns of adaptive variance was investigated. Resulting adaptive patterns were also utilized for estimates of relative seed-transfer risks and to map zones of certain transfer risks in geographic space. The two regions are described and compared with respect to their single-locus and multilocus genotypic structures. Measures of diversity and differentiation were analyzed with the traditional univariate procedures. Single alleles were also tested for spatial structures by spatial autocorrelation analysis and associations with climate variables were analyzed with multinomial response models and Mantel tests of matrix associations.

Ponderosa pine in Oregon maintains a high genetic diversity which is comparable to values reported for geographically adjoining parts of the range. Diversity estimates were in the upper range of values found for conifer species. Genetic diversity appeared to be high in quite small areas. Diversity was slightly higher in Southwest Oregon than in Central Oregon. Several loci showed a large differentiation among the two regions. Several alleles at many loci showed spatial variation patterns. Significant heterogeneities of allele frequencies combined with significant spatial patterns suggest that these patterns are rather the result of differential selection in different environments than a consequence of historic events (races).

Multilocus genotypic frequencies were moderately associated with habitat conditions. Climatic conditions at source location explained 33% of total variance of 41 allozyme variables. Within the regions, explained variance was about 23%. Overall, alleles at 13 out of 31 loci showed an association with climatic conditions. These results clearly argue for the adaptivness of certain alleles or allozyme loci. Especially enzyme systems involved in important physiological pathways such as MNR-1, MNR-2, LAP-2, PEP-3, G6P-1, ACP-1, GDH-1, IDH-1, PGM-1, SKD-2 and MDH-3 appear to be adaptive. Of the studied alleles, 42% were unrelated to the pattern, 16% showed a correlation of more than 10%, 31% of more than 20% and 11% of more than 30% with the adaptive pattern. Single alleles thus seem to behave like quantitative trait loci, each gene contributing only small amounts to the pattern of adaptive variance. All results clearly indicate that temperature at source location seems to be the most important environmental factor responsible for the observed adaptive patterns. Moisture characteristics of the site had only a minor overall effect on genotypic variation, except in Southwest Oregon. In Southwest Oregon, patterning was weak, highly complex and both temperature as well as moisture conditions of the site showed about an equal influence on multilocus frequencies. Strong moisture gradients combined with a relatively mild climate may lead to a more important influence of moisture characteristics of the site in this area. The distinct geologic substrate and the soil conditions may also play an important role. The adaptive pattern of allozyme variation in Central Oregon was primarily related to temperature. The pattern of multilocus frequencies was nearly identical with patterns based on seedling quantitative traits published by SORENSEN (1994).

Results from multivariate analyses suggest that observed differentiation among the two regions is nearly exclusively the result of natural selection in two contrasting environments than the consequence of evolutionary events. The existence of two races (Pacific race, North Plateau race) as a consequence of a different evolutionary past, as suggested by several authors, seems highly unlikely.

Transfer risk estimates revealed that seed should not be transferred across the Cascade Range. Within Southwest Oregon, seed should not be moved in the east-west direction. Current zones appear to be too large to guarantee an acceptable transfer risk. Seed transfer should rather be based on models of transfer risk than on discrete seed zones, since the delineation of such zones seems to be a highly complex task. The current seed zones in Central Oregon do not seem to reflect the major adaptive patterns. Current zonation seems to be con-

servative. Only 8 zones would be required to guarantee a maximum relative transfer risk of 30%.

As much as 77% of the patterned variation in growth and phenology traits could be explained by the climate conditions at source location. The major pattern of adaptive variance related growth vigor and growth timing to temperature conditions at source location. Growth vigor seems to be the result of adaptation to temperature constraints at source location. Since elongation potential appears to be genetically fixed, selection seems to act on phenology traits, attuning the sources to the length of the growing season. Within Southwest Oregon, patterns of adaptive variance were rather complex. Growth potential and growth timing were significantly but weakly associated with temperature at source location. A small scale, ecotypically organized pattern was apparent. Even if temperature clearly was the most important factor, about 25% of the significant model variance was related to the moisture characteristics of the site. Growth vigor and growth timing of the sources in Central Oregon were strongly related to temperature. Of the variation in growth potential, 45% could be significantly explained by temperature at source location. Growth potential appeared to be genetically fixed.

Based on quantitative traits, seed should not be moved across the Cascade Range. As a general rule, seed should not be moved in the west-east direction within Southwest Oregon and not in a northwest-southeast respectively southwest-northeast direction within Central Oregon. Due to a rather ecotypic pattern, even short distance transfers within Southwest Oregon may produce a considerable mismatch between source locations and planting sites. Current seed zones do not seem to adequately portray the major patterns. For ponderosa pine, seed zones in Central Oregon could be considerably larger in longitude and latitude than the current zones. A maximum of 7 to 8 zones are needed to guarantee a transfer risk smaller than 30%; for practical reasons 5 major zones seem to be sufficient.

Although quantitative traits reflected much higher proportions of adaptive variance than allozymes, the adaptive patterns in allozymes clearly approximated the patterns for seedling traits. For total sampling area, both patterns portrayed the same major ecological gradients of adaptive significance. With a coefficient of correlation of $R = 0.842$, predicted scores on the first allozyme variate were highly correlated with scores on the first trait variate. Patterns were also highly similar within Central Oregon ($R = 0.66$). With a coefficient of correlation of $R = 0.142$, predicted scores were, however, only weakly correlated within Southwest Oregon. According to the distinct situation in this area, several reasons for the low congruence of the two patterns are conceivable.

Allozyme data can contribute useful information to the understanding of adaptive patterns of genetic variation and may be of practical help for seed-zone formation or transfer risk estimates.

Kurzfassung

Die vorliegende Arbeit untersucht und beschreibt genetische Variationsmuster von *Pinus ponderosa Dougl. Ex Laws.* in zwei gegensätzlichen Regionen von Oregon, USA, einerseits in Zentral Oregon d.h. an den Ostabhängen der Kaskaden inklusive der Gebiete der Ochoco und Warner Mountains, andererseits in Südwest Oregon auf der Westseite der Kaskaden mit den Klamath und Siskiou Mountains. An genetischen Daten wurden Isoenzym-Marker und phänotypische Merkmale von Sämlingen derselben Mutterbäume analysiert, wobei schwerpunktmässige multivariate statistische Methoden verwendet wurden. Zusammenhänge zwischen Allelfrequenzen und Sämlingsdaten mit den Klimaverhältnissen des Ursprungsortes wurden mittels kanonischer Korrelationsanalyse untersucht, und die resultierenden Anpassungsmuster wurden als dreidimensionale Ebenen («trend surfaces») über den geographischen Koordinaten dargestellt. Die resultierenden multivariaten Variationsmuster von Allelhäufigkeiten wurden mit den Variationsmustern von Sämlingsmerkmalen verglichen, um deren Übereinstimmung zu prüfen und daraus die Nützlichkeit von Isoenzym-Markern für die Beschreibung von Anpassungsmustern zu beurteilen. Die Anpassungsmuster dienten zudem dazu, potentielle Risiken beim Verschieben von Saatgut zu schätzen und kartographisch darzustellen. Die so entwickelten Karten mit Isolinien bestimmter Risiken ermöglichen es, die Qualität der vorhandenen, rein geoklimatisch hergeleiteten Saatgutzonen für *Pinus ponderosa* zu beurteilen. Die beiden Regionen werden sowohl in Bezug auf ihre Einzel-Locus- wie Mehr-Locus-Strukturen verglichen. Für Einzel-Locus-Vergleiche wurden traditionelle Diversitäts- und Differenzierungsmasse verwendet. Die einzelnen Allele wurden zudem mittels Autokorrelationsanalyse auf das Vorhandensein von räumlichen Variationsmustern geprüft und Zusammenhänge zwischen Allelhäufigkeiten und Klimavariablen wurden mit multinomialen Modellen und mittels Mantel-Test-Statistik untersucht.

Pinus ponderosa in Oregon verfügt über eine hohe genetische Diversität, die im oberen Bereich liegt, der für Koniferen beschrieben wurde. Die gefundenen Werte sind vergleichbar mit Diversitätswerten in den benachbarten Gebieten. Die Diversität war in Südwest Oregon leicht höher als in Zentral Oregon. An mehreren Genorten zeigte sich eine erhebliche Differenzierung zwischen den beiden Regionen. Mehrere Allele an verschiedenen Genorten wiesen räumliche Variationsmuster auf. Die Kombination von heterogenen Allelhäufigkeiten mit räumlichen Variationsmustern lässt vermuten, dass die vorhandenen Variations- bzw. Differenzierungsmuster in erster Linie das Resultat von natürlicher Selektion sind und weniger die Folge von historischen Ereignissen.

Die Mehr-Locus-Genotypenfrequenzen zeigten moderate Zusammenhänge zu den klimatischen Bedingungen des Herkunftsortes. Die klimatischen Bedingungen erklärten 33 % der Gesamtvarianz von 41 allelischen Variablen. Innerhalb der Regionen betrug der erklärbare Anteil 23 % der Gesamtvarianz. Allelhäufigkeiten an 13 von insgesamt 31 Genorten zeigten eine Beziehung zu den klimatischen Bedingungen des Ursprungsortes. Diese Ergebnisse sprechen für die Adaptivität gewisser Isoenzyme. Speziell Allele an Genorten, deren Enzyme wichtige physiologische Abläufe katalysieren wie MNR-1, MNR-2, LAP-2, PEP-3, G6P-1, ACP-1, GDH-1, IDH-1, PGM-1, SKD-2 und MDH-3, scheinen adaptiv zu sein. 42 % der untersuchten Allele zeigten keinen Zusammenhang, 16 % einen von mehr als 10 %, 31 % einen von mehr als 20 % und 11 % der Allele einen von mehr als 30 % zum adaptiven Variatiosmuster. Enzym-Genorte verhalten sich daher mehrheitlich wie Genorte, die quantitative Merkmale kodieren; jedes Gen trägt lediglich einen kleinen Anteil zur Gesamtvarianz bei. Alle Ergebnisse deuten klar darauf hin, dass die Temperaturverhältnisse des Herkunftsortes als die wichtigsten Selektionsfaktoren für das Entstehen der vorhandenen Anpassungsmuster zu betrachten sind. Niederschlag und Wasserangebot hingegen scheinen im allgemeinen nur einen unbedeutenden selektiven Einfluss auf die genotypischen Strukturen auszuüben. Südwest Oregon unterschied sich diesbezüglich allerdings von Zentral Oregon. Die Zusammenhänge zwischen Mehr-Locus- Genotypenstruktur und Klimaverhältnissen waren in Südwest Oregon weniger stark ausgeprägt und Temperatur- und Niederschlagsverhältnisse zeigten einen Zusammenhang von etwa gleicher Stärke. Stark ausgeprägte Niederschlagsgradienten in Verbindung mit relativ milden Temperaturverhältnissen mögen verantwortlich dafür sein, dass die Niederschlagsverhältnisse in diesem Gebiet eine wichtigere Rolle spielen. Zusätzlich mögen die speziellen geologischen Verhältnisse die selektive Bedeutung des Wasserfaktors verstärken. Das Anpassungsmuster in Zentral Oregon war hauptsächlich durch die Temperaturverhältnisse bestimmt. Das Muster, welches sich aus den Mehr-Locus-Genotypenfrequenzen ergab, war praktisch identisch mit

dem von SORENSEN (1994) für das selbe Gebiet puplizierten Anpassungsmuster, hergeleitet von Sämlingsmerkmalen.

Aufgrund der Ergebnisse der multivariaten Analysen scheint die beobachtete Differenzierung in den allelischen Strukturen zwischen den beiden Gebieten vor allem das Resultat von natürlicher Selektion in zwei gegensätzlichen Umwelten zu sein und nicht die Konsequenz einer unterschiedlichen evolutiven Vergangenheit. Die Annahme, dass auf beiden Seiten der Kaskaden zwei verschiedenen Rassen («Pacific race», «North-Plateau race») als Folge einer verschiedenen Vergangenheit existieren, wie sie von verschiedenen Autoren postuliert worden ist, erscheint wenig wahrscheinlich.

Gestützt auf die Isoenzym-Anpassungsmuster sollte Saatgut nicht über die Kaskaden hinweg verschoben werden. Innerhalb von Südwest Oregon sollte Saatgut nicht in West-Ost Richtung transferiert werden. Die existierenden Saatgutzonen scheinen hier für *Pinus ponderosa* zu gross zu sein, um ein akzeptables Risiko zu garantieren. Die Saatgut Verwendung in diesem Gebiet sollte nicht durch fixe Saatgutzonen, sondern eher durch Modelle beschrieben werden, da die Abgrenzung von diskreten Zonen sehr schwierig scheint. Die vorhandenen Saatgutzonen in Zentral Oregon scheinen die wesentlichen Anpassungsverhältnisse ebenfalls schlecht widerzuspiegeln. Aufgrund der Isoenzym-Anpassungsmuster scheinen sie sehr konservativ zu sein. Für ein Risiko von maximal 30 % wären lediglich 8 Saatgutzonen notwendig.

Bis zu 77 % der Variation in den Sämlingsmerkmalen liess sich mit den Klimaverhältnissen am Ursprungsort erklären. Die wichtigste Beziehung ergab sich zwischen Wuchskraft bzw. phänologischen Merkmalen und den Temperaturverhältnissen. Die Wuchskraft scheint das Ergebnis einer Anpassung an die Temperaturen des Ursprungsortes zu sein. Die natürliche Selektion wirkt offensichtlich auf phänologische Merkmale, indem sie die Herkünfte an die Länge der Vegetationsperiode anpasst. Die Anpassungsmuster innerhalb von Südwest Oregon waren wiederum komplex. Obwohl die Temperatur auch hier der wichtigste selektive Umweltfaktor war, waren etwa 25 % der Variation der Sämlingsmerkmale mit der Wasserbilanz des Standortes erklärbar. In Zentral Oregon waren die Temperaturverhältnisse hingegen klar der wichtigste Faktor; 45 % der Variation in der Wuchskraft bzw. in den phänologischen Merkmalen liessen sich mit den Temperaturverhältnissen erklären.

Aufgrund der Anpassungsmuster der Sämlingsmerkmale ergab sich im allgemeinen eine identische Risikobeurteilung für die Saatgutverwendung wie sie gestützt auf Isoenzym-Muster eingeschätzt worden ist.

Obwohl die Beziehung zwischen Klimaverhältnissen und Variation von Sämlingsmerkmalen stärker war als jene zwischen Klima und Mehr-Locus-Genotypenfrequenzen, glichen sich die beiden multivariaten Anpassungsmuster sehr stark. Beide Anpassungsmuster widerspiegelten im wesentlichen dieselben Gradienten von adaptiver Bedeutung. Die Werte auf den beiden "trend surfaces" waren, mit Ausnahme von Südwest Oregon ($R = 0.142$), hoch korreliert (Gesamtgebiet: $R = 0.842$, Zentral Oregon: $R = 0.66$). Mehrere Ursachen für die geringe Übereinstimmung in Südwest Oregon sind denkbar und werden diskutiert.

Isoenzym Daten können nützliche Information für das Verständnis von Anpassungsmustern liefern. Sie können auch für praktische Problemstellungen wie das Ausscheiden von Saatgutzonen oder für Modelle zur Risikobeurteilung eingesetzt werden.

1. Introduction

1.1 General introduction

A major focus of forest genetics is the assessment of patterns of genetic variation within and between populations. Knowledge of genetic variation patterns as well as information on the degree of adaptation to contrasting environments, is essential for sound management practices, effective seed zone delimitation, seed and plant transfer guidelines and tree improvement programs. Gene ecology, therefore, provides an important basis of knowledge for all silvicultural management decisions.

Throughout the ages, natural selection has physiologically attuned populations to their environments. Therefore, for each population there exists an optimal range of environments for which individuals and their offspring are adapted. Movement of seed or plants beyond their optimal range results in maladaptation and exposes them to risk of poor survival or growth. Optimal productivity from artificial reforestation or regeneration, as well as management investments, are guaranteed only if maladaptation can be avoided or limited to a small extent. To accomplish this, transfer of seed and plant material must be controlled and must be based on knowledge of genetic variation patterns.

Provenance trials, followed over a long period of time, potentially can yield the best information on genetic variation, but are limited by time, money and the labor needed to perform such experiments. Knowledge based on such provenance trials is therefore sparse and often limited to few provenances and test sites. In addition, the results often cannot be generalized or extrapolated to a wide area of site conditions due to genotype-by-environment interactions.

An alternative method of gaining insight into genetic variation patterns over a variety of environments is seedling common garden studies. Seeds from different mother trees, sampled over a part of the species range, are collected and sown under controlled conditions in one or more nursery test environments. On the basis of a variety of quantitative traits measured on the seedlings over a period of 1 to 5 years, genotypic values of the mother trees are estimated and patterns of genetic variation are mapped in terms of topography and other indexes of environmental variation (CAMPBELL, 1979, 1986, 1991; CAMPBELL and SUGANO, 1987; REHFELDT, 1979, 1986, 1989, 1990).

Yet another method of detecting and describing genetic variation is the use of molecular gene markers such as proteins, DNA fragments (restriction fragment length polymorphisms (BOTSTEIN et al., 1980) or random amplified polymorphic DNA (WILLIAMS et al., 1990)). Electrophoresis of isozymes, a technique which has become widespread since the late 1960's, has been extensively used to estimate and describe genetic variation in plants and animals. With this technique, enzyme proteins are separated in a gel medium using an electric current. Since enzymes are coded by DNA segments found at certain chromosome locations (loci) they may be used as markers for genes. Various forms of enzymes (allozymes) are expressed when alternative gene variants (alleles) are present. Allozymes differ in their electrical charge and migrate at different rates through the gel, and can thus be separated in an electric field by electrophoresis (SOLTIS and SOLTIS, 1989; CONKLE et al., 1982; O'MALLEY et al., 1980). Allozyme data allow accurate estimates of genetic variation and other population genetic parameters. DNA markers differ from isozyme markers in that they represent variation in DNA sequences themselves (by cutting the DNA at restriction sites using restriction enzymes and separating the fragments with electrophoresis), whereas isozymes are transcribed products of DNA sequences. DNA markers are essentially used for the same purpose as isozymes, but they extend the measurable variation to the total genome whereas isozyme markers are limited to detectable enzymes.

A large number of allozyme studies describing patterns and amounts of genetic variation in natural populations of forest trees have been reported. These studies have been directed toward a variety of goals; e.g., assessing the amount of variation present, defining population structure, defining taxonomic relationships, and demonstrating or testing the adaptive significance of protein variation. Results of these studies are reviewed by HAMRICK et al. (1981, 1992), HAMRICK and GODT (1990), MITTON (1983) and STRAUSS et al. (1992). Allozymes

might also be useful for establishing breeding zones or seed zone boundaries, gene conservation areas and for other purposes such as certifying the source of seedlots or for the identification of clones in breeding programs (FERET and BERGMANN 1976; ADAMS 1981; CONKLE and WESTFALL 1984; MILLAR and WESTFALL 1992; WESTFALL and CONKLE, 1992; WESTFALL 1992). Mapping of genetic patterns of geographical variation can be carried out quickly and easily from allozyme analysis. To be useful for mapping adaptive variation, allozyme loci or genes closely linked to these loci must be subject to selection pressure. For seed certification, populations must at least be discernible on the basis of their allozymes.

Twenty-five years after the first occurrence of electrophoretic studies using isozymes, the adaptive significance of protein polymorphism is still controversial. In fact, few subjects in biology have been more strongly debated than the evolutionary significance of protein polymorphism. Most of the debate centers around two opposing views: The "selectionist" and the "neutralist". The selectionist view asserts that natural selection maintains and favors protein polymorphism, whereas the neutralist view assumes that the vast majority of such variation is selectively neutral or nearly neutral and is just an accumulation of random mutations (KIMURA, 1968, 1983; KIMURA and OHTA, 1971; KING and JUKES, 1969). The ecological significance of observed allozyme variation is therefore still unclear. If the neutral theory is correct, then the observed variation is more or less irrelevant to adaptive evolution, at least in the ecological conditions prevailing at present. Consequently, allozyme gene markers are of little use for ecological genetics in describing and explaining environmentally significant variation patterns.

1.2 Adaptive significance of protein polymorphisms

The nature of protein polymorphism has been constantly debated since isozyme techniques were applied to study genetic variation in the mid 1960s. The numerous studies that followed made evident that protein structure varies substantially within natural populations. At first, this newly discovered variation seemed to confirm the importance of balancing selection. However, it was soon realized that selection, acting simultaneously at so many loci, would depress the mean fitness to such an extent that populations would be unable to survive. This led to the theory that most protein variation is selectively neutral (KIMURA, 1983). The controversy has continued since and was reviewed recently by NEVO et al. (1984), NEI and GRAUR (1984) and NEI (1987). More than 100 proteins from over 1000 species of plants and animals have meanwhile been surveyed and the question is still unresolved. Apart from theoretical and mathematical considerations and predictions, there have been two main approaches to testing selectionist and neutralist theories with real-world data: 1) Macro and micro geographic surveys of protein variation combined with inferences about the associations of protein variation with environmental variation, and 2) studies of biochemical kinetics and physiological functions of enzymes.

In the majority of investigations where the adaptive significance of allozymes has been studied, no adaptive variation has been demonstrated. Allozyme polymorphisms thus appear to be mostly neutral to selection pressure. The vast amount of literature where no adaptive variation has been demonstrated, cannot be reviewed here; instead only a number of results showing exceptions or good support for occasional polymorphisms having significant adaptive consequences will be cited. The following presentation thus has mere practical reasons; it does not intend to adopt a selectionist perspective.

The techniques used to demonstrate genetic-environmental patterns include geographic (spatial) mapping of allele frequencies and gene-environment associations. Spatial allozyme patterns associated with climate have been described, for example, for *Avena barbata* (CLEGG and ALLARD, 1972; KAHLER et al., 1980), *Hordeum spontaneum* and *Triticum dicoccoides* (NEVO et al., 1981, 1988), *Oenothera biennis* (LEVY and LEVIN, 1975), *Lycopersicon pimpinellifolium* (RICK et al., 1977) and many other herbaceous plants. In animal species, associations of protein polymorphisms, associated with environmental variation, were found for the freshwater fish *Castostomus clarkii* (KOEHN, 1969), the harvester ant *Pogonomyrmex barbatus* (JOHNSON et al., 1969), the fruit fly *Drosophila melanogaster* (JOHNSON and SCHAFFER, 1973, OAKESHOTT et al., 1982), the fathead minnow *Pimephales proleus* (MERRITT, 1972), the marin fish *Fundulus heteroclitus* (MITTON and KOEHN, 1975; POWERS et. al., 1991; CASHON et al., 1981) and many others. These associations of protein poly-

morphism with environmental variation are primarily descriptive rather than deterministic. They do not establish direct causal relationships between genetic variation and environment, however they inferentially demonstrate that enzyme polymorphisms, or at least parts of it, might be maintained or influenced by some form of natural selection.

Complementary evidence for a possible, occasional adaptive significance of enzyme polymorphism comes from a large number of biochemical and physiological studies on allozymes. Many of these studies have demonstrated *in vitro* kinetic differences between allozyme variants (KOEHN, 1969; GIBSON, 1970; MERRITT, 1972; HARPER and AMSTRONG, 1973; PLACE and POWERS, 1979; GROSSMAN, 1980; GRAVES and SOMERO, 1982; GRAVES et al., 1983; BERGMANN and GREGORIUS, 1993; MITTON, 1983; MITTON et al., 1977; NARISE, 1979; CHARLES and LEE, 1980; HOFFMAN, 1981; POWERS et al., 1991). *In vivo* differences among allozyme variants, suggesting that at least some of the *in vitro* kinetic differences might give rise to *in vivo* differences affecting fitness, have been reported, among others, by KOEHN (1978), PAPEL et al. (1979), ANDERSON and MCDONALD (1981), CAVENER and CLEGG (1981), DIMICHELE and POWERS (1982a), MARSHALL et al. (1973), POWERS et al. (1991), CLARKE (1975), SCHWARTZ (1960,1973), SCHWARTZ and LAUGHNER (1969), BROWN et al. (1976), RAINEY et al. (1987).

Fitness differences related to differences in allozymes have been reported for many species. In *Drosophila*, kinetic differences between allozymes of alcohol dehydrogenase were associated with differences in survival and developmental time (DALY and CLARKE, 1981; DORADO and BARBANCHO, 1984; VIGUE et al., 1982). Kinetic differences in alpha-glycerophosphate dehydrogenase allozyme variants were correlated with environmental temperature, flight metabolism and power output (CURTSINGER and LAURIE-AHLBERG, 1981; MILLER et al., 1975; O'BRIEN and MACINTYRE, 1972; SACKTOR, 1975). Allozyme activity and other kinetic parameters of esterase-6 were associated with reproduction (RICHMOND et al., 1980). Differences in metabolic flux of glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase were reflected in fitness differences (HUGHES and LUCCHESI, 1977; BIJLSMA, 1978; CAVENER and CLEGG, 1991). In *Mytilus edulis,* osmoregulation and phenotypic fitness were associated with activity differences between allozymes of aminopeptidase (HILIBISH and KOEHN, 1985; KOEHN and IMMERMANN, 1981). In *Colias* butterflies, kinetically different allelic variation of glucose phosphate isomerase was correlated with survival and mating success (WATT et al., 1983, 1985). In *Metritium senile*, differences in allozyme variants were associated with a different modulation of the pentose-shunt metabolism (ZAMER and HOFFMAN, 1989). Allelic activities of glutamate-pyruvate transaminase in *Tigriopus californicus* were associated with different rates of alanine accumulation and different responses to hyperosmotic stress (BURTON and FELDMANN, 1983). In *Fundulus heteroclitus*, differences in the kinetics and other biochemical properties of several enzymes involved in the glycolysis, the pentose shunt and the malate/isocitrate shuttles were reflected in differences in metabolism, oxygen transport, swimming performance, developmental rate and relative fitness (DIMICHELE et al., 1986, 1991; DIMICHELE and POWERS, 1982a, 1982b, 1984, 1991; PAYNTER et al., 1991; PLACE and POWERS, 1979, 1984a, 1984b; ROPSON and POWERS, 1988, 1989; VAN BENEDEN et al., 1989). The observed differences in fitness components associated with molecular variants support a possible adaptive significance of certain enzyme polymorphisms since natural selection might act on such fitness traits.

An adaptive significance of certain enzyme polymorphisms is also supported by observations that environmental and physiological stress may cause differences in fitness among allozyme variants of enzymes such as amylase and phosphoglucomutase in marine organisms (NEVO et al., 1977, 1978). Fitness differences could also be reproduced in laboratory experiments, demonstrating differential survivorship of allelic isozyme variants, singly or in multilocus structures, caused by single or two interacting pollutants (NEVO et al., 1980, 1981, 1983).

Numerous studies showing positive correlations between allozyme heterozygosity and fitness traits in many different species may be viewed as an indirect evidence for the adaptivness of protein polymorphisms (for an overview see ZOUROS and FOLTZ, 1987). Positive correlations with fitness could indirectely imply an advantage of having multiple alleles, which is in contradiction to the neutral theory. According to the prediction of neutrality, different molecular forms should not differ in their kinetic or functional properties, at least not for enzymes with strong

functional constraints, and heterozygote superiority therefore is unlikely to occur. Consequently, the occurrence of heterozygote superiority at a molecular locus may be taken as an indication that functional differences at the molecular level are non neutral and that such loci might be under natural selection since they seem to be related to fitness traits of the individual. Alternatively, however, allozyme heterozygosity may only be reflecting levels of inbreeding over the entire genome and positive associations with fitness traits would thus be a mere consequence of different inbreeding effects on these traits.

Positive correlations of allozyme heterozygosity and fitness traits such as growth, viability, fertility or homeostatic stability have been reported in many species and for many enzyme systems. For example, the fitness of homozygotes at the *Adh* locus in a population of *Drosophila melanogaster* appeared to be about 90% of of heterozygotes (MUKAI and YAMAZAKI, 1980). Positive correlations between heterozygosity and growth were found in plants and animals. Among plant species, growth rate was positively associated with allozyme heterozygosity in *Litaris cylindrica* (SCHAAL and LEVIN, 1976), *Pinus rigida Mill.* (LEDIG et al., 1983; BUSH et al., 1987), *Populus tremuloides* (MITTON and GRANT, 1980), *Zea mays* (KAHLER and WEHRHAN, 1986; EDWARDS et al., 1987) and many others. Among animal species, correlations of growth and allozyme heterozygosity have been reported for American oyster (*Crassostrea virginica;* SINGH and ZOUROS, 1978; ZOUROS et al., 1980; FOLTZ et al., 1983), blue mussel (*Mytilus edulis;* KOEHN and GAFFNEY, 1984), mollusks (*Macoma balthica;* GREEN et al., 1983 and *Haliotis discus;* FUJINO, 1978), mosquito fish (*Gambusia affinis;* SMITH and CHESSER, 1981), herring (*Clupea herengus;* KING, 1985), sheep (BAKER and MANWELL, 1977), pig (MAKAVEEV et al., 1977), mouse *(Peromyscus polionotus;* GARTEN, 1976) and many others.

Positive associations between heterozygosity and age have been found in many species, suggesting that heterozygote individuals may exhibit higher survival rates (ELLSTRAND et al., 1978; SCHAAL and LEVIN, 1976; FARRIS and MITTON, 1984; WATT, 1977; ZOUROS et al., 1983; GREEN et al., 1983; DIEHL et al., 1985; MITTON and KOEHN, 1975). In addition, there are many reports of size-dependent differences in mean heterozygosity found in natural populations, implying that heterozygote individuals may not only exhibit better growth but also better survival. Species in which single- or multilocus heterozygosity increase with size, include the abalone (*Haliotis discus*; FUJINO and SASAKI, 1984), mussels (*Modiolus demissus;* KOEHN et al., 1973 and *Mytilus edulis*; JOHNSON and UTTER, 1975), oyster (*Crassostrea gigas;* FUJIO, 1982), sand flounder (*Rhombosolea plebeia;* SMITH and FRANCIS, 1984) and many others. In the blue grouse (*Dendragapus obscurus)*, heterozygote individuals had an advantage in viability (REDFIELD, 1974). The same result was found for western toad (*Bufo boreas)* by SAMOLLOW and SOULE (1983). Among chickens, heterozygous individuals at the major histocompatibility complex B locus show an advantage in viability over homozygotes during embryonic development (FUJIO, 1971; MORTON et al., 1965).

Heterozygote advantage in fertility was reported for *Litaris cylindrica* by SCHAAL and LEVIN (1976). Flowering plants were more heterozygous than non-flowering individuals and the number of ovules was positively associated with heterozygosity at 15 allozyme loci. In *Pinus ponderosa Dougl. ex Laws.*, trees with high fertility (cone production) showed an excess of heterozygotes at three loci, compared to Hardy-Weinberg expectations (LINHART et al., 1979). In *Drosophila euronotus*, heterozygous females caught in the wild carried more sperm and produced more progeny than did homozygous females (STALKER, 1976). Male *Colias* butterflies, heterozygous at *Pgm* and *Pgi* loci, had a mating advantage over males homozygous at both loci (WATT et al., 1985). In pure-bred pigs, average litter size was positively associated with paternal and maternal heterozygosity (RASMUSEN and HAGEN, 1979). The calving interval in cattle was found to decrease with increasing maternal heterozygosity. In addition, the total number of calves born per female increased with heterozygosity (HIERL, 1976; SCHLEGER et al., 1978). For the white tailed deer (*Odocoileus virginianus*) JOHNS et al. (1979) found that females carrying twins were significantly more heterozygous than those carrying a single fetus.

Differences between heterozygous and homozygous individuals in performance and metabolism have been reported in many studies. Vegetatively regenerated *Norway spruce* cohorts in high altitude locations were mostly heterozygous for a 6-phosphogluconate dehydrogenase locus whereas individuals grown from seed were all homozygous at the same locus. Higher ca-

pacities in energy production, hormone transfer or better nutrient allocation of the founder tree, which had to support the new vegetative offsprings over a long period of time, is postulated as a possible explanation (STIMM and BERGMANN, 1994). In the oyster (*Crassostrea virginica)*, oxygen consumption was least in more heterozygous individuals (KOEHN and SHUMWAY, 1982). A similar correlation was reported for *Mytilus edulis* by DIEHL et al., (1985). Growth efficiency increased with individual heterozygosity for the clam, *Mulinia lateralis* (GARTON et al., 1984). In the gastropod, *Thais haemastoma*, heterozygotes had a greater metabolic efficiency than homozygotes (GARTON, 1984).

The neutral theory, in all its modifications, asserts that most molecular diversity in nature is non- selective and that this diversity is maintained in populations through mutational input and random fixation since molecular variants are equal or similar in function. While neutrality and random processes may be important factors in molecular evolution, they can hardly explain the rich genetic diversity prevalent in natural populations nor the differences of fitness among molecular forms found in the many cases cited above. Even if the cited examples for a possible adaptive significance of allozyme variation are rare given all the loci and the vast amount of species studied, they provide good support for occasional allozyme polymorphisms having siginficant adaptive consequences. Consequently they could be potentially useful for mapping ecologically significant adaptation patterns in nature.

1.3 Enzyme polymorphisms in forest tree species

Studies of allozyme polymorphism have shown that forest trees are among the most genetically variable organisms known (HAMRICK and GODT, 1990). There has been a great deal of interest in whether observed allozyme variation reflects the work of natural selection and consequently would be useful for describing the mode and extent of population adaptations. The search for adaptive explanations for the large amount of genetic diversity observed in most populations has met with only limited success. Moreover, in spite of considerable work in this area, there is still very limited knowledge about the phenotypic effects associated with allozyme variation in forest trees. The large number of allozyme studies in forest tree cannot be reviewed in detail. A summary of the most important results and a number of illustrating examples will be given instead.

When allozyme data over many loci are summarized and expressed in a composite index of differentiation, most allozyme variation is found within rather than among populations, even when samples are taken across the geographic range of the species. Low levels of interpopulational variation signify that most of the alleles are found in similar frequencies in the different populations sampled. Exceptions to this generalization have been found for red pine (*Pinus resinosa Ait.;* FOWLER and MORRIS, 1977; ALLENDORF et al., 1982; SIMON et al., 1986), torrey pine (*Pinus torreyana Parry ex Carr.;* LEDIG and CONKLE, 1983) and western red-cedar (*Thuja plicata Donn ex D. Don;* COPES, 1981) which all show little variation, and for stone pine, *Pinus cembra L.*, where 32% of the variation resided among populations (SZMIDT, 1982) or Douglas-fir, *Pseudotsuga menziesii (Mirb) Franco.*, which has a high level of differentiation over its entire range (LI and ADAMS, 1989). Genetic drift as well as bottlenecks were hypothesized to explain the lack of variability for red pine, torrey pine and western red-cedar, while geographic isolation and possibly separate refugia are believed to explain the high differentiation among populations of stone pine.

Population differentiation is commonly measured using the univariate approach by NEI (1973) or WRIGHT (1965), in which all loci are weighted equally and the mean over all loci is calculated to give the coefficient of genetic differentiation G_{st} (=F_{st}). The coefficient G_{st} measures the proportion of total gene diversity found among populations; it therefore reflects the degree of differentiation which might be due to natural selection or genetic drift.

The average G_{st} value was estimated as 0.076 for 23 wind pollinated conifer species and as 0.075 for nine species of angiosperms such as Alnus, Populus and Quercus (MUONA, 1990). According to these G_{st} values, 7 percent of the total gene diversity as detected by allozyme markers, on average reside among different populations, while 93% is found within populations. With the exceptions mentioned above, interpopulational differentiation in forest tree species

ranges generally between 1% and 15% of the total gene diversity as the following examples demonstrate.

G_{st} values for lodgpole pine, *Pinus contorta Dougl. ex. Loud.,* range from 1% to 6% (KNOWLES, 1984; WHEELER and GURIES, 1982; DANCIK and YEH, 1983; YEH and LAYTON, 1979). Similar values have been reported for Ponderosa pine (*Pinus ponderosa Dougl. ex Laws.*), however, the reported values vary substantially: O'MALLEY et al. (1979) found 12% inter-population differentiation for 10 populations sampled in the Rocky Mountains, while WOODS et al. (1983) reported only 1.5% for six small, isolated stands located within a nine-kilometer radius. G_{st} values for Monterrey pine, *Pinus radiata D. Don,* vary between 5.1% and 16.2% for different studies (PLESSAS and STRAUSS, 1986; MORAN et al., 1988). The differentiation among ecologically different stands is as low as 2%. G_{st} values for Scots pine (*Pinus sylvestris L.*), vary widely between 2% and 16% (GULLBERG et al., 1985; RUDIN et al., 1974; MEJNARTOWITCZ, 1979). In black pine, *Pinus nigra Arnold,* interpopulational differentiation was 6% for five populations sampled in the two different subspecies, *nigra* and *laricio,* while G_{st} values within subspecies were only 0.5% to 2% (SCALTSOYIANNES et al., 1994). For Norway spruce, *Picea abies (L.) Karst.*, estimates of G_{st} varied between 2% and 6% (LUNDKVIST and RUDIN, 1977; TIGERSTEDT, 1974; BERGMANN, 1974, LAGERCRANTZ and RYMAN, 1990). According to different studies, black spruce (*Picea mariana (Mill.) B.S.P.*), showed between 1% and 7% interpopulational differentiation (BOYLE and MORGENSTERN, 1987; DESPONTS and SIMON, 1987; O'REILLEY et al., 1985; YEH et al., 1986; YEH, 1981). Differentiation among populations of sitka spruce, *Picea sitchensis (Bong.) Carr.*, was reproted to be 7.9% (YEH and EL-KASSABY, 1980). The values for Douglas-fir range between 0.1% and 7.1% for coastal, and between 4.3% and 12.2% for interior populations (MERKLE and ADAMS, 1987; LI and ADAMS, 1989). The low value of 0.1% was reported for 22 coastal Douglas-fir populations sampled in a small area in Southwest Oregon (MERKLE and ADAMS, 1987). For populations from different elevations, EL-KASSABY and SZIKLAI (1982) reported a differentiation coefficient of 6.8%. Over the entire range of coastal Douglas-fir, LI and ADAMS (1989) found a very high G_{st} value of 24.1%, 75% of which, however, was to due to differentiation between the two varieties.

Results for angiosperm trees show a similar range of gene diversity among populations. *Populus trichocarpa Torr. & Gray* showed a G_{st} value of 6.3% in the Pacific Northwest (WEBER and STETTLER, 1980). In an animal pollinated, widespread Eucalyptus species, MORAN and HOPPER (1987) found an interpopulation differentiation of 10% to 12%. In two species of tropical acacias in Australia and New Guinea, the G_{st} values ranged from 0.9% to 18% (MORAN et al., 1989a). In two European oak species, *Quercus robur L.* and *Q. petraea (Matt.) Libel.*, differentiation was between 0.5% and 5% (KREMER et al. 1991). In six American oak species the average G_{st} value was 8.6% (MANOS and FAIRBROTHERS, 1987). In *Fagus sylvatica L.*, the average differentiation calculated over 140 populations was 5.4% (MÜLLER-STARCK, 1991). In chestnut, *Castanea sativa L.*, a differentiation of 12.2% was reported by PIGLIUCCI et al. (1991) and VILLANI et al. (1991).

In summary, based on Nei's G_{st} or Wright's F_{st} coefficient, only limited amounts of allozyme variation exist among populations of most forest tree species. These coefficients of genetic differentiation, however, clearly underestimate the real inter-population differentiation since they weight loci equally and take the mean over all loci. In addition, both G_{st} and F_{st} attain a value of 1 if, and only if, all populations are fixed for different alleles at all loci (GREGORIUS, 1978; GREGORIUS and ROBERDS, 1986). Consequently, the small amounts of reported differentiation might be more the result of using an inadequate measure to describe the differentiation than a lack of differentiation itself. Studies, using the coefficient proposed by GREGORIUS and ROBERDS (1986), clearly demonstrate that subpopulation differentiation can vary widely for different loci studied. In *Liriodendron tulipifera L.*, for example, the coefficients of differentiation by GREGORIUS ranged from 1% for a peroxidase locus to 28.1% for an esterase locus (BROTSCHOL et al., 1986). In ponderosa pine, the coefficient of differentiation ranged from 4.8% for a phosphohexose-isomerase locus to 17% for a fluorescent esterase locus (GREGORIUS and ROBERDS, 1986). A coefficient of differentiation calculated over multiple loci, therefore, has to weight the loci according to their frequencies in the gene pool in order to reflect these differences in a composite index. In several studies, which compare the differentiation index by GREGORIUS with the G_{st} coefficient by NEI, the average differentiation

index over multiple loci was about 3 times higher than indicated by G_{st} (GREGORIUS and ROBERDS, 1986; GREGORIUS et al., 1986; MÜLLER-STARCK and GREGORIUS, 1986; MÜLLER-STARCK, 1987). The method proposed by GREGORIUS and ROBERDS clearly has greater sensitivity in portraying population differentiation.

However, all measures of diversity are composite indices which in one way or another take an average over loci or populations. Although the different coefficients of diversity differ somewhat in their sensitivity, they do not tend to reflect real differentiation among populations, especially regarding the existing differences among the different gene loci. Patterns of variation may differ at individual loci which are not be reflected in a composite index (For other methodological problems regarding allozyme markers see *section 1.4*). The reported results, therefore, may not portray very well small interpopulation differentiation. Consequently, small G_{st} values are not proof against the adaptivness of allozyme markers as is sometimes concluded.

Many studies in forest tree populations were conducted in order to find associations of allozyme variation with environmental variables. There does not appear to be any general relationship between allozymes and ecological variables since many of these studies failed to find such associations, but some have reported significant association between allele frequencies and environmental conditions. Frequently however, there is a confounding of geographic continuity and environmental continuity. Consequently, other processes than natural selection such as genetic drift or past migration history could account for the observed patterns. In some circumstances, however, allozymes appear to be associated with environmental conditions as the following examples should illustrate.

GURIES and LEDIG (1981) reported highly significant correlations of allele frequencies for six out of 11 loci with climatic variables, such as winter temperature and snowfall, that parallel clines in cone serotiny (LEDIG and FRYER, 1972), wood specific gravity (LEDIG et al., 1975) and height growth (LEDIG et al., 1976) in pitch pine, *Pinus rigida Mill.* Significant correlations were found for malate dehydrogenase (*Mdh*), isocitrate dehydrogenase (*Idh*), 6-phosphogluconate dehydrogenase (*6pgd*), leucine aminopeptidase (*Lap*), glutamate-oxaloacetate transaminase (*Got*) and aconitase (*Aco*) loci. Many other authors have proposed adaptation to local environments as an explanation for a change in allele frequencies across large geographic distances, altitudes, changes in slope and soil or climate conditions.

In Norway spruce, strong clinal variation in an acid phosphatase locus (*Acp*) was reported by BERGMANN (1975a, 1978) for a latitudinal transect as well as for two altitudinal transects in the Austrian and Swiss Alps. The observed variation was hypothesized to be a result of selection caused by several temperature variables. The clinal variation in the Austrian Alps paralleled the variation found in seedling traits and a very high correlation ($r=0.99$) was reported between the allele frequencies at this locus and time of budset (HOLZER, 1982). Two more loci, 6-phosphogluconate dehydrogenase (*6pgd*) and glutation reductase (*Grd*), showed clinal variation in the same altitudinal transect (BERGMANN, 1988). With the exception of ACP, these clinal trends could not be verified in another study in the German Alps (RUETZ and BERGMANN, 1989). The association of allele frequencies with elevation found for the *Acp* locus was applied to investigate the degree of autochthony of stands in German mountain forests (RUETZ and BERGMANN, 1988, 1989). Clinal variation in allele frequencies at four loci (*Lap*, *Gdh* (glutamate dehydrogenase), *Got* and *Acp*) associated with altitude was reported for four populations of *Picea abies* in Sweden (LUNDKVIST, 1979). Clinal variation for an esterase (*Est*), an acid phosphatase (*Acp*) and a leucine-aminopeptidase (*Lap*) locus in Norway spruce populations in Scandinavia was also found by BERGMANN (1973, 1975b). Genetic distance between 11 populations was associated with the geographic distances in a north to south direction. LAGERCRANTZ and RYMAN (1990), investigating variation patterns based on data of 70 populations from the natural range of Norway spruce, found a very high correlation of allozyme frequencies with latitude and longitude ($r^2=0.80$); with clinal variation oriented in a southwest-northeast direction. According to the authors, this clinal variation pattern reflects the effects of evolutionary factors, imposed by recent historical events related to the last glaciation, and to a lesser extent adaptation to environmental gradients. STUTZ (1990) found evidence for the adaptive significance of *Lap*, *6pgd* and *Fe* (fluorescent esterase) loci in an investigation of Norway spruce populations in Switzerland. Factor analyses as well as multiple regression analyses revealed that *Lap* was associated with continentality of the climate (precipitation during vegeta-

tion period, July temperature and soil-water stress). Heterozygosity of *Fe* was related to precipitation during the vegetation period and *6pgd* was influenced by soil properties (see later).

A survey over the range of European silver fir, *Abies alba Mill.*, showed overall clinal variation for the *Idh-B* locus, with frequency of the most common allele ranging from 0.80 in the north to about 0.20 in its southernmost distribution in Calabria, Italy (MOLLER, 1986; BERGMANN et al., 1990). This ecogeographical variation pattern paralleled a temperature gradient from north to south. In a later study, clinal variation in *Idh-B* was confirmed using additional population samples. In addition, enzyme kinetic differences in thermostability and catalytic efficiency between the alleles were found (BERGMANN and GREGORIUS, 1993). These different kinetic properties of the allelic forms in combination with the congruence of the frequency distribution with temperature, strongly suggest that temperature is the selective force acting directly on the *Idh-B* locus of silver fir. It is interesting to note in this context that relationships between temperature-dependent kinetic properties and specific temperature adaptations of genotypes were found for the same enzyme system in trout species by MOON and HOCHACHKA (1971, 1972). Similar north-south clinal trends for the locus was verified in a seperate study of *Abies alba*, by KONNERT (1992), for a much smaller area in Bavaria, Germany.

Clinal variation patterns were also reported for black spruce, *Picea mariana*, in Newfoundland by YEH et al. (1986). Relatively strong correlations of allele frequencies were found for *G6pd* (glucose-6-phosphate dehydrogenase) with longitude, for *Pgi* (phosphoglucose isomerase) and *G6pd* with latitude, and for *Aco* (aconitase) and malate dehydrogenase (*Mdh*) with elevation. According to the authors, it is not likely that one would find significant correlations in five out of 13 loci, if random processes alone were involved in generating the patterns. They therefore conclude that the complex patterns are the product of underlying genetic processes such as natural selection and past migration history.

MITTON et al. (1977) reported striking genetic differentiation at a peroxidase allozyme locus (PER) for ponderosa pine on the eastern slopes of the Colorado Front Range. The frequency of the most common allele increased steadily with elevation over a distance of only 3 km. In cool and mesic mountain sites, the frequency was high and the genotypic structure met Hardy-Weinberg expectations. On warm and xeric sites in the plain, however, the frequency was low and an excess of heterozygotes was observed. These results have been confirmed by later studies. In addition, congruent differentiation for the same locus was found between adjacent north- and south-facing slopes. On warmer and dryer south-facing slopes, an excess of heterozygotes and a low frequency of the common allele was found, while on cool, north-facing, slopes the frequency was high, with no deviation from Hardy-Weinberg expectation. Moreover, the adaptive significance of the peroxidase locus was supported by enzyme kinetic studies which showed a difference in the temperature optimum between the allelic forms of the enzyme (MITTON et al., 1977; MITTON et al., 1981; MITTON et al., 1980; LINHART et al., 1979, 1981; BECKMAN and MITTON, 1984). Significant differences in allele frequencies between sites on mesic, north-facing, slopes and xeric, south-facing, slopes were also observed for loci coding colormetric esterase (CE), fluorescent esterase (FE), phosphoglucomutase (PGM) and glutamate dehydrogenase (GDH) by HAMRICK et al. (1989).

Similar results for peroxidase have also been reported for beech, *Fagus sylvatica,* in Europe (THIEBAUT et al., 1982; COMPS et al., 1987). A frequency cline at this locus paralleled the temperature gradient from southern to northern locations. In addition, the same associations were found when temperature changed due to elevation or aspect of sites. Clinal variation associated with temperature was also found for a *Got* locus, with a north-south as well as an altitudinal trend (COMPS et al., 1987; THIEBAUT et al., 1982). GÖMÖRY et al. (1992) reported correlations of the frequencies of several alleles with mean annual temperature and precipitation in beech populations in France: allelic frequencies at *Got*, *Acp*, *Idh* and *Mnr* (menadione reductase) loci were significantly correlated with temperature. Frequencies at *Got*, *Per*, *Acp*, *Idh*, *Mnr* and *Pgm* were also correlated with precipitation. Some of the reported correlations, however, might be redundant because of intercorrelations among the two climate variables.

Highly significant correlations of isozyme band frequencies with geographic variables were described by FRYER (1987) for the natural range of *Pinus rigida.* Some isozyme bands of ACP, PER and CYO (cytochrome oxidase) were associated with altitude, while some of LAP, ACO,

ACP and PER showed clinal variation with latitude or longitude. Forty of the 76 isozyme bands tested had significant correlations with at least one of the 17 climatic variables. Winter daily mean temperature, summer mean monthly minimum temperature, winter mean monthly maximum temperature and winter mean monthly minimum temperature were most strongly associated with the isozyme bands.

In lodgpole pine in British Columbia (*ssp. latifolia*), YEH and LAYTON (1979) reported clinal variation patterns for an aspartate aminotransferase (*Aat*) locus with altitude and a 6-phosphogluconate dehydrogenase (*6gpd*) locus with latitude.

Eight out of 13 loci studied by YEH and O'MALLEY (1980) in Douglas-fir from British Columbia showed significant correlations with geographic variables: the frequencies of the most common alleles at *Mdh* and *Pgm* loci were correlated with latitude; at *Aco*, *Ald* (aldolase), *Mdh* and *Me* (malic enzyme) loci with longitude; and, at *Ald*, *Me*, *Pgi* and *6pgd* loci with altitude. The authors concluded that, for the number of loci chosen at random, such correlations would not predominate if the alleles were neutral.

Substantial differentiation of a peroxidase locus (*Per*) along an elevational gradient was found in samples of different growth forms of Engelmann spruce, *Picea engelmanii (Parry) Engelm.*, and subalpine fir, *Abies lasiocarpa (Hook) Nutt.*, at tree line by GRANT and MITTON (1977). The observed heterogeneity of allele frequencies showed a consistent linear trend for two alleles in both species. The enzymic differentiation, in this case, corresponded to clear morphological differentiation in growth form.

Although other studies have found significant congruence between allozyme loci and morphological characters, there does not appear to be any general relationship between allozymes and quantitative traits. The reason for this might at least partly be due to analytical difficulties in assessing such relationships. Only one study directly compared the zymograms of parent trees with those of their progeny and with growth traits of these progenies (FERET, 1974). Another study by FRYER (1987) compared allozyme data and quantitative traits for the same seedling. In several other studies comparing allozymes and traits, however, the data were assessed using different samples for the allozyme analysis and assessment of quantitative traits (for a further discussion of the analytical problems see *section 1.4*).

FERET (1974) studied three stands of *Picea pungens Engelm.* and progeny derived from the stands for an esterase (*Est*) and a peroxide (*Per*) locus, as well as for some morphological characteristics. Some isozyme and morphological variation was detected among parent trees and individual progenies, but there were no differences whatsoever among stands on the basis of their progeny. The relationship between isozymes and other traits therefore was not conclusive. No correlation analysis to study the association among the two data sets was performed.

Clusters based on morphological traits and based on isozyme bands from the same seedlings of pitch pine were in substantial agreement in the study of FRYER (1987), however, the isozymes used in this study were only phenotypicaly assessed bands from zymograms; *i.e.* the bands were not interpreted in terms of allelic variants.

Good agreement between isozyme data and quantitative traits of eight-week old seedlings from half-sib families in a Douglas-fir population was reported by EL-KASSABY and SZIKLAI (1982). LAGERCRANTZ and RYMAN (1990) reported a very high correlation ($r=0.78$) between morphological and allozyme data for 48 populations of Norway spruce from Europe. Good congruence of juvenile characters and allozyme frequencies was found in a cluster analysis of 25 Eucalyptus populations from Australia (*E. grandis Hill ex Maiden* and *E. saligna Sm.*) by BURGESS and BELL (1983). KONNERT (1991) reported significant correlations between allele frequencies at a *G6pd* locus and several traits, such as phenology and growth, in one-year old seedlings of Norway spruce. VON WÜHLISCH and KRUSCHE (1991) reported significant additive effects at *Mdh* and *6pgd* loci on diameter growth in 21-year old Norway spruce provenance plots. WHEELER and GURIES (1982) found a moderately high correlation between the genetic distance based on 42 allozyme loci and the morphological distance based on 12 cone and seed traits in *Pinus contorta.* Cluster analysis produced the same grouping for both data sets.

Many other studies, however, have not found relationships between allelic and morphological data, for example RAJORA et al., (1991), DICKINSON et al., (1988), EL-KASSABY (1982), LINHART et al. (1989), GURIES and LEDIG (1982) and FALKENHAGEN (1985).

Associations of enzyme polymorphism with soil conditions, suggestive of an adaptive significance of certain enzymes, were found in several investigations. SHEA (1985) described local differentiation patterns in Engelmann spruce, *Picea engelmanii,* which were partly induced by differences in soil-water availability. Populations on wet sites showed a significant deficiency of heterozygous individuals at a *Pgm* locus, while populations on dry sites had an excess of heterozygous. This result was later confirmed by investigations in neighboring populations carried out by STUTZ and MITTON (1988) and MITTON et al. (1989). In Norway spruce, STUTZ (1990) found an association of a *6pgd* locus with the parent rock formation: the allele frequency changes at this locus paralleled a gradient from granite to limestone rock formations. Adaptation to ultramafic soils is also postulated as a plausible explanation for observed allele frequency differences in jeffrey pine, *Pinus jeffreyi Grev. and Balf.*, populations in Southwest Oregon (FURNIER and ADAMS, 1986). BROTSCHOL et al. (1986) found indications of an adaptation to soil conditions which was expressed in allozyme differentiation among stands of *Liriodendron tulipifera L.* In a recent study on jack pine, *Pinus banksiana Lamb.*, XIE and KNOWLES (1992) found significant associations of two allozyme loci with soil nutrients. An aldolase (*Ald*) allozyme variant appeared more often on soils with high concentrations of potassium and low concentrations of silicon and titanium, while an allelic form at the *Acp* locus was positively associated with the concentration of heavy metal ions in the soil solution.

The adaptive significance and potential of enzymes in forest tree species have also been studied under conditions of stress such as air pollution. Several of these studies have found relationships between allozyme genotypes and stress tolerance. SCHOLZ and BERGMANN (1984) were able to show differences in allele frequencies and genotypic structure between tolerant and sensitive clones of Norway spruce after fumigation with sulfur dioxide. Particularly striking were the results found for a *G6pd* locus which seemed to be directly involved in the selection process induced by the pollutant. Besides *G6pd*, three more loci (*Got*, *Pepca* (phosphoenolpyruvate carboxylase) and *Grd* (glutathion reductase)) differed between the two groups. Generally higher diversity and heterozygosity were observed in the tolerant group compared with the sensitive group (BERGMANN and SCHOLZ, 1987). Strong indications of selection by air pollutants were also found by MEJNARTOWICZ (1983), who studied the allele frequency variation at an *Acp* locus among more or less tolerant trees of Scots pine in the vicinity of a factory emitting SO_2 and fluorides. KONNERT (1992) observed associations of allele frequencies at *Got* and *6gpd* loci with the degree of needle loss on damaged silver-fir trees in Germany. In addition, the degree of needle loss was negatively correlated with multilocus gene diversity. Tolerant and sensitive groups were also significantly different in their genotypic frequencies for two *Idh* loci and a *Got* locus, however, no difference in genetic diversity and heterozygosity was observed for these loci. MÜLLER-STARCK (1989) and MÜLLER-STARCK and ZIEHE (1991) reported a higher heterozygosity over 17 loci in a tolerant group of adult beech, compared with a sensitive group of trees, under the influence of air pollution. Differences in genotypic frequencies at a *Grd* locus were also found by RUETZ and BERGMANN (1989) who compared damaged and undamaged groups of adult trees of Norway spruce in a mountain forest affected by air pollution. Differences in allele frequencies among tolerant and sensitive groups of trees of Norway spruce were reported by LÖCHELT (1994) for the four loci *Got*, *6pgd*, *Pgi* and *Nadh* (nicotinamid-adenin-dinucleotid dehydrogenase). In all 24 sampled populations, frequency differences among the two groups of trees were observed but no consistent trend was apparent over all the populations.

MÜLLER-STARCK and HATTEMER (1989) found indirect evidence of directional selection (viability selection) for a *Lap* locus, comparing seed and adult trees of the same beech stands in Germany. In addition, average heterozygosity over 17 loci increased under the natural conditions of stress. KIM (1980) observed that seedlings, homozygous at a *Lap* locus, had higher viability under homogenous glass house conditions whereas heterozygous seedlings revealed higher viability under the heterogeneous conditions in the forest. Indirect evidence for viability selection on a *6gpd* locus was also reported by KONNERT (1991) for Norway spruce.

Several studies report a positive relationship between the number of heterozygous allozyme loci and growth rate. Two central hypotheses may be stated which explain such relationships, dominance or overdominance. While the dominance hypothesis suggests that heterozygote superiority is a consequence of the less frequent occurrence of deleterious recessive alleles as homozygotes in the more heterozygous genotypes, overdominance suggests that having multiple alleles is advantagous, causing higher fitness. Under the dominance hypothesis, relationships between the number of heterozygous allozyme loci and fitness traits are expected to be negative because each heterozygous locus causes a slight reduction of fitness compared with homozygotes for the common alleles because the rare alleles occur predominantely in the heterozygous condition. Association of heterozygosity with fitness traits may then be viewed as a mere reflection of the level of inbreeding over the entire genome. Under the overdominance hypothesis, however, associations between allozymes and fitness traits are expected to be positive; polymorphisms may thus be interpreted as an indirect evidence for the adaptivness of certain allozymes and as a result of balancing selection.

MITTON and GRANT (1980) found a significant positive association of heterozygosity and growth, as estimated by the average width of the annual rings, in quaking aspen, *Populus tremoloides Michx.*, in Colorado. Protein heterozygosity was related to the variability of growth rate in Ponderosa pine and in lodgpole pine, but in neither of the species was mean heterozygosity related to growth rate of mature trees (KNOWLES and MITTON, 1980; KNOWLES and GRANT, 1981; MITTON, 1983; MITTON et al., 1981). In Ponderosa pine, highly heterozygous individuals exhibited higher growth variability than predominantly homozygous individuals, while in lodgpole pine the opposite occurred. These contradictory results may be due to the indirect influence of fertility differences related to heterozygosity and growth rate. A negative relationship between growth rate and relative levels of female cone production was found for Ponderosa pine, while there was no such relationship for lodgpole pine. Ponderosa pine is known for their sporadic reproductive output during mast years, which leads to reduced radial growth in years with reproduction. Lodgpole pine, on the other hand, produces cones regularly every year which allows regular annual energy allocation to radial growth each year (LINHART et al., 1979; LINHART and MITTON, 1985). In Engelmann spruce, taller three-year-old seedlings had consistently higher levels of heterozygosity than the normal control seedlings (MITTON and JEFFERS, 1989). In pitch pine, results on relationships between heterozygosity and growth rate were inconclusive, since in 5 stands the correlation was positive and in three stands it was negative. In addition, the relationships were highly dependent upon the age and the density of the stands. Correlations between heterozygosity and growth increased with stand age, density, and environmental unpredictability, observations showing that relationships are complex, but also suggesting that heterozygotes were better able to withstand greater environmental variability over time (LEDIG et al., 1983). In 18 year old beech populations, THIEBAUT et al. (1992) observed significantly more heterozygote individuals in the group of tall individuals than in the group of small ones. In jack pine, the relationships between heterozygosity and quantitative traits were highly non-linear and the degree of association changed under different environments; maximal expression of heterosis for morphological traits was observed under stressful conditions (GOVINDARAJU and DANCIK, 1987). MORAN et al. (1989b) found a positive correlation between heterozygosity and height at the population level in *Casuarina cunninghamiana Miq.* In knobcone pine, *Pinus attenuata Lemm.*, heterozygosity was positively correlated with trunk growth, but negatively with cone production (STRAUSS, 1986). Eighty percent of the founder trees in Norway spruce cohorts at tree line, which produced and supported vegetative propagules over some period of time, were heterozygous for a *6pgd* locus, whereas all other trees without vegetative propagation were homozygous for this locus (STIMM and BERGMANN, 1994). The authors speculated that being heterozygous at this locus means an advantage in energy production or dry matter allocation which is necessary to support the vegetative propagules for a period of time. In radiata pine, STRAUSS and LIBBY (1987) found a highly significant positive linear correlation between clonal heterozygosity and growth. A significant negative quadratic correlation, however, was indicating that the effect of heterozygosity diminished as the level of heterozygosity increased. The increasingly negative relationship of heterozygosity to growth is explained by negative effects of deleterious alleles, since each heterozygous locus causes a slight depression in growth compared with homozygotes for the common alleles and since the rare alleles occur predominantly in heterozygous condition. The contributions of individual loci to heterosis were highly variable. The results were interpreted by the authors as evidence against overdominance and in favor of the neutrality of the examined

alleles. A different conclusion is reached by BUSH and SMOUSE (1992) in a study on 19 year old loblolly pine, *Pinus taeda L.* Survival, growth and fecundity were decidedly non-random with respect to allozyme genotypes. At a *Pgi* locus they found overdominance, with respect to height growth and fecundity, of a modestly rare (frequency of 0.18) allelic variant in the heterozygous form with the most common allele. Although this rarer allele suffered severe survival disadvantages in homozygous form and a mild disadvantage even in heterozygous form, the net effect over the whole life cycle is shown to be a mild overdominance which leads to balancing selection of the two allelic forms. In agreement with other investigations, BUSH and SMOUSE (1991) found severe survival selection against low-frequency alleles, both in heterozygous and homozygous form. The authors conclude that while polymorphic alleles are obvious candidates for balancing selection due to overdominance, rare alleles are not. Negative correlations of growth with heterozygosity for highly polymorphic loci with rare alleles are explained by deleterious effects of such rare allelic variants while positive correlations between heterozygosity and fitness are hypothesized to be the result of overdominance involving more frequent alleles or other genes which are tightly linked to these alleles.

Examples of associations of heterozygosity and other fitness traits, such as higher tolerance to air pollution or higher fertility, have already been discussed. In summary there is some evidence that higher fitness, higher stability and higher plasticity under stressful conditions and heterogeneous environments may be related to heterozygosity of certain allozymes in forest tree species. Although many questions, especially about the mechanism causing this heterotic effect, are still open to discussion, this evidence is in good agreement with the results of other, non forest tree species presented in *section 1.2.* In addition, the results showing associations of allozymes with ecological variation are suggestive of the adaptivness of certain enzyme systems. Although other factors might be involved in shaping the observed patterns, the fairly congruent results, regarding the enzyme systems which show such environmental associations, over many species and study areas suggest that they may be adaptive.

1.4 Potential and limits of isozyme markers

Isozymes behave as genetic traits in a strict sense. Polypeptides are the primary products of coding regions. Since the nucleotide sequence of the polypeptide is collinear to the sequence of the gene, a change in the coding region of the gene may directly lead to a change in the amino acid sequence of the polypeptide. Since several different combinations of nucleotide triplets in the DNA code for the same amino acid, not all mutations produce a change in the amino acid sequence of the polypeptide. If such a change does occur, however, it may alter the net charge or the structure of the enzyme molecule (peptide), thus it can be detected by electrophoresis. Consequently, isozymes are markers which are as close as possible to the DNA level in the sense that differences in the markers are direct reflections of differences in the coding genes. Because isozymes are already transcribed products, they do have properties that are different from DNA markers. Electrophoretic detection, for example, is restricted to functional proteins; structural proteins as well as water-insoluble or cell structure-bound proteins cannot be utilized. Consequently, isozyme markers are less abundant than DNA markers. Isozyme loci represent only a small and non-random sample of all structural genes present in the genome. POWELL (1975) demonstrated that about 0.5% of the eukaryotic genome codes for all the proteins in an organism. In addition only about 25% to 30% of substitutions in the DNA give rise to amino acid substitutions which in turn lead to a change in the mobility (charge) of the enzyme molecules that can be detected by standard electrophoretic procedures. For these reasons only about 0.1% of all the nucleotide substitutions in the total genome can theoretically be assessed with electrophoresis. However, electrophoretic procedures are available only for a fraction of the potential proteins. In other words, only a very small fraction of the possible variation in the genome can be detected with isozyme markers. This lack of representativeness may bias the estimates of genetic diversity and heterozygosity. In addition, mutation rates and thus the variability of the different enzyme systems differ according to their function in metabolic pathways. GILLEPSIE and LANGLEY (1974) subdivided enzymes in two groups: group I and group II enzymes. Group I enzymes are characterized by a single physiological substrate. They are involved in primary metabolism. Group II enzymes are enzymes which use several different substrates; they are involved in secondary metabolism and their functions are often unknown. Group II enzymes are generally more variable than the group I enzymes, since the group I enzymes are more vital (GILLEPSIE and LANGLEY, 1974; JOHNSON, 1974). Con-

sequently, an overestimation of genetic diversity is possible whenever the enzyme groups are not surveyed equally. Differences in diversity measures reported by different authors for the same species may result from different numbers and types of enzyme loci surveyed in different studies.

The genetic control and Mendelian inheritance of isozymes can be verified easily, especially in conifers using haploid megagametopyte tissue. In many of the cases one structural gene codes for one enzyme, thus allowing for a simple genetic interpretation. There may, however, be more than one enzyme in the same enzyme system. In this case, isozyme band patterns have to be verified for their Mendelian inheritance before they can be used for genetic studies. Allozymes, e.g. isozymes which are encoded by different alleles of the same structural gene, are usually codominantly expressed, thus homozygous and heterozygous individuals are easily distinguished from each other. In general, isozymes show complete penetrance, without pleiotropic or epistatic effects.

Quantitative traits, such as morphological, physiological or chemical characters, on the other hand, are not genetic traits in a strict sense (with the exception of simply inherited morphological markers), since they are under obscure, mostly polygenic control and strongly influenced by the environment. Consequently, the relationships between their phenotypic and underlying genetic control remain unresolved. Quantitative traits are the expressed sums of many different and interacting genes in a specific environment. LEWONTIN (1984) has demonstrated that these fundamental differences between allozymes and quantitative traits do have implications for the statistical power of detecting differences. Based on theoretical grounds, LEWONTIN showed that significant differences among the frequencies of individual loci affecting a quantitative trait are much more difficult to detect statistically than differences in the trait itself. Components of fitness or physiological traits might have heritabilities as low as 1%. There might be more than a hundred loci relevant to the character. Hence, on the level of the markers, the character is broken down into more than a hundred genes, with each gene contributing only very little to the expression of the character. In a random sample of allozyme markers, however, it is very unlikely that all, or even a small number, of these genes would be included in the sample. Consequently, the detection of any relationship among allozyme markers and phenotypic traits is very unlikely. The many failures to find such associations are therefore primarily a consequence of these fundamental differences between the two data sets. Allozyme surveys will generally contain an "average" sample of structural genes affecting many different characters of different heritabilities. In practice, the comparisons of gene frequency differences with metric trait differences is a comparison of the "average gene" with the "average trait". As a consequence of these fundamentally different properties of allozymes and quantitative traits, results of studies based on quantitative traits are very likely to differ from studies based on allozyme markers, especially the degree of differentiation uncovered by each data set, and the strength of association with environmental variables. Even if an allozyme locus influences an adaptive quantitative trait, adaptive variation at the individual locus may be nearly impossible to detect. This does however not mean that an individual allozyme locus could not have a large effect on adaptation. Different results therefore are not proof that allozymes are not adaptively significant. Moreover, differing results do not necessarily indicate that evolutionary forces, acting upon the different sets of characters, differ from each other (see for example MITTON, 1983, page 462; MUONA, 1990, page 286).

Many studies, comparing geographic variation of allozymes with variation of phenotypic traits, do indeed find that quantitative traits exhibit a much larger degree of differentiation than protein markers. For example, MERKLE and ADAMS (1986), studying allele-frequency patterns at 27 loci in each of 22 breeding zones for Douglas-fir in Southwest Oregon, found that less than 1% of the allozyme variation was attributable to differences among the breeding zones. This result was in striking contrast with observed variation in seedling quantitative traits. Of the total variation among family means for six traits of one-year old seedlings, an average of 59% was associated with differences among breeding zones and the remainder with differences among families within zones (LOOPSTRA, 1984). Similar strong differentiation in quantitative traits among populations of Douglas-fir in Southwest Oregon have also been reported by other authors (HERMANN and LAVENDER, 1986; CAMPBELL and SORENSEN, 1978; WHITE, 1981; SORENSEN, 1983; CAMPBELL, 1986). In *Pinus contorta,* WHEELER and GURIES (1982) found that only 9% of allozyme variation was distributed among populations while 57% of morpho-

logical variation was caused by interpopulational differentiation. In Scots pine, *Pinus sylvestris,* several allozyme studies in Scandinavia have shown little latitudinal differentiation in allelic frequencies (CHUNG, 1981; GULLBERG et al., 1985; MUONA and SZMIDT, 1985). In contrast to these results, common garden studies revealed a very high correlation (r=-0.97) between budset and latitude (MIKOLA, 1982). Transfer experiments paralleled these findings: ERIKSON et al. (1980) found that a transfer of seedlings even over a very short range to the north or to an altitude higher than the source location results in increased mortality.

The special properties of allozyme markers must be considered in the choice of analytical methods used to describe patterns of variation, differentiation or associations with the environment. If populations differ greatly for a small number of marker alleles, differentiation is readily detectable with only these markers. On the other hand, if populations differ only slightly in their allele frequencies for each of a large number of loci, any differentiation will be detected only if all the loci are considered jointly in the analysis. Especially in forest tree species, where most of the loci share the same alleles in approximately the same frequencies over large geographic areas, univariate techniques of analysis may therefore not be sensitive enough to reveal the existing patterns of variation. Measures of genetic distance or similarity may obscure the underlying patterns because they weight all loci equally. Under these circumstances, single polymorphic loci are highly unlikely to reveal patterns or associations of any kind.

Human geneticists have faced the same problem in studying genetic differences among human races. In terms of allozyme frequencies, human races are remarkably similar at a large number of loci. Differentiation at single polymorphic loci between individuals from different races was not much larger than differentiation between individuals from the same local populations (LEWONTIN, 1972; CAVALLI-SFORZA, 1974; LATTER, 1980). This intriguing result prompted the application of multivariate statistical methods. Multivariate analyses of the same data sets were able to reveal significant patterns of differentiation between Europeans and South American Indians (MENOZZI et al., 1978; SMOUSE et al. 1982; SMOUSE and NEEL, 1977; WARTENBERG, 1985; BARBUJANI and MILANI, 1986). Multivariate techniques make use of the fact, that a large number of small differences is equivalent to a small number of large differences by treating all the loci jointly and equally in the same analysis. Cumulative effects of small differences in allele frequencies, summed across many loci, may result in detectable differentiation between populations at the multilocus level. Whereas the conventional univariate analyses are restricted to studying variation at individual loci or their average over loci, multivariate techniques make it possible to examine variation in multilocus sets, which include correlations between alleles or interactions among loci. Genes do not act independently from others in the genome. Coadapted gene complexes might only be detected when the alleles are analyzed simultaneously. Multivariate techniques seem to be able to overcome parts of the difficulties inherent in allozyme data. Compared to univariate methods, the chance of detecting significant differences or associations is largely improved by multivariate procedures.

Only few studies have used multivariate techniques to analyze allozyme data in forest tree species. GURIES (1984) reanalyzed published allozyme data of four different species by means of principal component analysis. In western white pine, *Pinus monticola Dougl.*, the first principal component accounted for 37% of the variation in allozyme frequencies and was strongly correlated with both latitude and elevation of populations, while the first three principal components cumulatively accounted for 66% of the variation in allele frequencies. Western white pine populations from southern Oregon appeared to form a partially differentiated cluster while no differentiation among northern populations was observable. In pitch pine, *Pinus rigida*, 70% of the variation in allele frequencies were accounted for by the first three principal components. The first two components revealed some distinct grouping; the patterns however did not follow clear environmental gradients. Stochastic processes such as genetic drift and "founder-effects" are believed to be the driving forces involved in producing these patterns. In lodgpole pine from British Columbia, 34% of the variation in allele frequencies was accounted for by the first two principal components. Correlations with latitude, longitude or elevation were only moderate. Nevertheless, ordination of the two components separated populations north of lat 53 ° N from those to the south. The patterning of variation is interpreted as a consequence of past migration history with populations from two different refugia meeting in central British Columbia. In eastern cottonwood, *Populus deltoides Bartr. ex. Marsh.*, the first three principal components accounted for 67% of the variation in allozyme frequencies. Ordination of the first two compo-

nents revealed several clusters of populations which were congruent with the major river systems. Very strong correlations between the first component and latitude (r=0.87), longitude (r=0.74) and several climatic variables were found. Patterning of variation is interpreted as partly due to selection, partly due to migration along riparian habitats.

YEH et al. (1985) used discriminant analysis to study differentiation among 17 populations of lodgpole pine (*ssp. latifolia*) in the Yukon and British Columbia. Two significant canonical discriminant functions accounted for 38% of the total variance in the 20 polymorphic loci. The first function paralleled a strong north to south geographic pattern. The second function had no apparent simple geographic interpretation, but separated populations according to their elevations. In spite of the fact that most of the variation at a single locus could be found within any population in lodgpole pine (WEEHLER and GURIES, 1982; DANCIK and YEH, 1983), the multivariate analysis uncovered a rich structure of genetic variation which was associated with geography and climate conditions.

O'REILLY et al. (1985) used discriminant analysis to study differentiation of upland and lowland black spruce stands in northern Ontario. The discriminant functions based on allozyme variation were able to correctly classify about 70% of the sampled trees into their respective upland and lowland origins. In comparison, only 52% of the one-year old progenies from the same populations were correctly classified into the upland/lowland groups by discriminant function based on morphological seedling traits.

With conventional single-locus techniques, MERKLE and ADAMS (1987) found only about 1% of genotypic variation attributable to differences in breeding zones of Douglas-fir in southwest Oregon. Discriminant analysis for the same region produced 4 significant canonical functions. The breeding-zone contribution to explaining the variation in the scores on these functions ranged from 22% in the first to 12% in the fourth canonical function. However, the first two discriminant functions represented only 25.6% of the variation in the allozyme data set. Thus, only 3.3% of the total genetic variation could be assigned to breeding zones. Although it was possible, by means of canonical discriminant analysis, to assign a larger proportion of allozyme variation to differences among breeding zones, the patterns of variation revealed only moderate associations with geographic variables. Moreover, the magnitude of differentiation was in striking contrast to results from common garden studies which, using the same families, showed strong clines over the same environmental gradients.

CONKLE and WESTFALL (1984) and WESTFALL and CONKLE (1992) were able to detect geographic patterns in several forest tree species using a variety of multivariate analysis techniques. Using canonical correlation analysis, the authors were able to account for 49% of the genotypic variation with two canonical variates in a study on ponderosa pine in California's Sierra Nevada. The two canonical variates were associated with latitude, longitude and elevation. Based on the scores of both canonical variates, groups of similar multilocus genotypes were formed, which related to seed zones and elevations. Moreover, the average probability of correctly allocating an individual tree on the basis of its multilocus genotype to one of these groups, using discriminant functions, was 51%, which is high precision compared to the random expectation of 6%. In another survey, WESTFALL and CONKLE (1992) analyzed multilocus patterns in three more species of the mixed conifer zone in the Sierra Nevada of California: White fir (*Abies concolor Lindl.*), sugar pine (*Pinus lambertiana Dougl.*), and Douglas-fir. Geographic patterning was strongest for Douglas-fir and weakest for white fir with 43% versus 14% of variation accounted for by the first three canonical vectors. There was a striking similarity in the trend-surface among all four species, ponderosa pine included. The patterns were all saddle-shaped surfaces with the long axis of the saddle oriented in a northwest-southeast direction. These saddle-shaped forms of the allozyme patterns were also shared with those of growth traits from progeny tests in California (KITZMILLER, 1990), whereby growth decreases along the Sierras to the northwest and increases to the southwest. Eight groups of multilocus contour intervals were formed based on the canonical scores of the first three vectors. The proportion of correctly classified trees into these eight groups by discriminant analysis, using the original multilocus genotypes of each tree, were relatively high (34% to 53% for the different species) compared with the random classification rate of 14%. A separate analysis of coastal Douglas-fir from southwestern Oregon and northern California by WESTFALL and CONKLE (1992) produced no significant canonical vectors. The patterning was very weak and paralleled the results

of MERKLE and ADAMS (1987). In all the allozyme data, the correlation between a single allele and the canonical model only occasionally exceeded 0.25. Although very small proportions of the variation in any one allele is associated with geography (environment), the aggregate multivariate patterns were much stronger. Allozyme markers behave, as was outlined before, indeed in the way expected for quantitative trait loci. Moreover, it is interesting to note, that allele frequencies at several loci are related to the pattern in most of the samples, especially at loci coding enzymes associated with glycolysis and the Krebs cycle, such as ACO, IDH, MDH, 6PGD, PGM, LAP and GOT.

The multilocus allozyme patterns of white fir, mentioned above, were used by MILLAR and WESTFALL (1992) to develop gene conservation strategies for this species. Based on transfer-risk analysis (CAMPBELL, 1986; WESTFALL, 1992) from multilocus allozyme patterns, three major groups were defined. These groups served as starting points for defining genetic resource management units e.g. gene conservation areas in which average transfer-risk was greater than 6% and smaller than 12%. The allozyme patterns were validated by comparing them to results from common garden studies and plantations. Average transfer-risk per 300 m elevation was 18% for allozymes and 19% for the common garden data. This close correspondence between the results of the different data sets confirmed the value of multilocus allozyme mapping and transfer-risk analysis in white fir of the Sierra Nevada populations in California.

In summary, there is evidence that multivariate analysis of allozyme loci is able to reveal greater levels of differentiation, stronger associations with ecological factors and more subtle gradients or patterns of variation than the conventional univariate procedures. Mapping of adaptive multilocus allozyme variation seems to have potential for contributing to seed zone delimitation, breeding zone formation, transfer-risk analyses and development of gene conservation strategies or other objectives related to gene-ecology.

1.5 Objectives

The main objective of this study was inspired by the promising results of recent multivariate allozyme studies. Accordingly, its major goal was to further evaluate the potentials and limits of multivariate statistical analyses applied to allozyme data. Different multivariate procedures were applied to single tree allozyme genotypes in order to reveal and describe patterns of variation, to test for their adaptive significance and to evaluate their relevance to practical forestry. Present patterns of variation may be the result of different evolutionary forces. In order to assess the importance of the most relevant underlying processes, patterns of genetic variation were studied in the context of one central and two alternative hypotheses. The central hypothesis was:

1) ***Geographic variation in multilocus allozyme genotypes is the result of recent adaptation to current environments, caused by natural selection***

The two alternative hypotheses to test for were:

2) *Geographic variation in multilocus allozyme genotypes is the result of a different evolutionary history, caused by migration from different refugia populations*

3) *Geographic variation in multilocus allozyme genotypes is the result of random processes, caused by genetic drift among populations*

Observed single-locus and multilocus patterns of genetic variation as well as associations with geography and climate were interpreted for their congruence with the expected patterns under each of the three hypotheses. The proportion of variation caused by adaptation to current environments (central hypothesis) was primarily tested by relating the allozyme genotypes to climate variables which were considered to be the most important determinants for an eventual adaptive pattern. In order to assess the importance of past migration history and genetic drift for current patterns of variation (alternative hypotheses), total variation was partitioned into the proportions of variation which 1) are shared by the environmental variation 2) are shared by both environmental and geographic variation and 3) are not shared by the environmental variation. Since effects of migration or drift are expected to result in patterns which are unrelated to

environmental variation, they should be reflected in the proportion of variation not shared by the environmental data.

To further evaluate the adaptive significance of multilocus allozyme variation, patterns of allozyme variation were compared to patterns of variation from quantitative traits, assessed in a common garden study. Since seedling quantitative traits are often closely associated with ecological factors of their seed source location, it is commonly agreed that their variation patterns largely reflect adaptation. Moreover, seedling traits are at present extensively used for the purpose of establishing seed-zones and seed-transfer rules, which are applied in practical forestry. The degree of congruence between the patterns of variation in allozymes and quantitative traits was used to elucidate the adaptive significance of allozyme variation and to help in the interpretation of the existing patterns. In the present work, the quantitative traits were assessed from the progeny of the same mother trees used to assess allozyme genotypes. In this respect, this study adds a new dimension to the investigation of genetic variation patterns. If multivariate analysis of allozymes should turn out to be a useful approach for revealing patterns of adaptive variation, the technique might eventually be able to replace common garden procedures or at least to provide additional complementary information.

More specifically, this study describes and interprets patterns of genetic variation in ponderosa pine in two contrasting regions of Oregon, USA, namely the east slope of the Cascade Range, including parts of the high desert (Warner and Ochoco mountains), and the Klamath and Siskiou mountains in Southwest Oregon, from the Cascade Range to the Pacific ocean. Patterns of genetic variation, both for allozyme markers and seedling quantitative traits, are analyzed for both regions separately and partly for the total sampling area combined. Associations with geographic and climatic variables are studied and compared with each other as well as among the two regions. Explicitly, the study attempts to achieve the following primary objectives:

- A description of allozyme variation and differentiation with conventional univariate techniques.
- An examination of spatial structures of single-locus allozymes and their univariate relationships to geography and ecology.
- A comparison of the two regions with respect to their single- and multilocus genotypic structures.
- An assessment of associations between multilocus allozymes, geography and climate conditions.
- An interpretation of the underlying processes which may have generated the present patterns of variation.
- A comparison of the allozyme patterns with multivariate patterns of seedling quantitative traits.
- An evaluation of the adaptive significance of multilocus allozyme variation and an assessment of the potential and utility of the multivariate allozyme approach for gene-ecology and practical forestry.
- An evaluation of the existing seed zones with respect to the observed variation patterns of allozyme and quantitative traits.

2. Ponderosa pine

2.1 Distribution, characteristics and ecology

Ponderosa pine is the most widely distributed pine in the western United States. Sometimes it is even called the most widely distributed and most important pine in North America (HARLOW and HARRAR, 1969; FOWELLS, 1965). In terms of area, ponderosa pine has the single largest range of any western species. In terms of acres where ponderosa pine predominates, it is second only to the Douglas-fir type. Overall, ponderosa pine is the principal species on some 11 million hectares (27 million acres) in the western United States. The species, however, occurs on several million additional acres where it does not constitute the plurality of stocking. Throughout the western United States, ponderosa pine is the fourth most abundant species, accounting for about 1.1 billion m^3 (38 billion cubic feet) of growing stock (VAN HOOSER and KEEGAN, 1987). It extends from latitude 51° N and about longitude 123° W in the Fraser River drainage in British Columbia to approximately latitude 24° N and longitude 105° W in west-central Mexico, and from latitude 42° 30' N and longitude 99° 30' W in north-eastern Nebraska to the Pacific Coast in California (*Figure 1, p. 19*). The elevation distribution ranges from sea level (Tacoma, Washington) to about 2700 m in California, Colorado and Arizona. Within this vast area of distribution, commercial stands are found in British Columbia, Washington, Oregon, Idaho, Utah, Montana, South Dakota, Colorado, Nebraska, New Mexico, Arizona, California and Mexico. The heaviest concentrations of ponderosa pine occur in northern California (2 million ha) and eastern Oregon (1.9 million ha). Arizona and New Mexico also contain substantial acreages, with each exceeding 1.3 million hectares. Ponderosa pine is a major lumber species and the source of some of the highest value softwood lumber. Its economic importance is considerable: 7,000 to 9,000 people are involved in harvesting and hauling of ponderosa pine, while 10,000 to 15,000 workers are employed in sawmills, and 10,000 in the pulp and paper production, depending on this species. All in all, there are about 35,000 jobs in the primary forest product industry depending on ponderosa pine, with an estimated 750 million dollars annually in total salaries (VAN HOOSER and KEEGAN, 1987).

Ponderosa pine was first recorded in the journal of the Lewis and Clark expedition of 1804, but was named and described only in the late 1820's by David Douglas. For many years it was called western yellow pine but the name was changed in 1932 to ponderosa pine. Ponderosa pine is subdivided into three different varieties: Pacific ponderosa pine (*Pinus ponderosa var. ponderosa*), Rocky Mountain ponderosa pine (*Pinus ponderosa var. scopulorum*) and Arizona ponderosa pine (*Pinus ponderosa var. arizonica*). Pacific ponderosa pine is a three-needle pine; but two-needle fascicles can be found at low frequencies of 1% to 3% (HALLER, 1965; WEIDMANN, 1939; READ, 1980). Rocky Mountain ponderosa pine, on the other hand, has a moderate to high proportion (20% to 85%) of 2-needle fascicles, and its needles are much shorter than elsewhere. In addition, it has stiff needles with sunken stomata, while the Pacific race has flexible needles with surface stomatas. Arizona ponderosa pine normally has five needles per bundle. Some authors consider it a distinct species, while others consider it a variety of the ponderosa pine complex.

Ponderosa pine grows to impressive size: stems with diameters as great as 263 cm (103.5 inch at breast height) and total height of 72m (232 feet) have been recorded (AMERICAN FORESTRY ASSOCIATION, 1956), while diameters of 75 to 125 cm and heights of 30 to 40 m are common throughout its range. In Oregon, the tallest tree is 2.7 m in diameter and 49.5 m in height; it is located close to the Deschutes River south of Bend in eastern Oregon (source: Oregon Forestry Education Program). Trees often reach ages of 300 to 600 years. In eastern Oregon a tree of 726 years of age has been recorded by KEEN (1940).

The ecological niche occupied by ponderosa pine is one characterized by high temperatures and low water availability. Pacific Ponderosa pine grows in the drier portions of a maritime climate regime. Extended cover of clouds and rainfall in winter, with prolonged drought and few clouds in summer, is a common weather pattern. At typical locations in six western states and British Columbia, average annual precipitation varies from 250 to 550 mm with 50 to 150 mm during the growing season (May through August). Commonly, July and August precipitation is about 25 mm or less, and in some places there is none at all. In contrast, Rocky Mountain ponderosa pine grows in a more continental climate. In these areas, winter precipitation is normally

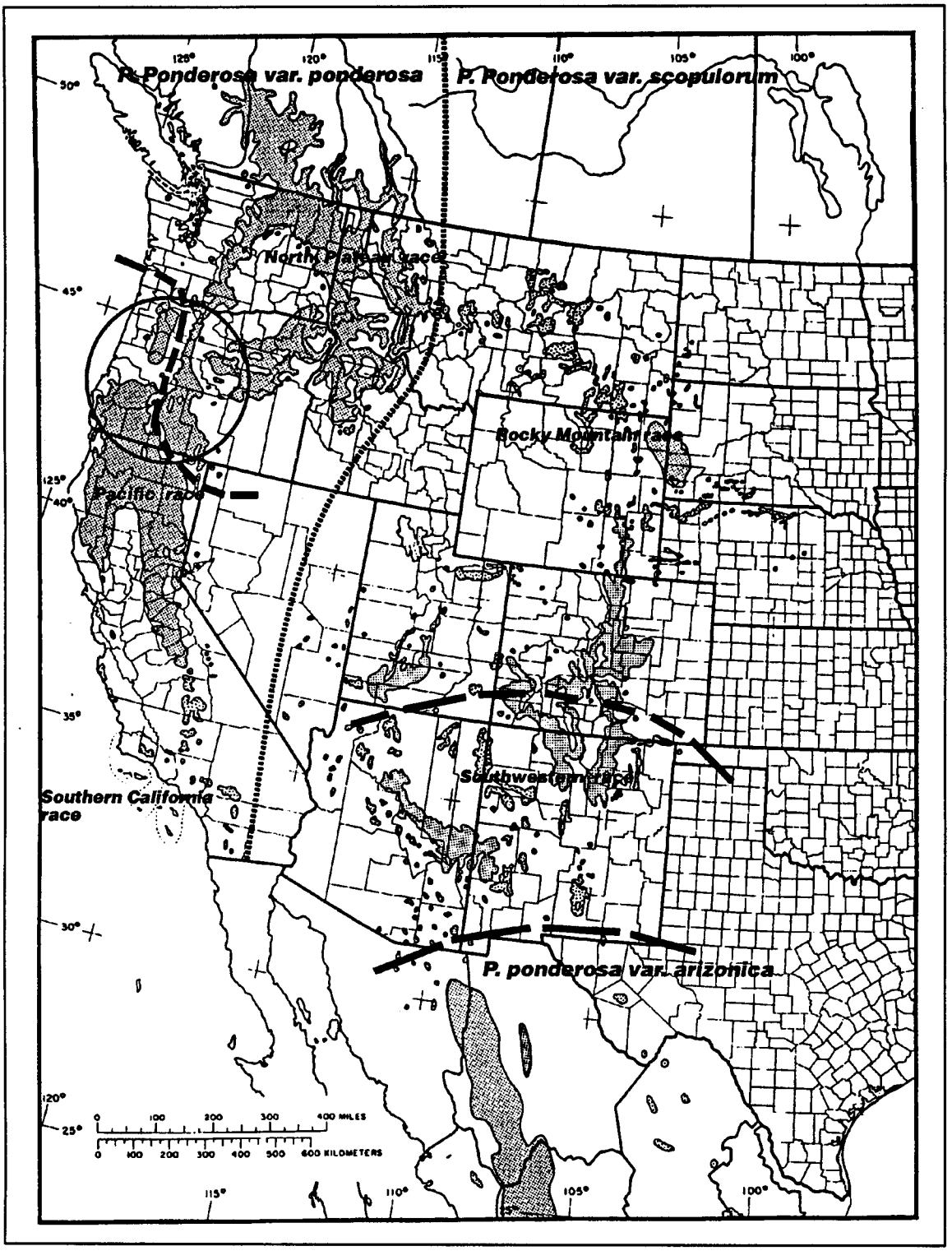


Figure 1: *Distribution of ponderosa pine in the United States, Canada and Mexico (from LITTLE, 1971). Races according to CONKLE and CRITCHFIELD (1988). The circle indicates the area investigated in the present study*

less than in the maritime climate, but storms deliver more moisture in mid-to-late summer (STEELE, 1987). Average annual temperatures are between 5.4°C and 9.9°C, and average July-August temperatures between 16.7°C and 20.9°C. Annual extremes are from -44°C to 43° C, or perhaps even lower and higher. Ponderosa reflects the warm, dry limits of coniferous forests and forms the lower timber-line throughout its geographic distribution. At lower timber-line,

it merges with grassland, shrub land or woodland. The success of ponderosa pine on hot, dry sites is a function of its ability to control water loss and remain photosynthetically active during periods of water stress. Ponderosa pine has the ability to close its stomata at night, minimizing nighttime water loss, to store large quantities of water in both sapwood and heartwood and to locate water in the soil due to its great rooting potential. In addition, temperature optima for photosynthesis are high (25°C to 35°C) in comparison to other species (BASSMAN, 1987).

Ponderosa pine occupies a climax role toward the severe limits of its environmental range and becomes increasingly seral with increasingly favorable environment. On more favorable sites, ponderosa encounters severe competition from other species and may become established only when disturbance reduces competition. At its northern limits, ponderosa pine is outcompeted on cool, moist sites by Douglas-fir (*Pseudotsuga menziesii (Mirb.) Franco*), grand fir (*Abies grandis (Dougl. ex D. Don) Lindl.*), western red cedar (*Thuja plicata Donn ex D. Donn*), or occasionally by subalpine fir (*Abies lasiocarpa (Hook) Nutt.*) (MCLEAN, 1970; PFISTER et al., 1977; COOPER et al., 1985). Farther south, the pine's cool, moist limits occur on sites where the climax is Douglas-fir, white fir (*Abies concolor (Gord. & Glend.) Lindl. ex. Hildebr.*), blue spruce (*Picea pungens Engelm.*), and occasionally subalpine fir (ALEXANDER et al., 1984; YOUNGBLOOD and MAUK, 1985). On severe sites, however, ponderosa pine is climax by default because other conifers are unable to survive. Sites where ponderosa pine is climax typically occur within the lower forest zone, where it grades into non-forest communities. Where it is climax, ponderosa pine has few associated conifers. Junipers are the most common associates, but in Arizona and New Mexico, an occasional Arizona cypress (*Cupressus arizonica Greene*) may appear with ponderosa. In California and Oregon, knobcone pine (*Pinus attenuata Lemm.*), digger pine (*Pinus sabiniana Dougl.*), jeffrey pine (*Pinus jeffreyi Grev. & Balf.*) and Washoe pine (*Pinus washoensis Mason & Stockwell*) may be present in stands where ponderosa pine is climax, while to the north and east, lodgpole pine (*Pinus contorta Dougl. ex. Loud.*) or limber pine (*Pinus flexilis James*) may occur. As a major seral, ponderosa pine occurs on sites with frequent surface fires which kill competing conifers and prepare a seedbed (COOPER, 1960, STEELE et al., 1986).

In summary, ponderosa pine is a species associated with warm and dry habitats, where it occurs as a climax species. Generalizations about its ecology are difficult to apply, however, since it occurs in strikingly different climatic regimes, soil conditions, plant communities and successional roles over a large area in the western United States.

2.2 Races and genetic variability

The large ecological variability present in the area of distribution leads to a strong genetic differentiation within the ponderosa pine complex. Based on morphological and biochemical evidence, CONKLE and CRITCHFIELD (1988) subdivided Pacific ponderosa pine (*P. ponderosa var. ponderosa*) into three major geographic races: North Plateau race, Pacific race and Southern California race. Rocky Mountain ponderosa pine (*P. ponderosa var. scopulorum*) consists of two major races: Rocky Mountain race and Southwestern race (*Figure 1, p. 19*). Provenances from Southern California are the least well known today. At present, their strong differentiation from all other races is based solely on distinct resin patterns.

Pacific race pines have relatively large needles, cones, seeds, and are rapid growing, but lack the thick hypodermal cell layers in needles that contribute to stiffness and cold resistance (WEIDMANN, 1939); they are the least cold resistant in provenance tests outside the Pacific region. In morphological characteristics, the North Plateau race is closely aligned with the Pacific race. Needle hypoderms, however, are thicker and trees of this race are more adapted to cold. Seedling growth starts earlier in the North Plateau race, but growth rate is less than Pacific race seedlings. The Rocky Mountain race is distinguished by smaller cones and seeds, compact foliage, high proportions of 2-needle fascicles, shorter and stiffer needles, and sunken stomata. Seedlings from this race are adapted to the cold, slow growing and growth starts earlier in spring. Compared with the Rocky Mountain race, the Southwestern race has only moderate proportions of 2-needle fascicles (5% to 15% versus 60% to 85%), and has long needles, more open foliage, and a higher growth rate. Compared with Pacific ponderosa pine, the Southwestern race has stiffer needles, deeper stomata and smaller seeds (HALLER, 1965; PFISTER et al., 1977; READ, 1980; WEIDMAN, 1939).

Based on the analysis of xylem monoterpene variation in a large sample of 75 trees at each of 70 locations in the western United States, five broad regional groups with similar amounts of resin components within groups, and significant differences among the groups, were revealed (SMITH, 1977; CONKLE and CRITCHFIELD, 1988). These regional groups, named in the same terminology already used above, were:

1) Southern California group;
2) Pacific group;
3) North Plateau and Rocky Mountain group;
4) Southwestern group, and
5) a group consisting of a broad transition zone between the Rocky Mountain and Southwestern groups.

The patterns found by SMITH were distinct and transitions were sharp between:

1) the Pacific group and the two groups of North Plateau and Rocky Mountains, combined,
2) the Southern California group and the Pacific group,
3) the two groups of Southern California and Pacific combined and the two groups of Rocky Mountains and Southwestern combined.

A broad transition zone, however, separated the combined North Plateau and Rocky Mountain groups from the Southwestern group. The three sharp transition zones correspond to physiographic barriers which prevent gene flow; such as the crest of the Sierra Nevada and Cascade Range, the Transverse Ranges at the narrow southern tip of the Sierra Nevada and the Mohave-Sonoran desert region south of the Great Basin. Distinct differences in monoterpene composition between populations from the east and west sides of the Cascade Range in Oregon and northern California were also reported by STURGEON (1979), the boundary between the chemical regions following the crest of the Cascade Mountains. In addition to barriers preventing gene flow, some races may have come together only recently and may have had little time for gene exchange (CRITCHFIELD, 1984). Arizona ponderosa pine (*P. arizonica Engelm.* or *P. ponderosa var. arizonica*), often considered a variety of ponderosa pine, is quite distinct from the Southwestern race of ponderosa pine and is recognized, based on its monoterpene components, as a unique species (PELOQUIN, 1984).

Analyses of genetic distance between the races, based on allozyme allele frequencies, showed a close genetic relationship between the Pacific and the North Plateau race (Nei's genetic distance was 0.011). The genetic distance between the Rocky Mountain race and the western races was sizable and significant, with a coefficient of 0.082 between the Rocky Mountain race and the Pacific race, and of 0.060 between the Rocky Mountain and the North Plateau races (CONKLE AND CRITCHFIELD, 1988; NIEBLING AND CONKLE, 1990). Six gene loci especially showed significant differences in their frequencies between eastern and western races: alcohol dehydrogenase (*Adh*), isocitrate dehydrogenase (*Idh*), glutamate dehydrogenase (*Gdh*), glutamate-oxaloacetate transaminase (*Got*) and phosphoglucomutase (*Pgm*). The differences revealed from allozyme frequencies, contrasted with the resin patterns, which indicate no differentiation to the west and east of the Continental Divide. Isozyme gene flow must be minimal between western and eastern populations however, because alleles at some of the loci such as *Adh* and *Idh*, have significant frequencies in the two western areas, whereas they are absent or in low frequencies in the Rocky Mountain populations. According to the ranges given by NEI (1974), species are characterized by distances of 0.1 to 1.0, subspecies and varieties by 0.02 to 0.2, and races by 0.01 to 0.05. While such ranges alone cannot be used to determine taxonomic relationships, they provide useful guidelines for interpreting electrophoretic data. The large genetic distance found between the Rocky Mountain and North Plateau and Pacific races are within the range of distances reported for varieties. The observed distances of 0.060 to 0.082 even exceed some inter-species distances reported for other conifers, for example between *Pinus clausa Chapm.* and *P. virginiana Mill.* with a distance of 0.014 (WHEELER et al., 1983) or between *Abies balsamea var. balsamea (L.) Mill.* and *Abies fraseri (Pursh) Poir.*, with a reported distance of 0.060 (JACOBS et al., 1984). The genetic distance between the Pacific and North Plateau races is in the range of races, although allele frequency differences were only low to moderate. Differences between the two races were found for *Idh* (0.84 compared to

0.58), *6pgd* (0.39 compared to 0.31), *Adh* (0.55 compared to 0.59) and *Pgm* (0.76 compared to 0.60).

Racial differentiation between ponderosa pine from the east and west sides of the Sierra Nevada and Cascade Ranges is also supported by results of SMITH (1981) for immature cone color. Congruent with the results found by STURGEON (1979), a sharp separation was reported between the east and the west sides of the crest in southern Oregon and northern California; trees from the west side had light colored immature cones while trees from the east side had dark colored ones.

Several range-wide provenance tests have been conducted with ponderosa pine (WEIDMAN, 1939; SQUILLACE and SILEN, 1962; WELLS, 1964a, 1964b; READ, 1980, 1983; VAN HAVERBEKE, 1986). In general, the results of the different trials agree as to the geographic patterns of variation, e.g. the already described races. KORSTIAN (1924) provided the first explicit comparison of the western (Pacific and North Plateau) and eastern (Rocky Mountain) races of ponderosa pine. He described differences in growth, morphology, anatomy, chemistry, foliage, buds, cones, seed and bark. WEIDMAN (1939) made the first attempts to delineate racial differences in the ponderosa pine complex, by analyzing 20 range-wide sources, mostly from the North Plateau, west of the Continental Divide. Utilizing data from a 25-year old provenance test, he was able to delineate six distinct geographic groups. The west side of the Sierra Nevada and Cascade Range was subdivided into two groups (northern and southern Pacific), the area to the east as far as the Continental Divide into three groups (northern, central and southern Plateau) and finally into a group east of the Divide. SQUILLACE and SILEN (1962) reported 40-year results form Idaho and 30-year results from Oregon and Washington provenance plantations. Their results provided evidence for inherent differences in growth rate among trees of the various regions (36% of total variation). However, they questioned the delineation of WEIDMAN' s racial boundaries. Instead they found clinal east-to-west gradients in growth (64% of total variation) which were closely associated with climate variables such as winter precipitation and April to May temperature. WELLS (1964a, 1964b) reported results of a 2-year old common garden study with 60 origins from a range-wide collection. The variation patterns could be divided into five well-defined ecotypes: California, Northern Plateau and Willamette Valley, Southern Interior, Central Interior and Northern Interior. It is interesting to note that the two populations from the Willamette Valley in Oregon more closely resembled the populations from California than the populations from the east side of the Cascade Range. Although he found seedling traits to be associated with certain climatic factors within and among ecotypes, he concluded that the large differences between the western and the Rocky Mountain varieties are of more ancient origin. READ (1980, 1983) reported results of a common garden study with 3-year old seedlings from 80 origins throughout the range of the species. Based on 13 traits, READ delineated a California and possibly a Pacific group on the west side of the Sierra Nevada and Cascade Ranges and a North Plateau race on the east side. A small transition zone was found between the two varieties. The Rocky Mountain variety was divided into 6 ecotypes: Southern Rockies, Central Rockies, Colorado Plains, East Low Elevation, High Plains and Central Montana.

In summary, variation of traits show ecotypic as well as clinal patterns. Obvious differences exist between the two varieties *ponderosa* and *scopulorum*. Apart from the morphological differences already mentioned, provenances from the coastal variety have a slower germination rate, produce seedlings with more numerous cotyledons, form no mature needles in the first year and produce lamas shoots. In addition, coastal seedlings have a lower degree of branching, fewer and longer primary needles, longer and wider secondary needles and a greater density of stomata (GRANT et al., 1989). The shorter, narrower needles, coupled to a lower stomatal density, and higher frequencies of 2-needle fascicles in the Interior variety strongly support the interpretation, that the scopulorum form has undergone selection induced by aridity of the habitat, and that the morphological characteristics are a consequence of adaptation to warm and dry habitats. This interpretation is supported by physiological differences between the varieties (MONSON and GRANT, 1989): *Ponderosa x ponderosa* crosses showed lower water-use efficiency and higher stomatal conductance than *ponderosa x scopulorum* crosses.

Results from different provenance trials revealed also a north-south differentiation (CONKLE and CRITCHFIELD, 1988). Provenances from the Pacific and the Southwestern race had tall 1-

year seedlings which failed to produce winter terminal buds. Northern sources were shorter, produced winter buds and started growth earlier in spring. Other differences in morphology in north-south direction have already been mentioned above.

Reports on adaptive differentiation within racial ecotypes of ponderosa pine are given by several authors (CALLAHAM and LIDDICOET, 1961; MIROV et al., 1952; ECHOLS and CONKLE, 1971; JENKINSON, 1974; AGER and STETTLER,1983; RHEFELDT, 1984, 1986a, 1986b, 1990a, 1990b). Populations from contrasting environments differ genetically for numerous traits which convey adaptation. Adaptive differentiation, in general, is related to elevation and geographic origin of seed sources. Populations from low elevations, for example, tend to have high growth potential, while populations from areas with short growing seasons (high elevation and inland) are slower growing. On the Colorado Plateau, significant genetic variation was observed for 19 of the 20 traits analyzed (REHFELDT, 1990a). Patterns of genetic variation were pronounced and were related to elevation and geographic location. Seedlings from populations distributed along elevational gradients displayed adaptations to growing season length. When growing in the same common garden environment, populations from low elevations expressed a high growth potential and grew for a longer period of time. Populations adapted to short growing seasons ceased growth early in summer and thus tended to be small. Populations separated by 250 m in altitude tended to be genetically different; 250 m difference in altitude is related to a change of 20 days in the frost free period. REHFELDT also observed two other geographic clines, both with gentle slopes. The first associated traits influencing growth potential with the length of the frost-free period, while the second related growth rhythm and leaf length to the transition between summer droughts that characterize the Rocky Mountains and winter-spring droughts which characterize the Southwest. Similar elevational and geographic gradients were also found for smaller geographic areas in Idaho and Washington (REHFELDT, 1986a, 1986b). Elevational clines were also reported for the Sierra Nevada populations (CALLAHAM and LIDDICOET, 1961; CONKLE, 1973), the Inland Northwest (MADSEN and BLAKE, 1977) and the Rocky Mountains (REHFELDT, 1984).

It is noteworthy that similar genotypes tend to recur across the landscape where similar environments recur. For example, seed transfer models developed by RHEFELDT (1990b) from seedling common garden experiments, indicate that for a given site at 1100 m altitude in northern Idaho, genetically compatible sources are widespread, spanning an elevational distribution of nearly 750 m. To the north of the planting site, compatible populations are found at higher elevations because the environment is milder than at the site itself, whereas to the south of the site, compatible sources must come from elevations lower than the site because the environment is more severe. At the location of the site, however, compatible sources are restricted to an elevational band of only ±175 m (REHFELDT, 1990b).

2.3 Evolutionary history

Fossils of ponderosa pine are rare. In addition, its pollen, needles, seed wings and cone scales are similar to those of other western yellow pines. Tracing the evolutionary history of ponderosa pine therefore is difficult and speculative. No macrofossils of ponderosa pine have been found for the Pleistocene. Ponderosa pine was absent from Southwestern U.S. until after the last glacial period about 10,500 years ago (VAN DEVENDER et al., 1984). During full glaciation, ponderosa pine was probably restricted to only a few refugia. The divergent monoterpene types in the south and the west-east isozyme differences in the north suggest that there may have been two distant progenitor lines, one perhaps derived from a west coast and another from a more central progenitor in Mexico (CONKLE and CRITCHFIELD, 1988). The existence of morphologically distinct races supports this conclusion.

Ponderosa pine has a relatively recent history throughout much of the present range, which was colonized perhaps within the last 6,000 to 8,000 years. The primary massive range extension of ponderosa pine dates back to the Xerothermic period 3,000 to 8,500 years ago. From this relatively short history of immigration in the northern part of its range, it follows that complete adaptation to the variability of habitats may not have occurred. Moreover, some transition zones between different progenitor lines may still represent sharp genetic gradients because of recent contact. The small number of generations after re-colonization of the northern area has probably not permitted extensive gene exchange (CRITCHFIELD, 1984). Therefore, genetic

variation, as observed today, might at least partly be caused by different evolutionary history. Consequently, interpretation of variation patterns must account for the past migration history of the species. The present study will have to address this question in particular since the sampled area comprises two different races, the Pacific race in Southwest Oregon and the North Plateau race on the east side of the Cascade crest. In the study area, patterns of adaptive variation may be confounded with patterns due to a different evolutionary background, especially regional differentiation.

3. Materials and methods

3.1 Mapping genetic variation - general methodology

The study of associations between genetic variation of plants in relation to environments has been a major focus of genecology for several decades (TURESSON, 1923). Classically, this relationship has been examined by seeking correlations between plant type and habitat type. Consistently found correlations between genotypes and habitat conditions suggest adaptation of populations to their environments. Adaptively significant genetic variation can then be mapped in terms of geography and other indexes of environmental variation. Patterns of genetic variation may provide insight into the evolutionary forces responsible for the genetic structure of populations. In particular, selective forces may be understood, the degree of adaptation to ecological factors may be quantified and gradients and direction of changes may be assessed. Patterns may serve as a basis for the selection of optimal seed sources for a given environment. In addition, zones of similar genetic composition may be delineated and may serve to guide seed transfer and to minimize poor adaptation and maximize productivity of future stands. Groups of genotypes genetically suited to the environments of given sites as well as planting sites environmentally suitable for a particular group of genotypes may be predicted from seed transfer models and risks of transferring seed sources estimated (CAMPBELL, 1974, 1979, 1983, 1986, 1991; CAMPBELL et al., 1989; ADAMS and CAMPELL, 1981; REHFELDT, 1990b). From such models, the concepts of adaptability and geographic (environmental) zones are defined continuously rather than discretely. A continuous approach is better suited to describing reality since most genetic variation is distributed in macro- and micro-geographic gradients and not in patches (CAMPBELL and SORENSEN, 1978; CAMPBELL, 1979; GRIFFIN, 1978; GRIFFIN and CHING, 1977; SILEN and MANDEL, 1983; SORENSEN, 1979, 1983; WHITE, 1981). Presumably, such patterns result from natural selection by a complex environment that changes gradually along gradients associated with major topographic and geographic factors (*i.e.* topoclines). From the fact that adaptive variation is usually clinal rather than ecotypic, it follows that mapping must be based on regression procedures rather than on classification models. Genotypic values are estimated for location and environmental variables using regression. The rationale behind this approach is the principle that regression procedures give the most probable value of one variate given values of other variates regardless of any casual relationship. The model describes the difference in average genotypic value associated with a differences in location (environment). If a significant proportion of adaptive variation can be attributed to geographic and environmental variation, such models can be used to predict average adaptive values for each location of a topocline.

Two basic approaches have been used to asses patterns of adaptive genetic variation, with their major difference being sampling design:

1) Long-term provenance trials based on limited samples of populations and test sites, and
2) Seedling common garden studies based on grid samples of single trees from many locations over a sampling area.

The classical provenance test samples several to many trees from populations (provenances) at several to many locations. Progenies of the sample trees are planted in replicated designs at several plantation sites. Traits are measured over a period of time and source variation related to environmental conditions of the source locations. Degree of adaptation to the planting environments can be directly measured. Theoretically, this approach seems ideal since sources are tested under field conditions and in a range of sites. However, mapping complex patterns of genetic variation over complex environmental gradients is hardly possible since the high costs of establishment of field trials severely limit the number of sources as well as test sites. Extrapolation of results to broader areas covering a wide range of environments is often not possible or misleading.

Topographic, environmental and adaptive genetic variation are usually very complex within very small areas (CAMPBELL, 1979). If there is a gradient change in environments in adjacent locations there should also be a gradient change in genotypes, forming a cline (LANGLET, 1936; LEVINS, 1963). The proportion of adapted genotypes therefore changes in a continuous manner with gradual changes in site conditions. Furthermore, genetic variation among trees

within locations along the gradient can be expected to be proportional to the number of environments sampled at each location (LEVINS, 1969). Adaptive genetic variation attributable to parent tree locations (and respective environments) can best be described by a topocline or ecocline. To do this, collections of trees at many locations are needed to sample the several geographic, topographic and environmental variables required to index this variation. Ideally, to fully asses the complex patterns of variation in space, a systematic grid of sample locations is used and trees are collected and characterized genetically at all the grid points. Spacing of grid points depends primarily on the complexity of topographic and ecological variation and on the required resolution of the results. Regression procedures call for a large number of locations with different site variables. The ratio of observations to variables must be high to avoid erroneous conclusions from regression analyses, especially regarding test statistics. In addition, correlations among locations in terms of site variables must be kept to a minimum. On the other hand, sampling several trees per location ensures an adequate representation of within site variation. Phenotypic variation among individuals within sites comes from two sources: variation due to genotypic differences among individuals and variation due to environmental variation within the site. If genetic variation among populations reflects natural selection by an environment that varies across locations (a basic assumption for mapping adaptive variation), it may also be assumed, in absence of definite information, that the surviving genotypes in mature stands are adapted to their site conditions, and that differences among adapted genotypes at a location thus mainly reflect differences in environmental elements within the site which are active in natural selection (CAMPBELL, 1974, 1986). Surviving trees have been screened from many thousands of zygotes by the microenvironments at a location. In other words, at a specific location, genotypes and environments should have the same mean and variance. Consequently, sampling more trees within a site means primarily sampling additional environmental variation with a resolution too fine to be of practical importance. Optimal sampling, therefore, will rather sample few parents at many locations. Repeats at several locations are needed, however, to estimate the variance within sources and to test lack-of-fit in the regression analysis.

The second approach to mapping adaptive genetic variation is based on the rationale mentioned above. The mapping procedure is based on a systematic sampling of trees on a grid over a certain geographic area. Seeds from a single tree at many locations and from an additional tree (repeats) at a subsample of locations are collected and grown in common-garden nursery tests to estimate the genotypic values of the parent trees. The goal is to estimate a large fraction of the genotype in a test over a 2- to 3 - year period. As many traits as possible are therefore measured, including aspects of vigor and growth timing. To measure expression of genotypes as affected by environment, seedlings are grown in two or more contrasting environments which are designed to create extremes in growth rhythm. Such tests provide a genotypic value for many parent trees for several traits measured in two or more contrasting environments. In contrast to field experiments, adaptive variation is indirectly mapped because it is assumed that source variation is a result of natural selection, that the map of source variation is therefore a map of environmental variation, and that adaptation thus can be inferred from the patterns. This "indirect" approach to mapping genetic variation has, however, the advantage that genotypes from a very wide area, and thus from a wide array of environments, can be tested in a short period of time and at a reasonable cost. Furthermore, seedling tests provide important information on adaptive variation at a critical stage of development and at a reasonably fine scale. Nevertheless, the relevance of such patterns to long-term performance is only speculative until verified by long-term testing. With the ability to sample a wide range of environments rapidly and cheaply, seedling studies can be used to sort out those environmental gradients which are relevant to adaptation. Significant adaptive differences may then be evaluated for their practical importance using long-term field trials. Mapping genetic variation patterns using seedling common garden studies is widely used in the western U.S. in order to provide a basis for seed zone delimitation and to establish seed transfer guidelines (ADAMS and CAMPBELL, 1981; CAMPBELL, 1979, 1983, 1986; CAMPBELL et al., 1989; CAMPBELL and SUGANO, 1987; REHFELDT, 1979, 1983a, 1983b, 1983c, 1984, 1986a, 1986b, 1989a, 1989b, 1990a, 1990b). Nursery tests have their limitations, yet they provide an excellent means of mapping important adaptive patterns related to major environmental changes. Results of common garden seedling tests seem to produce results that agree, at least qualitatively, with long term field tests. In an elevational transect of ponderosa pine in California, representing 71 families from seven elevational zones planted in three different altitudes, the adaptational dif-

ferences expressed at 29 years of age, for example, strongly correlated with the clinal trend evident in 2 year seedling data (MIROV et al., 1952; CONKLE, 1973). In the Pacific Northwest, seedling data of Douglas-fir (*Pseudotsuga menziesii (Mirb.) Franco*) indicated higher risks in east-west transfer of provenances at high elevations than for latitudinal moves of comparable distances (CAMPBELL and SORENSEN, 1978). Results from 50 year old provenance tests corroborate these findings. Patterns of variation for Scots pine (*Pinus sylvestris L.*) in Sweden, evaluated from seedling test data, were in agreement with results obtained from seed transfer experiments in the field (LANGLET, 1936; HAGNER, 1970).

According to CAMPBELL (1986), mapping adaptive variation patterns, using a systematic sampling of trees on a grid over a certain area and evaluating seedling genotypic values in common garden tests, is based on the following assumptions:

- the area to be mapped has been sampled sufficiently to determine the true patterns of environmental and genetic variation;
- some adaptive variation is associated with the source location of parent trees and can be separated from other sources of variation;
- seed source variation can be characterized by measuring phenotypic traits in two or more nursery environments;
- seed source variation can be related to regional attributes such as longitude, latitude and elevation and can be mapped in terms of these attributes;
- the environment presumably is the impelling force in natural selection; the map of adaptive genetic variation is therefore also a map of the environmental complex active in natural selection
- a population is better adapted to its place of origin than is any other population.

The evidence for these assumptions has been discussed by CAMPBELL (1986).

3.2 Origin and collection of plant materials

The base materials for this investigation were parent trees and their progenies from two separate common garden seedling studies conducted by the US Forest Service, Pacific Northwest Research Station in Corvallis, Oregon. Each study represented a different region in Oregon: "Southwest Oregon" and "Central Oregon" (*Figure 2, p. 28*), but the objectives of both tests were the same: quantitative traits of 3-year old seedlings would be used to map adaptive genetic variation in order to provide basic knowledge for seed transfer. Following the rationale outlined in *section 3.1*, open-pollinated seed was collected from parent trees, sampled on grids within each region, including many locations with one to two trees per location.

Residual seed stocks from the common garden studies were used in the current investigation to asses allozyme genotypes of the sampled parent trees. Quantitative traits from 1 to 3 year-old progenies of these parent trees, as well as location information of the parent trees, were provided by the Pacific Northwest Research Station.

3.2.1 Central Oregon

The sampled area forms a rectangle of approximately 160 km in the east-west direction and 420 km in the north-south direction, with the western boundary at the crest of the Cascade Range and the northern limit formed by the Oregon-Washington border (*Figure 2, p. 28*). It includes the east slope of the Cascade Range, from the Columbia river to the border of California, the Warner Mountains, the Ochoco Mountains and parts of the high desert. Density of sampling was about one collection site per two townships, with distances between sampled locations of 16 to 32 km. The sampling grid was not exactly regular since collection sites had to be located in natural stands of ponderosa pine along accessible roads. Density of sampling was lighter in the Ochoco Mountains and in scattered stands in the high desert since not all potential grid points were situated within ponderosa pine stands. Moreover, sample locations were chosen to maximize the range of elevation and site conditions. The sampling covers an altitudinal

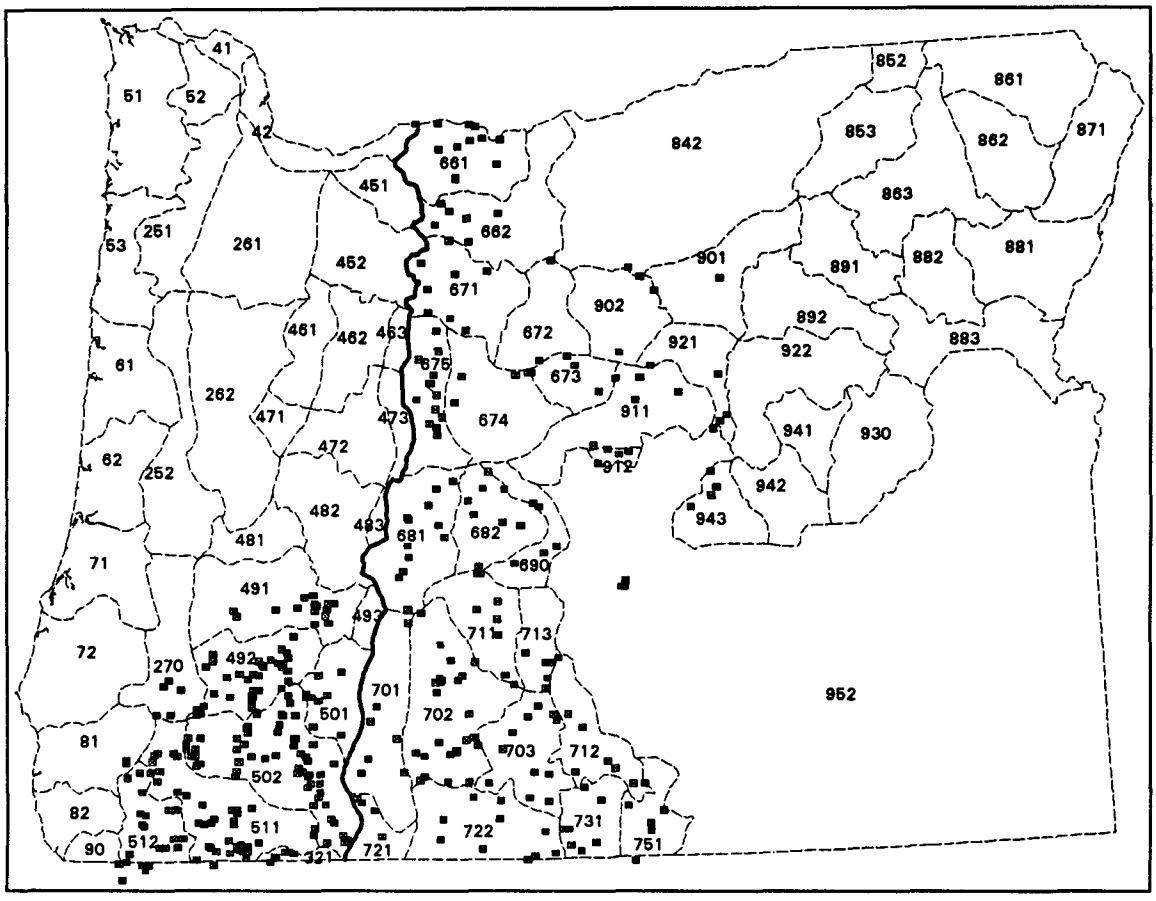


Figure 2: Map of Oregon showing locations (⊠) of sample trees in the two geographical regions investigated in the common garden and allozyme studies. The two regions are referred to as "Southwest Oregon" (west side of the Cascade Range) and "Central Oregon" (east side of the Cascade Range). The Cascade crest, separating the two regions is shown by the black line. The zones with numbers are the official seed-zones used for all species in Oregon

range from 50 m to 2,165 m above sea level, slopes ranging from 0% to 84%, and aspects from 0° to 360°.

At each of 227 sites, one or two parent trees were sampled, for a total of 307 families (80 locations were represented by two families). Two-tree samples were included in order to estimate within-location variation and to be able to perform lack-of-fit tests for models relating genetic variation to geographic and environmental variables. In order to minimize the chance of relatedness among trees, the minimum distance between the two trees at any one location was 50 m with aspect and slope still the same for both trees. Not all 307 families could be used in the analyses. Some of the families had no seed remaining for the allozyme assessment. Other families did not germinate in the common garden study and seedling traits were missing. A few families had missing values for some of the quantitative traits, and again others were eliminated from the analyses due to their peripheral location (*see section 3.5.2, p. 39*).

Seed of about sixty percent of the sample trees were collected in the years 1983, 1984 and 1985. The rest of the samples came from collections of selected trees already in storage at the US Forest Service, Dorina tree improvement center. The years of collection for those trees date back to the late 1970's. The goal of the collection in the 1990's was to obtain 500 filled seeds per sample tree. With an average of 40 to 50 filled seeds (and a variation of 10 - 125) per cone, ten cones was the minimum, 25 the average number of collected cones per parent tree. Cones were normally collected from only the upper portion of the crown; but when only a few cones were available, all cones from the entire crown were harvested.

3.2.2 Southwest Oregon

The area of sampling in Southwest Oregon is a rectangle 155 km north-south and of 145 km east-west, with the eastern boundary at the crest of the Cascade Range (*Figure 2, p. 28*). It includes the Siskiyou and Klamath mountains and the west slopes of the Cascades. Density of collection sites was not uniform over the region because ponderosa pine is not uniformly distributed. Sampling density on average was one collection site per 80 km^2 (31 square miles) of area, which makes roughly about one location per township or an average distance between locations of 8-10 km. Compared to Central Oregon, a higher sampling was needed because Southwest Oregon has rugged topography and extreme environmental heterogeneity (FRANKLIN and DYRNESS, 1973). Collection sites were located in natural stands of ponderosa pine along roads and were chosen so that the elevational and the site variation were maximally represented by the sampling. The sampled locations cover an altitudinal range from 245 m to 1,675 m above sea level, a range of aspects from 0° to 360° and a range of slopes from 0% to 70%. At each of 217 collection sites one or two trees were sampled, for a total of 277 families (60 locations represented two trees). For the same reasons mentioned in *section 3.2.1*, sample sizes varied for the different analyses. The two trees within a location were not separated by more than 800 m; no minimum distance was specified.

Ten percent of the parent-tree samples came from collections of selected trees (a very weak phenotypic selection) and were provided by the US Bureau of Land Management in Roseburg and Medford. Collections from these trees were made between 1974 and 1984. Cones of all other sampled trees were collected over a 3 year period (1984, 1985 and 1988) because of the generally poor ponderosa pine cone crop in this region. Fifteen to 25 cones per parent tree were collected from the upper third of the crown; by shooting for taller trees or by climbing, for smaller trees.

3.3 Assessment of allozyme data

3.3.1 Laboratory procedures

Genotypes of 488 parent trees from both sampling areas were assayed electrophoretically using six megagametophytes per tree. Because the megagametophyte tissue of seeds is haploid in conifers and is genetically identical to the maternal contribution to the seed, diploid maternal genotypes can be inferred directly using haploid seed tissue (TIGERSTEDT, 1973). Assuming a 1:1 segregation of allozymes in megagametophytes from heterozygous trees, the probability of mis-classifying the genotype of any one tree at a particular locus is less than:

$$\left(\frac{1}{2}\right)^{(k-1)}$$

where K is equal to the number of megagametophytes analyzed per tree. Thus, for six seeds the probability of mis-classifying a heterozygote at any single locus as a homozygote is less than 3.13%.

Seeds of each parent tree were soaked in a fungicide solution (2.5 g Captan / liter of water) for 2 days at 4°C in order to prevent development of fungi during germination. Seeds were kept separately by family in Petri dishes during all the treatments. After two days the seeds were stratified by soaking them in a 1% hydrogen peroxide solution (H_2O_2) for an additional 5 days at 4°C in order to improve germination. The hydrogen peroxide was replaced once by fresh solution after 2 days of soaking. After 5 days the hydrogen peroxide solution was poured off, seeds were placed on wet filter paper and the Petri dishes were then put in a seed germinator under a 26/21°C day-night temperature regime and a 12 hour light/dark cycle. Seeds were checked twice a day and moisture was added when needed. Germinated seeds were placed in moist Petri dishes at 4°C in a refrigerator as soon as the radicles had emerged 1 to 5 mm beyond the seed coats. Trees were used for electrophoresis when sufficient germinated seeds from a family were available. Only a few parent trees did not have sufficient germinants within 15 days. Of these trees, some had less than six viable megagametophytes and could not be used in the analysis, others could be genotyped using viable but ungerminated seeds since enzyme activity proved to be sufficient even in ungerminated seeds.

Six seeds per tree were dissected, and the megagametophytes were separated and placed in 96-well microtiter plates. Two drops of cold extraction buffer were added to each sample (*see Appendix I*). Microtiter plates were then placed in an ultra low freezer at -80°C until the next day when samples were thawed for 15 minutes before tissues were macerated mechanically. Grinding was carried out using a metal "replica plate" which fitted the 96-well microtiter plates. Using this replica plate as grinding tool, all 48 samples could be macerated simultaneously within less than 30 seconds. The crude extracts were absorbed onto 14 mm by 3 mm filter paper wicks (Whatman, number 2) and inserted into vertical slices in 12% hydrolized potato starch gels (Sigma Chemical Co., St. Louis, MO, USA). Gels were 21.5 cm by 12.5 cm in size and 8 mm or 12 mm thick, depending on the number of slices (enzymes) to be analyzed. Each gel could accommodate 54 wicks. With six megagametophytes per tree, eight sample trees could be assayed on a gel. Two of the remaining places were used for a red dye and four places for red pine (*Pinus resinosa Ait.*) standards. The dye was used to indicate the advancement of the front during electrophoresis. Red pine samples were used as a standard to facilitate the scoring, especially for alleles with very slight differences in migration rates. Scoring was carried out relative to the red pine standards which were placed at regular distances between the sample trees. Red pine is an ideal standard since it is monomorphic in all enzyme systems (only 1 exception at a locus for one seed was found).

Electrophoresis was conducted following standard procedures as outlined by CONKLE et al. (1982), O'MALLEY et al. (1980) and SOLTIS and SOLTIS (1989). Four gel buffer systems were used to reveal 18 different enzymes (*Table 1*). Details of gel and tray buffer solutions are given in *Appendix I*. The enzyme systems are described in detail in *section 3.3.2.*

Four gels for each of the four buffer systems were run per day, in a refrigerator at 4°C. Trays and gels were refrigerated overnight to cool them down before wicks were inserted in the morning. To keep the temperature sufficiently low during the run, gels were covered with frozen water bags or blue ice packs which were replaced each hour. Gels were run at 75 mA for buffer system A, 70 mA for buffer system B, and 60 mA for buffer systems D and E. Amperage was regularly adjusted, but voltage not allowed to exceed 320 V. Wicks were removed 15 to 20 minutes after the current had first been switched on. Progress of front migration was monitored regularly using the dye markers. Electrophoresis continued until the markers had migrated 8 cm from the origin for buffer systems A and B, and 6 cm for buffer systems D and E. After completion of electrophoresis, gels were sliced, stained and incubated at 37°C. Gels were scored the same day. Gels were then fixed and stored in a coldroom for eventual checking of scoring errors. Stain recipes basically followed CONKLE et al. (1982) with some modifications taken from SOLTIS and SOLTIS (1989), O'MALLEY et al. (1980) or ADAMS et al. (1990). For details on staining procedures *see Appendix I.*

System	Gel buffer [a]	Tray buffer [a]	Enzyme systems [b]
A	Tris citrate pH 8.3	Lithium borate pH 8.3	MNR, PEP, MPI, LAP
B	Tris citrate pH 8.8	Sodium borate pH 8.0	GOT, ACP, G6PD, GDH
D	Morpholine citrate pH 6.1	same as gel buffer	IDH, ACO, SKDH, PGM
E	Morpholine citrate pH 8.1	same as gel buffer	PGI, MDH, ADH, UGP, FUM, FDP, MNR

***Table 1**: Gel- and tray buffers used in electrophoresis*
[a] see Appendix I for details. [b] see Table 2 for explanations of the abreviations

3.3.2 Enzyme systems and loci

A wide variety of enzyme systems were tested in extensive preliminary test runs and stains and laboratory procedures were optimized before the sample trees were assayed. Enzymes which were included in the main investigation had to be polymorphic, had to stain consistently in germinated as well as in ungerminated seeds and had to fit into the practical laboratory procedure. Enzyme systems, which had to be dropped, for example, were G6PD (6-Phosphogluconic dehydrogenase), SRDH (Sorbitol dehydrogenase and ME (Malic enzyme) because their bands stained inconsistently. CAT (Catalse), GPT (Glutamate pyruvate transaminase) and FLEST

(Fluorescent esterase) were tested but later dropped because their staining was too time-consuming to fit into the simultaneous staining and scoring of sixteen gels. The following enzyme systems were included in the main investigation:

Enzyme	Abbreviation	E.C. number	Buffer System [a]	Number of scored loci
Acid phosphatase	ACP	3.1.3.2	B	1
Aconitase	ACO	4.2.1.3	D	1
Alcohol dehydrogenase	ADH	1.1.1.1	E	1
Fructose diphosphatase	FDP	3.1.3.11	E	1
Fumarase	FUM	4.2.1.2	E	1
Glucose-6-phosphate dehydrogenase	G6P	1.1.1.49	B	1
Glutamate dehydrogenase	GDH	1.4.1.3	B	1
Glutamate oxaloacetate transaminase	GOT	2.6.1.1	B	3
Isocitrate dehydrogenase	IDH	1.1.1.42	D	1
Leucine aminopeptidase	LAP	3.4.11.1	A	2
Malate dehydrogenase	MDH	1.1.1.37	E	3
Mannose phosphate isomerase	MPI	5.3.1.8	A	2
Menadione reductase	MNR	1.6.99.2	A/E	2
Peptidase	PEP	3.4.13.1	A	4
Phosphoglucomutase	PGM	2.7.5.1	D	1
Phosphoglucose isomerase	PGI	5.3.1.9	E	1
Shikimate dehydrogenase	SKD	1.1.1.25	D	2
UDP-glucose pyrophosphorylase	UGP	2.7.7.9	E	3
Total	***18***			***31***

Table 2: *Enzymes, abbreviations used in text, enzyme commission reference number (E.C.), buffer systems used for electrophoresis, and number of scored loci used in the present investigation*
[a] *see Table 1 and Appendix I for details on buffer systems*

An enzyme and its band phenotypes were designated by the enzyme's abbreviation in capital letters. When multiple loci (*i.e.* multiple zones of enzyme activity) occurred for an enzyme, the most anodally migrating zone (locus) was designated 1, the next slowest zone of mobility numbered 2, etc. Within each zone, the most frequent allozyme (allele) was assigned the value of 1, and other variants numbered sequentially in order of their appearance; they were also characterized by a mobility value expressed relative to the most frequent allele which was set to 100. In the case of ACO-1 and MNR-1, however, the allozyme originally designated as 1 turned out to be slightly less frequent than another allozyme, since both loci were highly polymorphic with two of the alleles in similar frequencies.

Without exception, inheritance and linkage of all 31 assayed loci were previously verified by a number of studies in different pine species, including ponderosa pine (*e.g.* CONKLE (1981), STRAUSS and CONKLE (1986), O'MALLEY et. al. (1979), CHELIAK et al. (1984), ECKERT et al. (1981), GURIES and LEDIG (1978), MILLAR (1985)). All enzyme systems have been widely used for many studies of pines (FURNIER and ADAMS, 1986; NIEBLING and CONKLE, 1990; STRAUSS, 1986; MILLAR et al., 1988; CONKLE, 1979). Related species often resemble one another in general isozyme characteristics. They can, however, differ in the number of loci for the same enzyme (ADAMS and JOLY, 1980; EL KASSABY, 1981). To aid interpretation of enzyme systems and loci, maps of isozyme phenotypes for ponderosa pine, established by the Pacific Southwest Forest and Range Experiment Station in Berkeley, California (Laboratory of T. Conkle), were available. Unless otherwise indicated, the variants in each zone of activity revealed for an enzyme, are controlled by a separate gene locus.

The following are comments on the zymograms of the 18 enzyme systems assayed in the present analysis:

MNR:

MNR displayed two zones of activity (*Figure 3, p. 32*). The most anodal zone was clearly and

consistently resolved on the A buffer system. The slower zone stained faintly on this system, but clearly on the E system. Therefore, one slice from both the A and E system were stained for MNR in order to facilitate scoring both loci. As suggested by STRAUSS and CONKLE (1986), MNR is identical with Diaphorase. Diaphorase was tested on about fifty trees in the preliminary test runs and the isozyme phenotypes were absolutely identical with those of MNR (*see also SOLTIS and SOLTIS, 1989*).

LAP:

Gels stained for LAP showed three zones of activity with the most anodal being monomorphic (*Figure 3*). A null variant (no band) was found at locus two. A silent variant at this locus was already reported by NIEBLING and CONKLE (1986) for ponderosa pine. Null alleles were also found for LAP in other pine species (GONCHARENKO et al., 1993; MORGANTE et al., 1993).

MPI:

MPI stained rather faintly but consistently in two zones (*Figure 3*). Two zones of activity were also reported by STRAUSS and CONKLE (1986). These authors observed an identical variation for MPI-2 with 6-PGD since both enzymes catalyze sequential steps in metabolic pathways. MPI is an enzyme which is rarely assayed and comparisons to other pine species and studies are not possible. NIEBLING and CONKLE scored only one locus in washoe pine (*Pinus washoensis Mason and Stockwell)* and in ponderosa pine.

PEP:

PEP showed four different zones of activity, all exhibiting variation (*Figure 3*). Four loci were

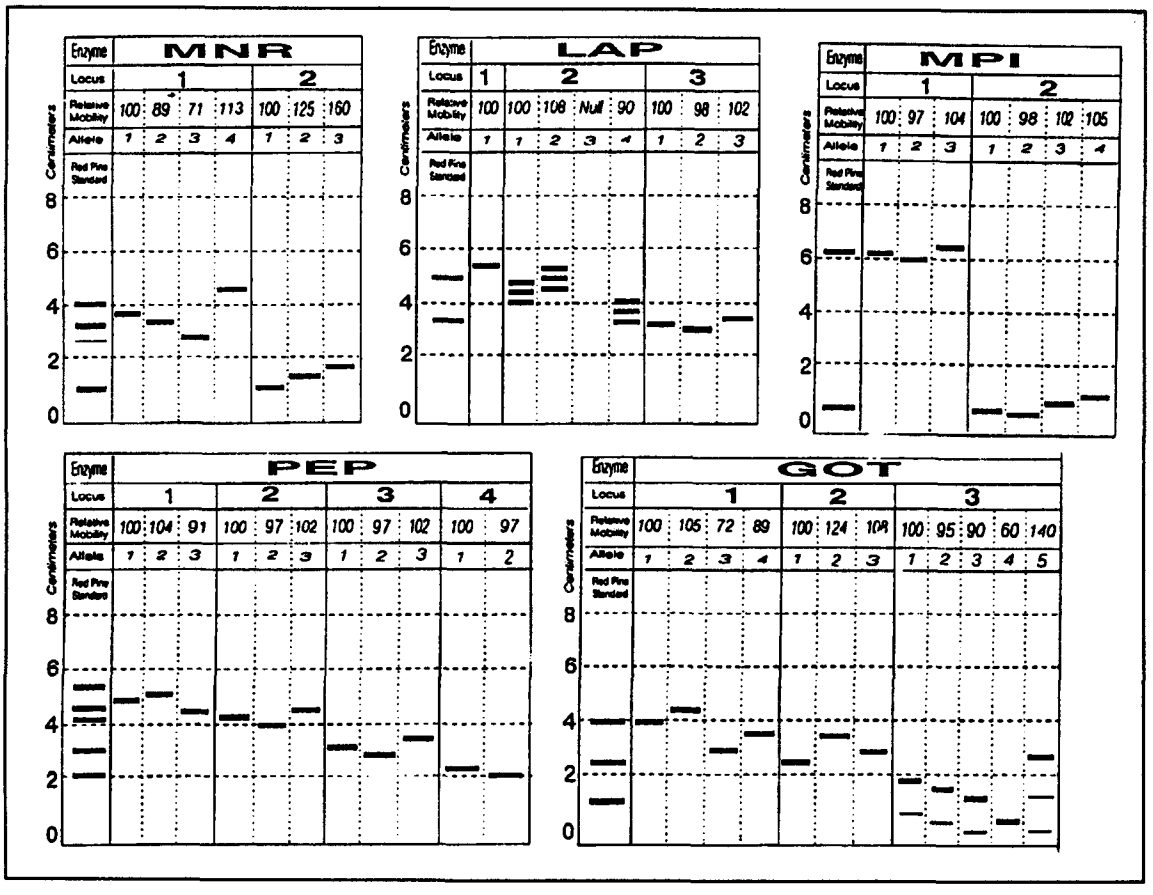


Figure 3: *Enzyme phenotypes of MNR, LAP, MPI, PEP and GOT. Mobility is expressed relative to the most frequent allele (No.1) whose mobility was set to 100. Band patterns of red pine are shown in the far left column.*

also reported for knobcone pine (*Pinus attenuata Lemm.*) by STRAUSS and CONKLE (1986).

GOT:

Three polymorphic loci were clearly resolved for GOT (*Figure 3, p. 32*). Three loci are common for most pine species (NIEBLING and CONKLE, 1986; MORGANTE et al., 1993; STRAUSS and CONKLE, 1986; MILLAR, 1985).

G6PD:

G6P showed two zones of activity with the most anodal zone staining very inconsistently and showing blurry banding patterns. Therefore, variants only at the slower zone were scored (*Figure 4*). Two zones of G6PD activity in pines were also described by other authors (O'MALLEY et al., 1979; STRAUSS and CONKLE, 1986; MILLAR, 1985; NIEBLING and CONKLE, 1990), with the most anodal zone being monomorphic in all cases.

ACP:

Only one zone of activity, showing two banded phenotypes, was observed (*Figure 4*). Only one *Acp* locus was described by MILLAR (1985) in *Pinus muricata D. Don* while CHELIAK et al. (1984) described two for *Pinus banksiana Lamb.*, ADAMS and JOLY (1980) found two for *Pinus taeda L.* and STRAUSS and CONKLE (1986) reported four loci in knobcone pine (*Pinus attenuata Lemm.*). A null variant was found in low frequency (0.025). A null variant for this enzyme was reported for *Pinus strobus L.* by ECKERT et al. (1981).

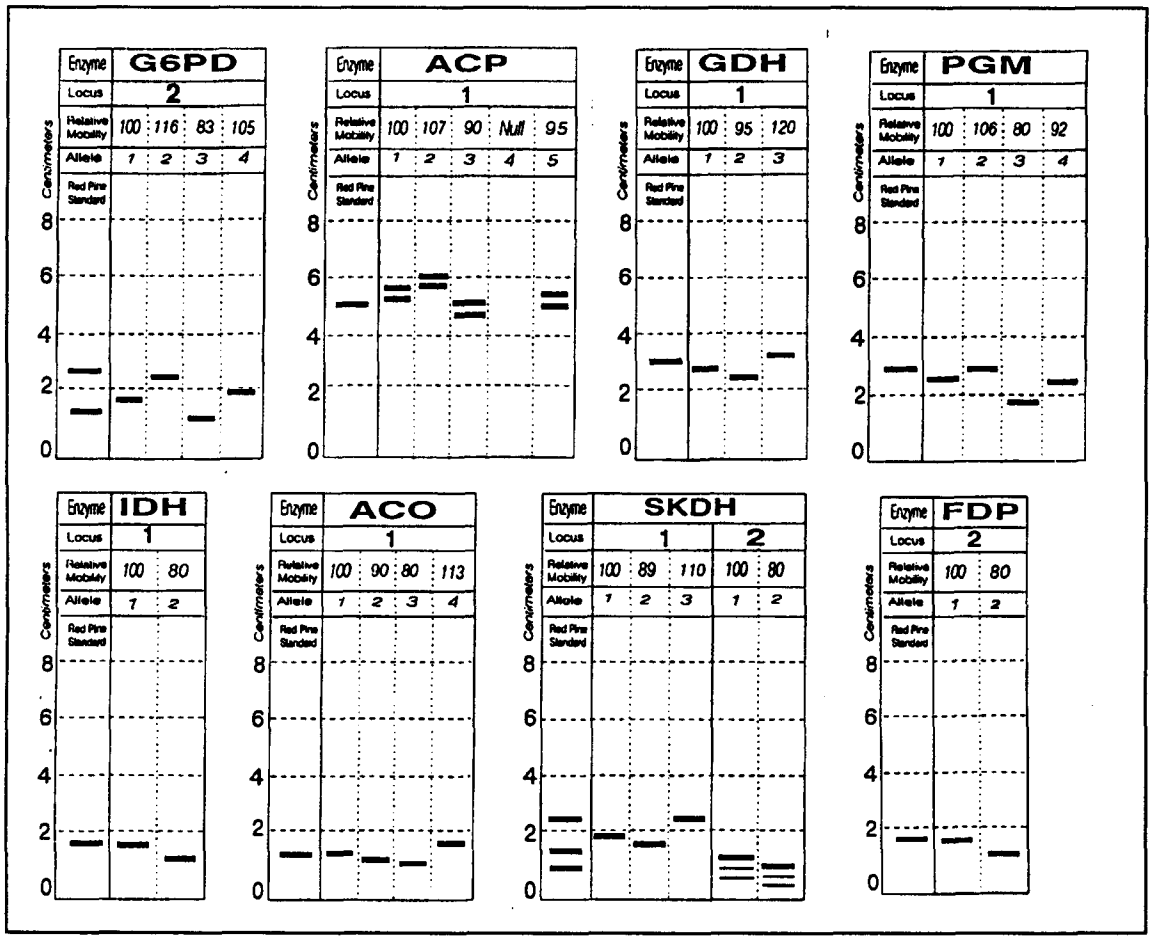


***Figure 4**: Enzyme phenotypes of G6PD, ACP, GDH, PGM, IDH, ACO, SKD and FDP. Mobility is expressed relative to the most frequent allele (No. 1) whose mobility is set to 100. Red pine phenotypes are shown in the far left column.*

GDH:

GDH stained rather faintly but consistently for one zone which is common for pine species (*Figure 4, p. 33*) (MILLAR, 1985; STRAUSS and CONKLE, 1986; MORGANTE et al., 1993; NIEBLING and CONKLE, 1990).

PGM:

PGM showed one zone of activity (*Figure 4, p. 33*). In pine species, both one and two zones have been observed for this enzyme. One zone of activity has been reported for ponderosa and washoe pine (NIEBLING and CONKLE, 1990), knobcone, lodgpole, loblolly and jeffrey pine (CONKLE, 1979) and knobcone pine (STRAUSS, 1986). Two zones have been described for pitch pine (GURIES and LEDIG, 1978; GURIES et al., 1978), eastern white pine (ECKERT et al., 1981) and bishop pine (MILLAR, 1985).

IDH, ACO:

There was a single zone of activity on gels stained for IDH and ACO (*Figure 4, p. 33*) which is common for pine species (GURIES and LEDIG, 1978; MILLAR, 1985; NIEBLING and CONKLE, 1990; CONKLE, 1979; STRAUSS, 1986; SHURKHAI et al., 1992).

SKDH:

SKDH is an enzyme infrequently assayed in forest trees. Two polymorphic loci stained faintly but consistently (*Figure 4, p. 33*). Two loci were also reported for other pine species by NIEBLING and CONKLE (1990), STRAUSS and CONKLE (1986), MORGANTE et al. (1993) and SHURKHAI et al. (1992).

FDP:

FDP showed two polymorphic zones of activity with the most anodal zone staining very inconsistently (*Figure 4, p. 33*). Consequently, only locus two was scored. Two zones of activity were also reported by STRAUSS and CONKLE (1986) for knobcone pine (*Pinus attenuata Lemm.*), with the most anodal zone being monomorphic. Two monomorphic zones were found in *Pinus muricata D. Don* (MILLAR, 1985).

UGPP:

UGPP revealed three polymorphic zones of activity (*Figure 5, p. 35*). This enzyme is rarely assayed in trees. STRAUSS and CONKLE (1986) are the only ones who have assayed this enzyme previously in a pine species. These authors described five zones of activity with four being invariant.

FUM:

Two zones of activity were observed for FUM. The most anodal zone, however, stained too inconsistently to be assayed (*Figure 5, p. 35*). STRAUSS and CONKLE (1986) described three loci, all being monomorphic in *Pinus attenuata Lemm.*

ADH:

ADH was polymorphic for two loci (*Figure 5, p. 35*). The most anodal zone however could not be scored due to inconsistent resolution. Two loci were reported in other pine species by STRAUSS and CONKLE (1986) and GONCHARENKO et al. (1993). NIEBLING and CONKLE (1986) also were unable to score the most anodal locus in ponderosa and washoe pine.

PGI:

Congruent with other pine species, PGI showed two loci (*Figure 5, p. 35*). Again, the most anodal zone stained too inconsistent to be scored. Therefore, only the locus coding the most

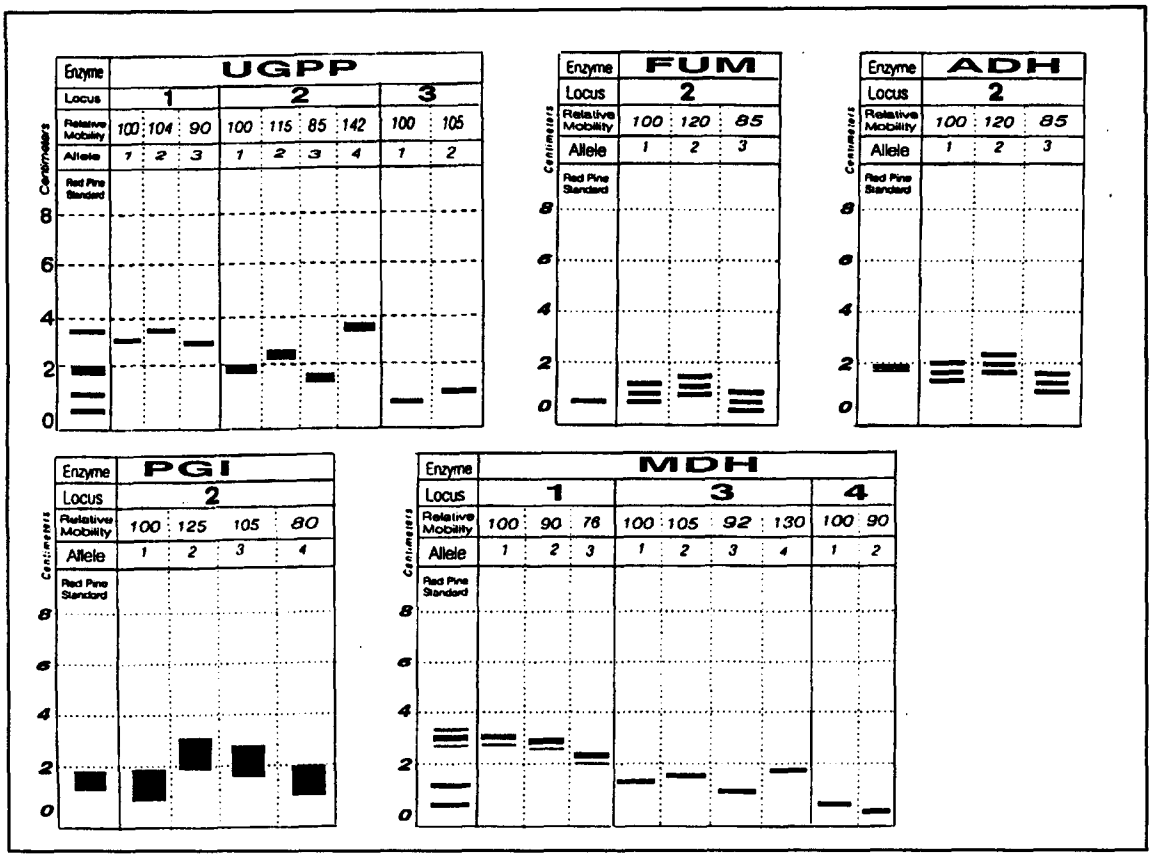


Figure 5: *Enzyme phenotypes of UGPP, FUM, ADH, PGI and MDH. Mobility is expressed relative to the most frequent allele (No. 1) whose mobility was set to 100. Band patterns of red pine are shown in the far left column.*

cathodal zone was assayed for PGI.

MDH:

MDH is the most complex enzyme system in pine species. Three complexities make allozymes of this enzyme difficult to assay. First, the two most anodal loci code allozymes having the same migration distances, thus the allozymes of one locus are often hidden by allozymes of the other locus. Second, interlocus hybrid bands between locus 1 and 3 are formed which complicate the scoring and third, a modifier locus is present which causes very slight shifts in certain alleles of both second and third loci (HARRY, 1983). In this situation, only three easily interpretable loci were scored (*Figure 5*). The hidden locus (number 2) as well as the modifier locus were not assayed.

3.3.3 Coding of allozyme genotypes for multivariate analyses

For use in multivariate analyses, it was necessary for genotypic data to be transformed into scores. Following SMOUSE et al. (1982), the diploid genotype of each parent tree was transformed into genotypic scores for each allele by assigning a value of 1, 0.5 or 0 depending on whether the allele was homozygous, heterozygous or absent in the diploid individual. This scoring system has the advantage of representing the frequency of any one allele at a given level of sampling (locus, tree, population). Moreover, variables which are transformed in this way, will show a multivariate normal distribution and will meet the assumptions for multivariate procedures (WESTFALL and CONKLE, 1992; WESTFALL, 1992). For each locus, the number of variables needed to code all observed genotypes is equal to the maximum number of observed alleles minus one. An example of the scoring system is given in *Table 3, p. 36* . A total of 71 variables was necessary to code the multilocus genotypes of all the 31 assayed loci.

Possible combinations of genotypes

Variables	11	12	13	22	23	33
Variable 1	1	0.5	0.5	0	0	0
Variable 2	0	0.5	0	1	0.5	0

Table 3: Transformation of genotypes into scores. Example of a locus with three alleles, numbered 1, 2 and 3. Two variables are needed to code all existing genotypes. Allele 3 is implicitly coded having a value of 0 for both variables

3.4 Assessment of seedling quantitative trait data

3.4.1 Common garden studies

Genotypic values of parent trees were evaluated in separate common garden seedling tests for the two geographic regions sampled. Seeds of parent trees were sown in replicated, randomized test designs in nursery beds in Corvallis, Oregon and seedling quantitative traits were measured for the first three years. Although the first years of establishment are critical for survival and adaptation, genotypic values of parent trees evaluated from such tests must be interpreted with some caution. Maladaptation to site conditions may not be revealed in the first few years. It may depend on the occurrence of rare climatic extremes or on the accumulation of seemingly minor damages. In harsh environments, detrimental effects and adaptive value of genotypes develop more rapidly than in favorable environments. Consequently, to evaluate the genotypic values of the parent trees, two contrasting environments were created in the nursery beds: a favorable environment with optimal growth conditions and a harsh environment with stressful periods for the seedlings. The two contrasting environments consisted of a cold air treatment in normal, uncovered beds and a warm air treatment, which was obtained by erecting polyethylene covered frames over the nursery beds. The covers were put on in mid November and removed at the beginning of April. The temperature differences between the covered and the uncovered beds were 5.5°C on a cold day (with an outside temperature of 4.5°C) and 19.5°C on a warmer day (with an outside temperature of 13°C). Budbreak in the warm treatments occurred about 15 days earlier than in the uncovered beds. In addition, warm air treatments were regularly watered and lightly fertilized with a balanced liquid fertilizer with microelements added. In contrast, the cold air treatments were only slightly fertilized in the first year. In subsequent years no fertilizer was applied and watering was reduced to the very minimum in order to create a stressful environment. The nursery beds (cold frames made of wood, 14.4 m by 1.50 m by 1.0 m) contained a 100 cm deep soil of agricultural origin. Before sowing, seeds were stratified at 2°Celsius for 60 days.

3.4.1.1 Central Oregon

Seeds from 307 parent trees were sown in two different years (1986 and 1987) to include year effects (replication in time). Tests were established as split plot designs using family sub-plots within environmental-treatment main plots. Four-seedling family-row-plots were used as experimental units. Two completely randomized blocks (replications) were used within each treatment and sowing year. The complete test therefore consisted of two sowing years, two treatments within each sowing year, two replications within each treatment and sowing year and four seedlings per family within each replication, eight seedlings per family within each treatment, sixteen seedlings per family within each sowing year and 32 seedlings per family for the complete test. Each block within a bed contained 78 rows with 16 positions within a row. Each seedling was characterized by its treatment, block number, row number, plot number within the row and position number within the plot. Two rows of border trees were used along the length of the beds, with 4 border rows at the ends.

Three ungerminated seeds per spot were sown in the first week of April in 1986 and in the third week of April in 1987. In the fall of the first year the most western (or a dead/damaged) seedling per spot was removed. In spring of the second year, the most eastern (or a dead/damaged) seedling per spot was eliminated (pruned). Spacing of the seedlings was then 7.5 cm between seedlings within rows and 9 cm between rows.

5% of the seedlings were attacked by insects, which affected the growth-form of the leader. 10% of the seedlings were lost due to drought in the fall of the third year.

Measurements were made over three years. The following traits were assessed for each seedling or family plot:

- EMERG: Date of seedling emergence, recorded every 3.5 days, as days after the emergence of the very first seedling of the test.
- GERM: Percentage of germination (plot means of the 12 seeds sown per plot).
- COTY: Number of cotyledons recorded on 8 seedlings.
- BUDSET1: Date of first-year budset, recorded once a week.
- SECNED: Presence and quantity of secondary needle fascicles on 1 year old seedlings.
- FHT1: Final first year height.
- H21: Early second year height.
- H22: Mid second year height.
- FHT2: Final second year height.
- DIA2: Final second year diameter.
- H31: Early third year height.
- H32: Mid third year height.
- FHT3: Final third year height.
- DIA3: Final third year diameter.
- TDWT: Final third year top dry weight (plot means for 4 seedlings).
- RDWT: Final third year root dry weight (plot means for four seedlings).

3.4.1.2 Southwest Oregon

Seeds from 268 parent trees were sown in spring 1989 (one year only). The test was established as a split plot design using family sub-plots within environmental treatment main plots. Four seedling family-row-plots were used as the experimental units. Each treatment was replicated in 3 complete randomized blocks. The complete test, therefore, consisted of two treatments, three replications within each treatment and four seedlings per family within each replication, 12 seedlings per family within each treatment and 24 seedlings per family for the complete test. Each block within a bed contained 67 rows with 16 positions within a row. Each seedling was characterized by its treatment, block number, row number, plot number within the row and position number within the plot.

Three ungerminated seeds were sown on April 17 and 18 in 1989. All other experimental conditions (treatment, soil, pruning) were the same as in the Central Oregon Mountain study, but spacing between seedlings was 7 cm between seedlings within rows and 9 cm between rows. No losses due to insects or drought occurred in this study.

Measurements were made over a three year period. The following traits were assessed for each seedling or family plot:

- EMERG: Date of seedling emergence (plot means only) in days after emergence of first seedling in the test.
- GERM: Percentage of germination (plot means of the 12 seeds sown per plot).
- BUDSET1: Date of first year budset, recorded once a week.
- FHT1: Final first year height.
- FHT2: Final second year height.
- BUDBUR3: Date of budburst at beginning of third growing season.
- FHT3: Final third year-height.
- DIA3: Final third year-diameter.

3.4.2 Preparation of seedling raw data for analyses

In a first step, all original seedling traits were tested for location effects within the nursery beds. All traits were regressed against the location variables (row, position, row^2, position^2 and row x position). Traits which were influenced by location effects within the nursery beds (R^2 greater than 0.10) were adjusted using regression procedures in order to remove these location effects.

In the Central Oregon study, only final third year diameter for the cold air treatment in replication 2 of the 1987 sowing showed location effects and was adjusted. In the Southwest Oregon study, final third year diameter in replication 1 of the cold air treatment and final second year height in replication 3 of the cold air treatment were adjusted for location effects.

In a second step, differences among the replications (nursery beds) were tested using analysis of variance procedures. Significant differences were found for the following variables:

Study - treatment - year	Significant differences among replications (*p=0.0001*)
Central Oregon - cold - 1986	EMERG, BUDSET1, FHT1, DIA2, H32, FHT3, DIA3, TDWT, RDWT
Central Oregon - warm - 1986	EMERG, BUDSET1, FHT1, HT21, HT22, FHT2, DIA2, H31, H32, FHT3
Central Oregon - cold - 1987	EMERG, FHT1, H21, H22, FHT2, DIA2, H31, FHT3 DIA3, TDWT, RDWT
Central Oregon - warm - 1987	EMERG, FHT1, H21, H22, H31, RDWT
Southwest Oregon - cold - 1989	EMERG, BUDSET1, FHT2, FHT3, DIA3
Southwest Oregon - warm -1989	EMERG, GERM, FHT1, BUDSET1, FHT2, FHT3

***Table 4**: Traits exhibiting significant differences among replications in the nursery tests. Abbreviations for traits are given on p. 37*

Because of heterogeneous means and variances, linking the replications (and different sowing years in the Central Oregon study) required that the separate data sets be transformed and scaled to obtain a single dataset. Following a method used by REHFELDT (1989a), traits measured within each replication, treatment (and sowing year for the Central Oregon study) were transformed in a third step to standard normal deviates:

$$Z_{ijkl} = \frac{(X_{ijkl} - \bar{x}_{ijk})}{\sigma_{ijk}} \qquad \text{or} \qquad Z_{ijk} = \frac{(X_{ijk} - \bar{x}_{ij})}{\sigma_{ij}}$$

for Central Oregon for Southwest Oregon

Where Z_{ijkl} is a standard normal deviate for seedling l from sowing year k, tested in replication i of treatment j in the Central Oregon study; X_{ijkl} is an original observation; and $\bar{x}_{ijk}$ and σ_{ijk} are the mean and standard deviation of all individuals of replication i in treatment j and sowing year k. For Southwest Oregon Z_{ijk} is a standard normal deviate for seedling k from replication i in treatment j; X_{ijk} is an original observation; and $\bar{x}_{ij}$ and σ_{ij} are the mean and standard deviation of all individuals of replication i in treatment j.

Finally, family mean values of these standard normal deviates for all traits were calculated for each study and treatment separately.

Since traits were measured in different years, some traits were assessed in one study but not in the other, and some of the traits were not assessed equally in both studies, the data sets of the two studies could not be combined into one dataset covering the total sampling area.

3.5 Geographic and environmental data

3.5.1 Geographic data

Each parent tree is characterized by a number of geographic variables which were recorded during sampling (*Table 5, p. 39*). Location information was transformed to metric units where necessary. Coordinates of all sample trees were input into a geographic information system (arcinfo). The official seed zone map of Oregon was digitized and tree and seed zone information was combined into a map and corresponding data base. Seed zone memberships for all sample trees were evaluated by the geographic information system and were taken from the

data base. Besides the variables listed in *Table 5,* more variables such as aspect, slope, distance from the Cascade crest, distance to the top and bottom of the slope and sun exposure were recorded but not used for the present investigation. Since primarily environmental variables were used for the analyses, no details are given on these variables. All geographic information of the parent trees was provided by the US Forest Service, Pacific Northwest Experiment Station, Corvallis, Oregon.

Location Variable	Description
Latitude	In $^{1}/_{100}$ of degrees N
Longitude	In $^{1}/_{100}$ of degrees W
Elevation	In m above sea level
Seed zone	Number from seed zone Map of Oregon (1973)

Table 5: Geographic information on sample tree locations used for the analyses

3.5.2 Climate data

Mapping adaptive genetic variation is commonly done by relating genetic variation to geographic variation expressed by variables such as latitude, longitude and elevation. Because climate variables are partly correlated with geographic variables, climate and location data are redundant to some extent and location variables may be used as a replacement for climate variables. Often however, environmental gradients are highly complex within small areas and location data alone may not describe environmental conditions adequately. Consequently, describing adaptive variation directly on the basis of environmental variation is preferable especially in areas with rugged topography and complex climate patterns. Information on climate conditions however are commonly available for established weather stations which are infrequent in forest areas. Hence, climate conditions must be estimated for tree locations, either by interpolation or the use of climate models.

In the present investigation, climate values for tree locations were estimated using local regression models (*for details on local regression models see section 3.6.3, p. 51*). Long term climate values (dependent variables) of 195 weather stations in Oregon (JOHNSGARD, 1963) were regressed on the latitude, longitude and elevation of the weather stations (independent variables). For each dependent climate variable, different local regression models were calculated and tested in order to optimize model fit. In addition, models were evaluated using the same basic data set of the 195 weather stations as input in the regression models and for comparing the predicted values from the model with the actual values. Optimized models were then used to estimate climate variables for all parent tree locations, based on their latitude, longitude and elevation. The weather stations are geographically well distributed and sample a range of elevations (3 m to 1,973) comparable to the range of tree locations (49 m to 1,963 m). Two examples of actual and modeled values for the weather stations as well as estimated climate values for parent tree locations are shown in *Figure 6 and 7, p. 40* and *41*. Local regression models are very powerful in modeling the complex climate changes over short distances as the two examples demonstrate. However, they tend to be sensitive to edge effects. At the edges, where the data basis is insufficient due to a lack of points outside the sampling area, the modeled e.g. interpolated values become unreliable. Consequently, climate values could not be estimated for all 488 tree locations. No climate values were estimated for trees close to the limits of the sampling area. Sample size therefore is reduced to 411 trees for all analyses using climate variables.

Derived climate values for tree locations are estimates only; they reflect general weather patterns, but do not reflect local climate which is certainly modified by slope and aspect. Since no information about slope and aspect of the weather stations was available, these variables could not be used for model building.

In total, 15 climate variables were estimated for the 411 tree locations using local regression models *(Table 6, p. 42)*.

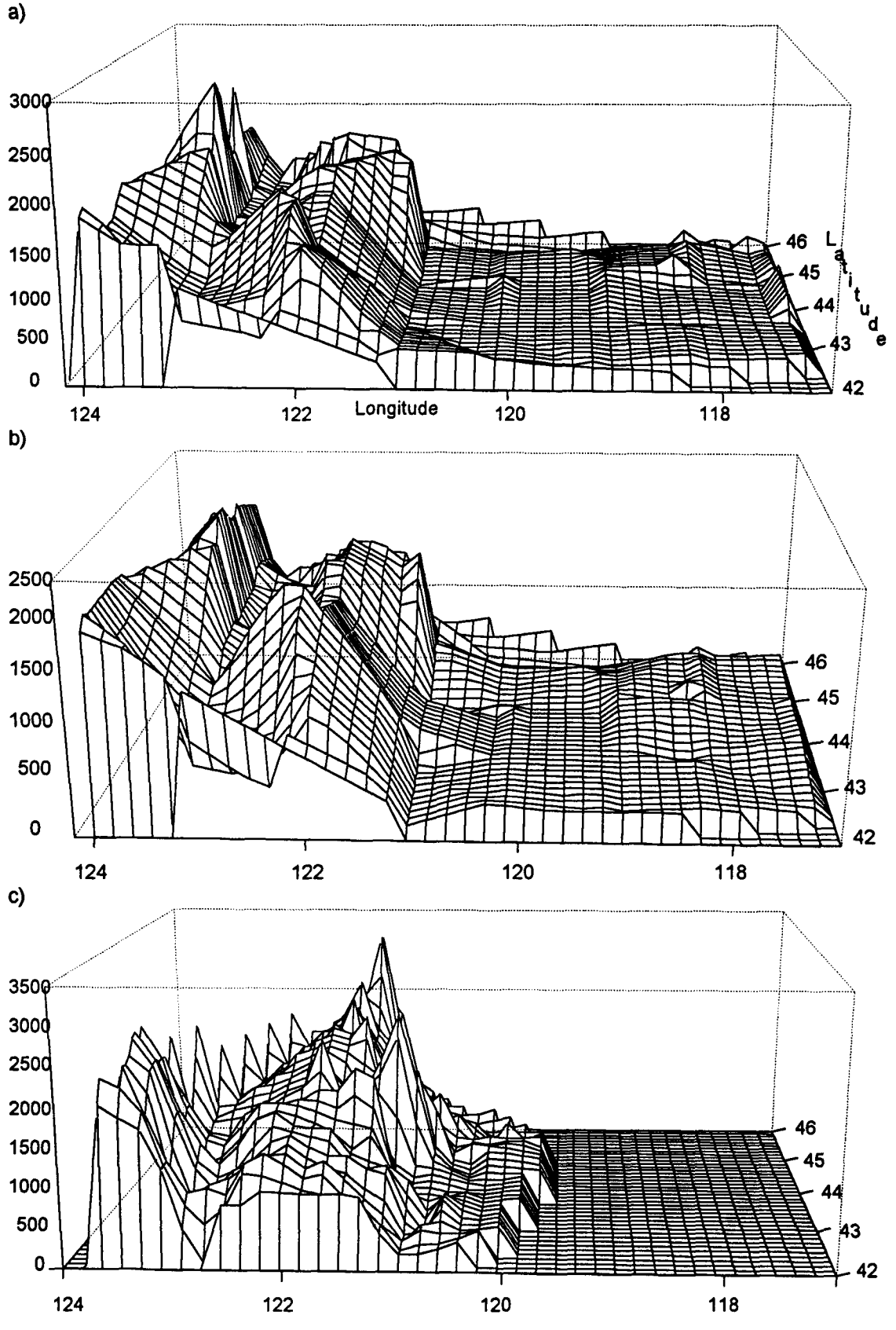


Figure 6: *Actual and estimated average annual precipitation [mm]:*
a) Actual long term values from 195 weather stations in Oregon b) Predicted values from local regression model for the same 195 weather stations c) Predicted values from local regression model for 411 tree locations

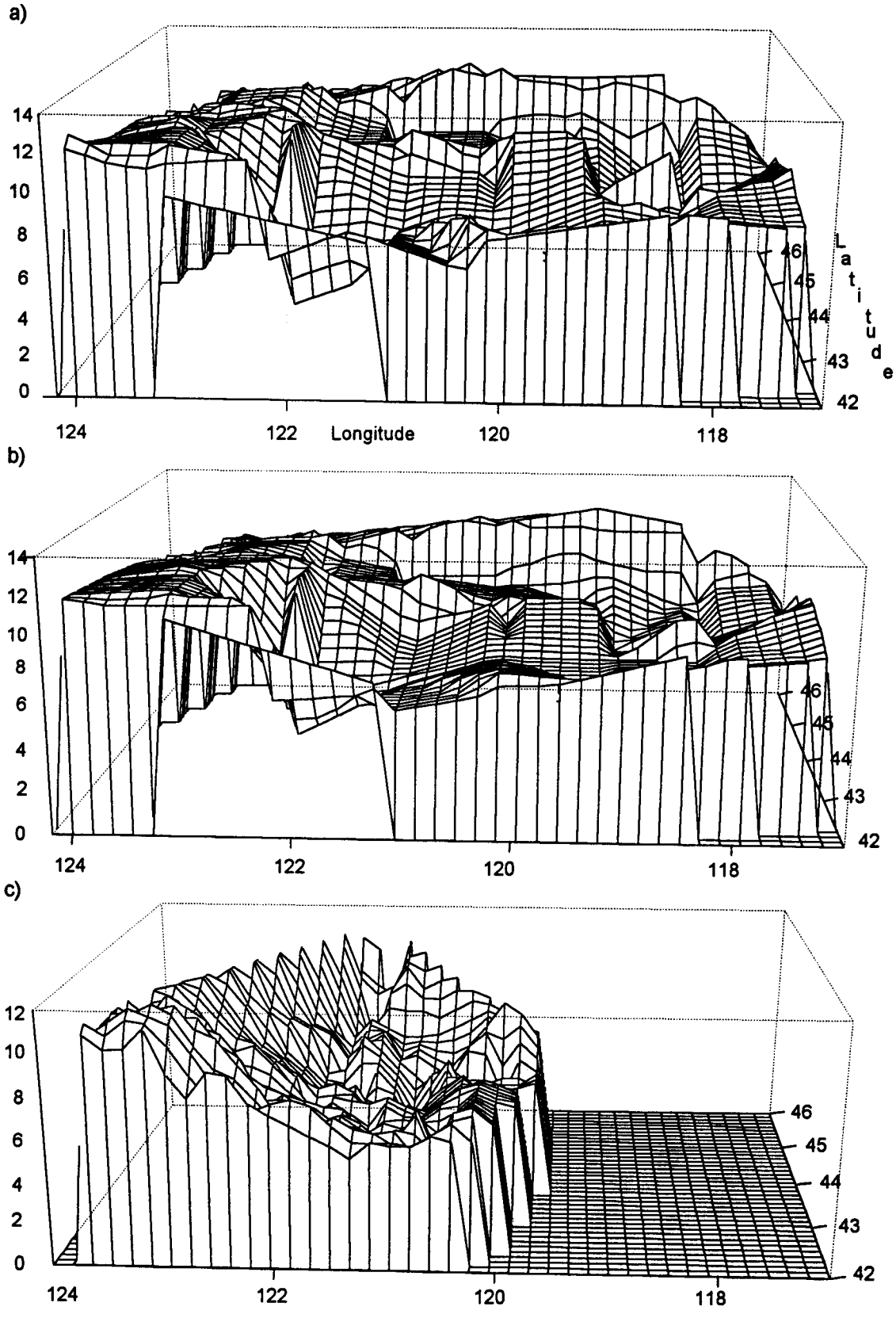


Figure 7: *Actual and estimated average annual temperature [degrees Celsius]*
a) Actual long term values from 195 weather stations in Oregon b) Predicted values from local regression model for the same 195 weather stations c) Predicted values from local regression model for 411 tree locations

REGION	SOUTHWEST OREGON				CENTRAL OREGON				Multiple model	Model Residual	Principal component Factor pattern		Final Communa-	VARIABLE
VARIABLE	MEAN	STD DEV.	MINIMUM	MAXIMUM	MEAN	STD DEV.	MINIMUM	MAXIMUM	r-square	Std Error	Factor 1	Factor 2	lities	
AVMIJA	-1.97	1.43	-5.15	0.98	-8.45	2.70	-13.50	-0.19	0.91	1.48	0.919	0.268	0.917	AVMIJA
AVMIJU	9.00	0.59	7.51	10.23	5.48	2.52	0.62	14.20	0.71	1.75	0.888	0.059	0.794	AVMIJU
AVMIYE	3.10	0.96	0.79	4.99	-1.39	2.40	-5.61	6.64	0.86	1.32	0.930	0.178	0.897	AVMIYE
AVMAJA	7.03	1.91	2.65	10.63	2.50	0.76	-0.02	4.94	0.80	1.53	0.936	0.092	0.885	AVMAJA
AVMAJU	27.60	1.56	23.78	30.82	25.55	1.94	20.88	29.91	0.86	1.51	0.722	-0.545	0.819	AVMAJU
AVMAYE	16.19	1.52	12.46	19.37	13.28	1.54	7.88	16.78	0.81	0.88	0.918	-0.225	0.894	AVMAYE
AVJA	2.11	1.43	-1.12	5.01	-2.90	1.67	-6.44	2.44	0.92	1.11	0.952	0.222	0.955	AVJA
AVJU	18.22	0.97	15.62	20.40	15.45	1.72	10.90	20.60	0.83	1.16	0.944	-0.254	0.956	AVJU
AVYE	9.92	1.21	6.61	11.94	5.92	1.57	1.98	10.97	0.90	0.73	0.994	-0.004	0.989	AVYE
APJA	224.87	113.47	43.40	480.30	161.13	104.52	14.98	423.60	0.93	33.4	0.359	0.903	0.945	APJA
APJU	11.71	6.83	1.64	39.42	20.78	10.11	4.60	63.29	0.76	4.14	-0.733	0.527	0.813	APJU
APYE	1362.87	619.15	354.50	2795.00	1137.06	679.40	141.60	2693.00	0.89	264.1	0.222	0.964	0.980	APYE
SURP	1047.16	598.72	62.00	2413.00	866.22	605.98	9.00	2285.00	0.91	202.8	0.212	0.956	0.959	SURP
DEFIC	309.57	54.86	157.00	430.00	244.05	97.99	17.00	443.00	0.91	43.9	0.513	-0.778	0.868	DEFIC
WATBI	737.56	617.46	-248.00	2148.00	621.00	697.87	-390.00	2266.00	0.92	228.7	0.127	0.978	0.974	WATBI
FACTOR 1	0.863	0.574	-0.589	1.886	-0.735	0.628	-2.552	1.160			8.569	5.081		Eigenvalue
FACTOR 2	0.037	0.817	-1.386	1.968	-0.031	1.134	-1.787	2.685			57.13%	33.90%		Proportion

VARIABLES AVMI: Average daily minimum temperature [Celsius] JA=January
AVMA: Average daily maximum temperature [Celsius] JU=July
AV: Average daily temperature [Celsius] YE= YEAR
AP: Average Precipitation [mm]

SURP: Average cumulative annual moisture surplus [mm]
DEFIC: Average cumulative annual moisture deficit [mm]
WATBI: Average annual water balance [mm]
Sample size: Total: 411
Southwest: 189
Central Oregon: 222

DATA:
Based on predicted values for 411 tree locations
Predicted values: From local regression models, developped from 195 weather stations in Oregon [see model parameters]

Factors derived by principal component analysis from the correlation-matrix of predicted variables from 411 tree loc.

Final communality: Proportion of a variables variance accounted for by the two retained principal components Factor 1 and Factor 2

Table 6: *Characteristics of estimated climate variables and derived principal components for the 411 tree locations by regions*

Estimated climate variables were then subjected to a principal component analysis (*see section 3.6.4.1 p. 54*) in order to remove strong intercorrelations among the single variables and to reduce the number of variables for further analyses. Two major principal components accounted for 91% of the variation in the original data set (*Table 6, p. 42*). High final communalities for all original variables indicate an excellent factor solution; on average more than 90% of the variance of each original variable is retained by the two principal components. In addition, the extracted principal components are readily interpretable: The first principal component (*Factor 1*), accounting for 57.13% of total variation, represents a temperature gradient. Increasing values of *Factor 1* mean milder winters but hotter and drier summers (as the negative and high loading for average July precipitation indicates), while decreasing values mean colder winters but cooler summers and a more balanced precipitation regime over the year. The second component (*Factor 2*), accounting for 33.9% of total variation, represents a moisture gradient which is independent of the variation already extracted by *Factor 1*. Increasing values of *Factor 2* mean higher rainfall in all seasons (but especially in winter) and lower overall moisture deficits (or a positiver water balance), while low values characterize drier climates with lower precipitation in all the seasons (but especially in winter), exhibiting an increasingly negative water balance over the year.

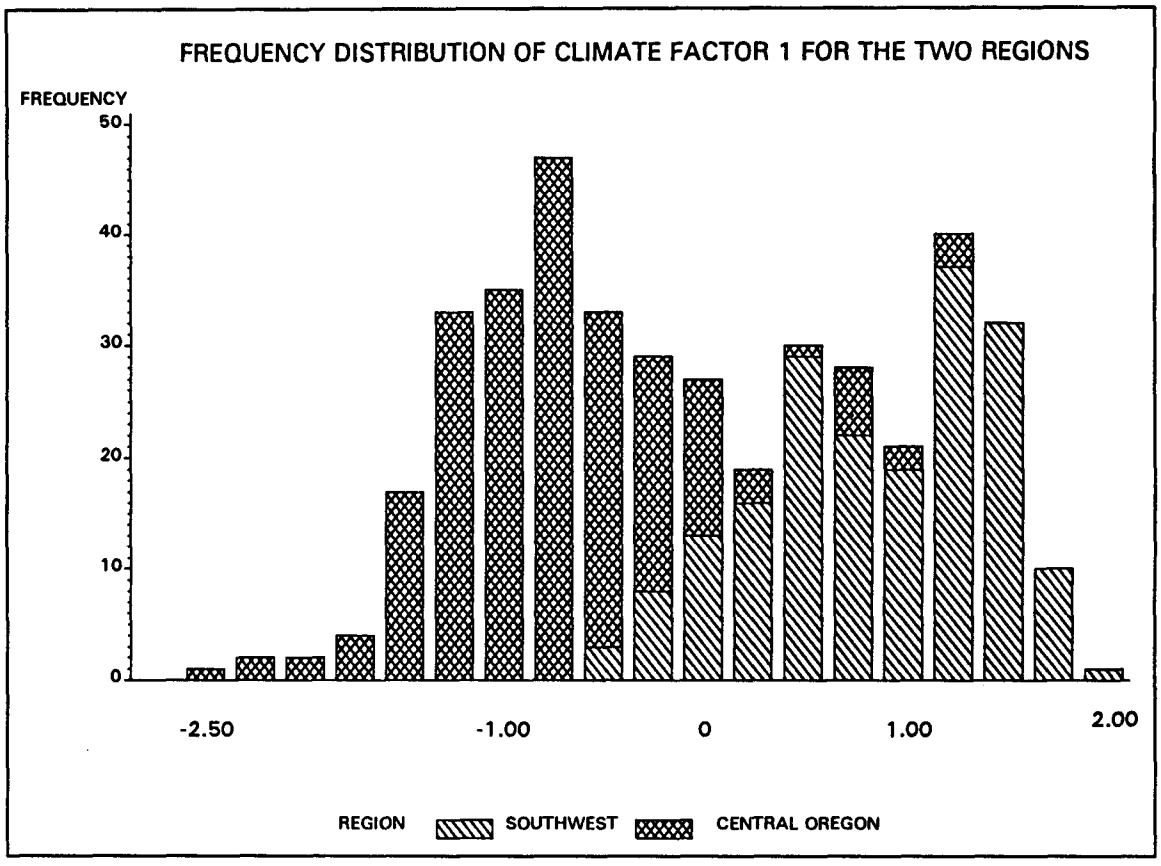
Values in *Table 6, p. 42* and the frequency distributions of the two principal components in *Figure 8, p. 44* demonstrate, that climatic conditions differ considerably between the two regions. Southwest Oregon has warmer winters, but hotter and drier summers than Central Oregon (Mean values on first principal component of 0.863 compared to -0.735). Climate is much more oceanic with less temperature extremes (especially in winter) than found east of the Cascade Range. Precipitation is slightly higher on average and less extreme regarding water balance than on the east side of the Cascades. However, the major part of precipitation occurs during the fall, winter and spring; summers are characterized by a period of very low precipitation. The east side of the Cascade Range has somewhat less pronounced seasonal differences in precipitation with higher precipitation during the summer months than on the west side of the range. Overall however, water balance is less favorable in areas to the east of the Cascades and in the high desert area. Mean, minimum and maximum values on *Factor 2* indicate a higher variability as well as greater extremes in the water balance than occurs to the west of the Cascade Mountains. Although Southwest Oregon has more favorable overall moisture conditions, dry sites with unfavorable water balance do occur. A zone situated between the Coast Range and the mounting slopes of the Cascade Mountain (longitude 122.5 to 123.5) is especially characterized by low rainfall and dry overall conditions (*Figure 6 c, p. 40*).

The considerable differences in climate between the two regions are also illustrated by results of a one-way analysis of variance, indicating the percentage of variance accounted for by regional differences (*Table 7, p. 45*). All temperature variables and the composite variable *Factor 1* show a highly significant influence of region, indicating the marked differences in the temperature regime between the two regions, especially during winter. Rainfall and water balance variables show less marked differences between the two regions. The highest difference is manifested in summer rainfall, being higher in Central Oregon than in Southwest Oregon. Overall, there is not much difference in water balance among the regions, however, since the difference in *Factor 2* is not significant ($p=0.41$).

3.6 Statistical methods and models

3.6.1 Testing spatial structures of genetic variation

The recognition of patterns of genetic variation in space is a first step toward an analysis of the causes that lead to the establishment of these variation patterns. In gene-ecology, such forces generating non-random patterns might, for example, correspond to population size and structure, adaptation to climate and edaphic gradients, isolation by distance for selectively neutral genotypes or historical migration patterns. Testing, describing and quantifying spatial structures in genetic data may therefore shed light on the evolutionary forces involved in shaping the patterns. Two different methods were applied to test the existence and to describe patterns of genetic variation related to location or environment:



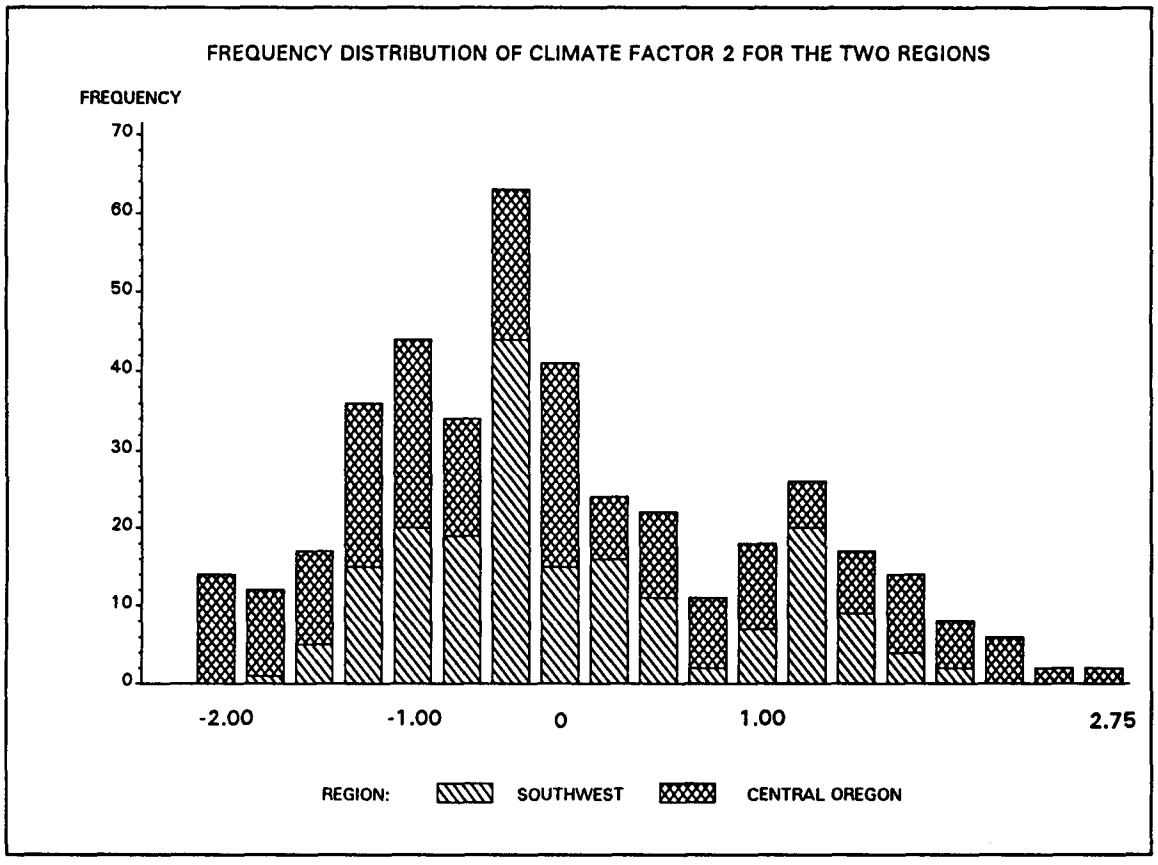


Figure 8: *Frequency distributions of climate principal components Factor 1 and Factor 2 for the two regions, based on 411 tree locations*

VARIABLE	SSQ REGION	SSQ TOTAL	F - VALUE	Prob p	% Var among REGIONS	% Var within CENTRAL	% Var within SOUTHWEST
AVMIJA	4349	6290	916.3	0.0001	69.1	24.6	6.3
AVMIJU	1298	2729	371.2	0.0001	47.6	49.3	3.1
AVMIYE	2101	3495	616.5	0.0001	60.1	34.6	5.3
AVMAJA	2086	2911	1034.8	0.0001	71.7	4.3	24.0
AVMAJU	420	1712	133.1	0.0001	24.5	48.7	26.8
AVMAYE	860	1816	367.7	0.0001	47.3	28.7	24.0
AVJA	2595	3574	1084.6	0.0001	72.6	16.5	10.9
AVJU	790	1606	395.7	0.0001	49.2	39.7	11.1
AVYE	1650	2448	845.5	0.0001	67.4	21.3	11.3
APJA	423996	5229288	36.1	0.0001	8.1	45.9	46.0
APJU	8329	39785	108.3	0.0001	20.9	56.8	22.3
APYE	5592039	178638550	13.22	0.0003	3.1	56.6	40.3
SURP	3553361	151136695	9385	0.0018	2.3	53.3	44.4
DEFIC	411921	3102657	62.6	0.0001	13.3	68.2	18.5
WATBI	1550884	179898896	3.56	0.0600	0.8	59.4	39.8
FACTOR 1	264.2	410	741	0.0001	64.4	20.4	15.2
FACTOR2	0.6581	410	0.66	0.4179	0.1	68.9	31.0

Table 7: *Analysis of variance, testing differences of climate variables among the regions. Partitioning of variance into components among and within regions*

1) Tests of matrix associations using genetic, geographic and climatic distance measures among the sample trees and their locations, and
2) Spatial autocorrelation analysis.

3.6.1.1 Mantel test for matrix associations

A useful, non-parametric method for comparing two distance matrices was introduced by MANTEL (1967). The procedure, known as the Mantel Test, was further developed by ROYALTEY et al. (1975) and SOKAL (1979a) and was applied to testing spatial patterns of genetic variation by different authors such as JONES et al. (1980), SOKAL et al. (1980), PIGLIUCCI et al. (1990) and others. Basically, the procedure tests the null hypothesis that differences between pairs of samples in a given variable are independent of their geographic or environmental differences. In a first step, different similarity or distance matrices from genetic and geographic or environmental data are constructed. In order to test and quantify any associations between genetic and geographic or environmental distance, the genetic distance matrix is compared in a second step to each of the other distance matrices. The two matrices are compared by computing the sum of the products of their corresponding elements over all rows and columns as

$$Z = \sum_i \sum_j x_{ij}\, y_{ij}$$

and comparing Z with that expected under an empirical null distribution (E(Z)). The expectation of Z is computed on the null hypothesis of random permutations of the rows and columns of one matrix, say X and the computation of the inner product of each of these permutations with the matrix Y. The test for significance of Z being different from random expectations is given as the departure t from random expectation relative to its permutational standard deviation, *i.e.*

$$t = \frac{[Z - E(Z)]}{\sigma(Z)}$$

which follows the standard normal distribution. Alternatively, the calculated z value may directly be compared with a generated and sorted z - distribution for say 1000 random permutations, and probabilities of obtaining a value greater or smaller than the calculated z may be directly inferred from this randomized distribution. Besides testing the significance of any association, the strength of association may also be quantified by calculating a matrix correlation which is the ordinary Pearson coefficient of correlation between the elements of the two matrices. The

formula for the computation of the correlation coefficient in a Mantel test is given by MANLY (1985).

A number of geographic, climatic and genetic distance matrices were constructed. Distances in kilometers between sample trees were calculated using the coordinates of the sample trees, given as latitude and longitude. A matrix of reciprocal distance was calculated as

$$\frac{1}{d+1}$$

with d being the distance in kilometers between respective sample trees. The reciprocal distance matrix considers all longer distances to be more or less equal, while emphasizing differences among shorter distances. Consequently, the reciprocal distance matrix focuses on local, small scale patterning while the normal distance matrix emphasizes long distance e.g. large scale patterning. Testing for regional differences was done by constructing a third distance matrix with a dummy coding for regional membership of each tree, assigning a value of 0 for trees coming from the same region and a value of 1 for trees coming from a different region (Central Oregon-Southwest Oregon).

Climate distance matrices were constructed following a method suggested by MANTEL and VALAND (1970). A distance matrix was calculated for each of the two principal components e.g. climate *Factor 1*, *Factor 2* and for both components combined, using ranks. The values for these variables were ranked from 1 to n (sample size) with any equal values being given an average rank. The distance between sample location i and j was then calculated as

$$d_{ij} = |r_i - r_j|$$

Four different distance matrices were constructed for genetic data. Genetic distances among individual sample trees were calculated using the two distance measures proposed by GREGORIUS (1974) and by NEI (1972, 1978), using either all analyzed 31 loci or only those 15 loci which showed significant spatial patterns in the spatial autocorrelation analysis (*see section 3.6.1.2, p. 47*). Genetic distance by GREGORIUS is defined as

$$d_{pq} = \sum_l 0.5 \sum_k |p_{lk} - q_{lk}|$$

with p_{lk} and q_{lk} being the allele frequencies of sample p and q for allele k at locus l

Consequently, it measures the proportion of alleles not shared by tree p and q. A value of zero is reached only if tree p and q have all alleles in common and the maximum value of 1 is obtained only if they have absolutely no alleles in common. In addition, distances are symmetrical and the distance between tree p and q cannot exceed the sum of their distances to a third tree r. It must be noted here, that this measure of distance is identical to the distance already used earlier by Prevosti (WRIGHT, 1978).

Genetic distance by NEI is defined as

$$d_{pq} = -\log_e \left(\frac{\sum_l \sum_k p_{lk} q_{lk}}{\sqrt{\sum_l \sum_k p_{lk}^2 \sum_l \sum_k q_{lk}^2}} \right) \quad \text{or as}$$

$$d_{pq} = -\log_e \left(\frac{(2n-1) \sum_l \sum_k p_{lk} q_{lk}}{\sqrt{\sum_l (2n \sum_k p_{lk}^2 - 1) \sum_l (2n \sum_k q_{lk}^2 - 1)}} \right)$$

with p_{lk} and q_{lk} being the allele frequencies for samples p and q for allele k at locus l and n being the sample size. The second formula is the unbiased genetic distance (NEI, 1978). Nei's genetic distance is zero when both samples have the same alleles at all loci; it is infinite when they have no alleles in common.

All distance matrices were calculated using especially developed SAS-IML programs (© SAS

Institute Inc., Cary, NC, USA). Computations of the Mantel tests were carried out using a FORTRAN program running on a mainframe computer. The FORTRAN program was developed based on a program listing given by MANLY (1985) which was adapted and extended to the special needs of the analysis, the input and output requirements and the FORTRAN version.

3.6.1.2 Spatial autocorrelation

Spatial autocorrelation is defined as the association of the value of a variable with the value of the same variable at geographically adjoining locations (SOKAL and ODEN, 1978a). Spatial autocorrelation analysis was developed by MORAN (1950), GEARY (1954), CLIFF and ORD (1973), and was adapted to ecological and biosystematic research mainly by JUMARS et al. (1977), JUMARS (1978), SOKAL and ODEN (1978a, 1978b), and SOKAL (1978). Spatial autocorrelation has been applied to the study of gene frequency distributions by different authors, such as SOKAL and ODEN (1978b), SOKAL and MENOZZI (1982), EPPERSON (1992), EPPERSON and ALLARD (1989), BARBUJANI and MILANI (1986), PERRY and KNOWLES (1991), PIGLIUCCI et al. (1990), GEBUREK (1993) and others.

The computation of spatial autocorrelation requires a set of localities represented as points in the plane. One or more variables are mapped onto these points with one value per variable for each point. Association among values of a variable with the values of the same variables at adjoining localities are then computed. If the value of the variable is dependent on the values from neighboring localities, the variable is said to exhibit spatial autocorrelation. Amounts, signs and changes of associations with distance classes may help in understanding the variation pattern of the respective variable in space. Spatial autocorrelation analysis may help in the description of gene frequency variation in space without any preliminary assumptions about the underlying population structure. Each allele and each gene locus must be taken into account separately, however. Hence, spatial autocorrelation is a multi-variable rather than a multi-variate approach.

Basically, allele frequencies for all sampled locations are compared to all other localities within suitably chosen distance classes, autocorrelations are computed for each distance class, and deviations from expectations are tested for significance. A graph of the relationship between spatial autocorrelation coefficients and distance is then constructed in order to help interpretation of the spatial structure of the data. Such graphs are called correlograms; they summarize the pattern of geographic variation exhibited by the response surface of a given allele (e.g. *Figure 9, p. 49*). Inferences about the spatial structure of the allele frequency may then be made based on the shape of the correlogram.

Autocorrelation normally is estimated using Moran's *I* coefficient which is defined as

$$I = \frac{n \sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i \sum_j W \sum_i (x_i - \bar{x})^2}$$

where in our case x_i and x_j are the allele frequencies of the *i* th and *j* th sample tree, $\bar{x}$ is the mean frequency of the *n* samples, w_{ij} is a weighting factor that was set to 1 for all pairs of sample trees falling in the distance class considered, and equal to 0 for all other pairs, and *W* is the number of pairs in the distance class of interest.

Geographic distance in kilometers among all sample trees were calculated and all pairs of trees were classified into five distance classes (*Table 8, p.48*). Class widths were chosen in order to have roughly the same numbers of pairs in each distance class.

Moran's *I* may range from a maximal value of 1 for a positive association of allele frequencies for a certain distance class to a minimal value of -1 for a negative relationship. Values significantly greater than the expected value of

$$E(I) = -(n-1)^{-1}$$

indicate a greater similarity of individuals within that distance class than would be expected by chance, while values of *I* significantly lower than *E(I)* signify less similarity among individuals within that class than would be expected were chance the only factor determining the spatial distribution of the allele in question. Differences between observed and expected values of *I* were tested for significance (for levels of *p=0.05* and *p=0.01*) following the formulas given in SOKAL and ODEN (1978a). For diallelic loci only one allele was considered as the second would contribute identical information.

Total sampling area	Southwest Oregon	Central Oregon
0-50 km	0-25 km	0-25 km
51-100 km	26-50 km	26-50 km
101-150 km	51-100 km	51-100 km
151-250 km	101-150 km	101-150 km
>250 km	>150 km	>150 km

***Table 8**: Distance classes used for spatial autocorrelation analyses of allele frequencies*

Three major patterns common in correlograms are described by SOKAL and ODEN (1978a) as follows:

- The "cline": characterized by a monotonic decrease of *I* with distance, starting from significant positive values and ending at significant negative values (e.g. *Figure 9, p. 49, G6p2-1*).
- The "depression": characterized by significant positive values of *I* in the first distance class, a near-zero non-significant *I* in the second class, a negative significant *I* in the third and fourth class and a non-significant near-zero *I* in the last distance class (*Figure 9, p. 49, Mnr1-1*).
- The "crazy quilt": characterized by high values of *I* surrounded by low values and vice versa without any clear ordering (*Figure 9, p. 49, Mnr1-3*).

SOKAL (1979b) has demonstrated that spatial correlograms indicate patch size e.g. the diameter of a homogeneous area in the surface of the variable. The distance at which the correlogram first intercepts the expected value *E(I)* represents the shortest side of an irregular patch on the surface of the variable. Patch size gives information about the scale and resolution of the variation in space and may help in interpretation of eventual shaping forces.

Morans' *I* for all the distance classes and significance test were computed using especially developed SAS-IML programs (© SAS Institute Inc., Cary, NC, USA) running on a mainframe computer.

3.6.2 Multinomial response models

Multinomial response models may be used to model categorical data, e.g. variables that can assume only a limited number of discrete values. Multinomial response models may be used to model either a single dichotomous target variable, or a single polytomous variable or several dichotomous or polytomous target variables. Furthermore, one, or more dichotomous or polytomous predictor variables may be introduced as covariates. One, or more of these predictor variables may be continuous. The large number of possible combinations of variables involved in modeling generates a whole family of multinomial response models which are all part of the generalized linear model family (HABERMAN (1978).

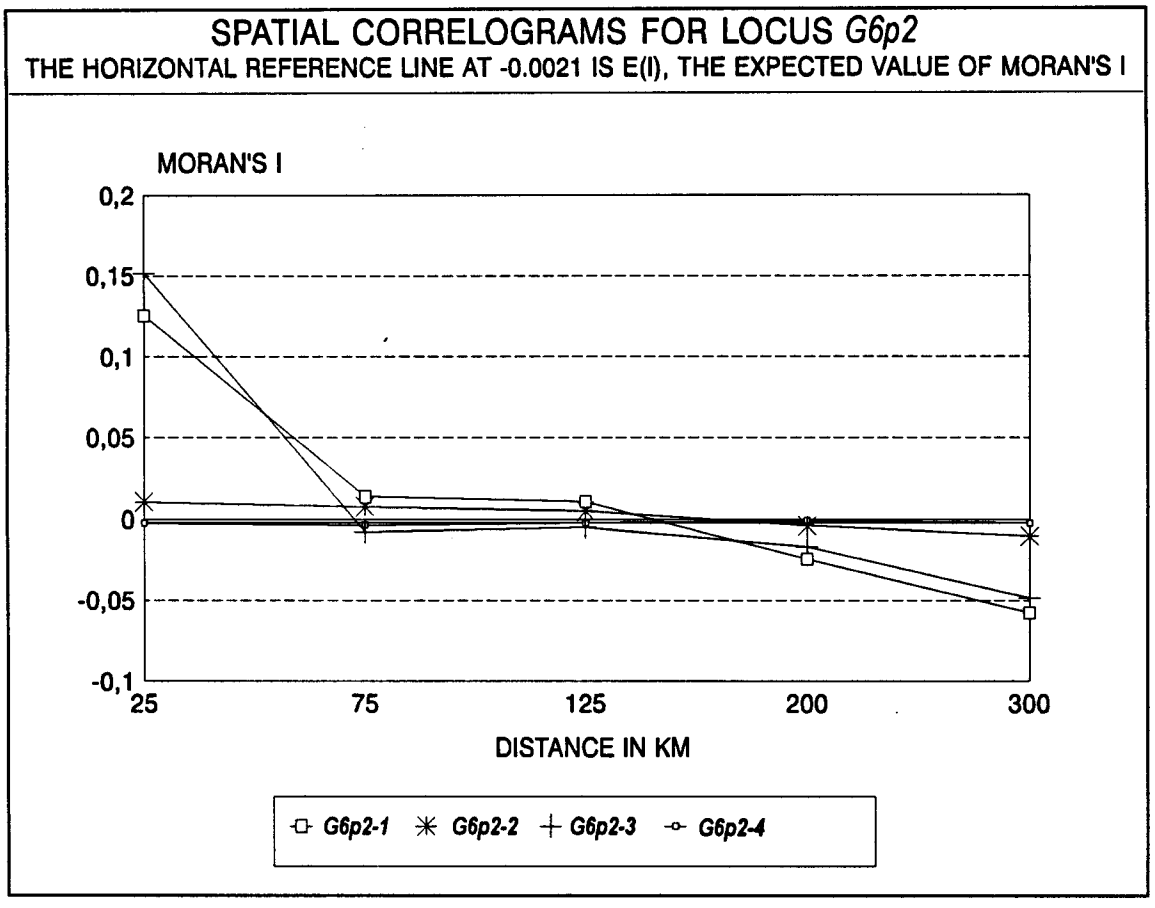
Generalized linear models as derived from NELDER and WEDDERBURN (1972) may be expressed as

$Y_i = \mu_i + \varepsilon_i$ where

Y_i is a response variable derived from the exponential family of probability distributions;

μ_i is the expected value (mean) of Y_i, *i.e.* $\mu_i = E(Y_i)$;

ε_i is a randomly distributed error;



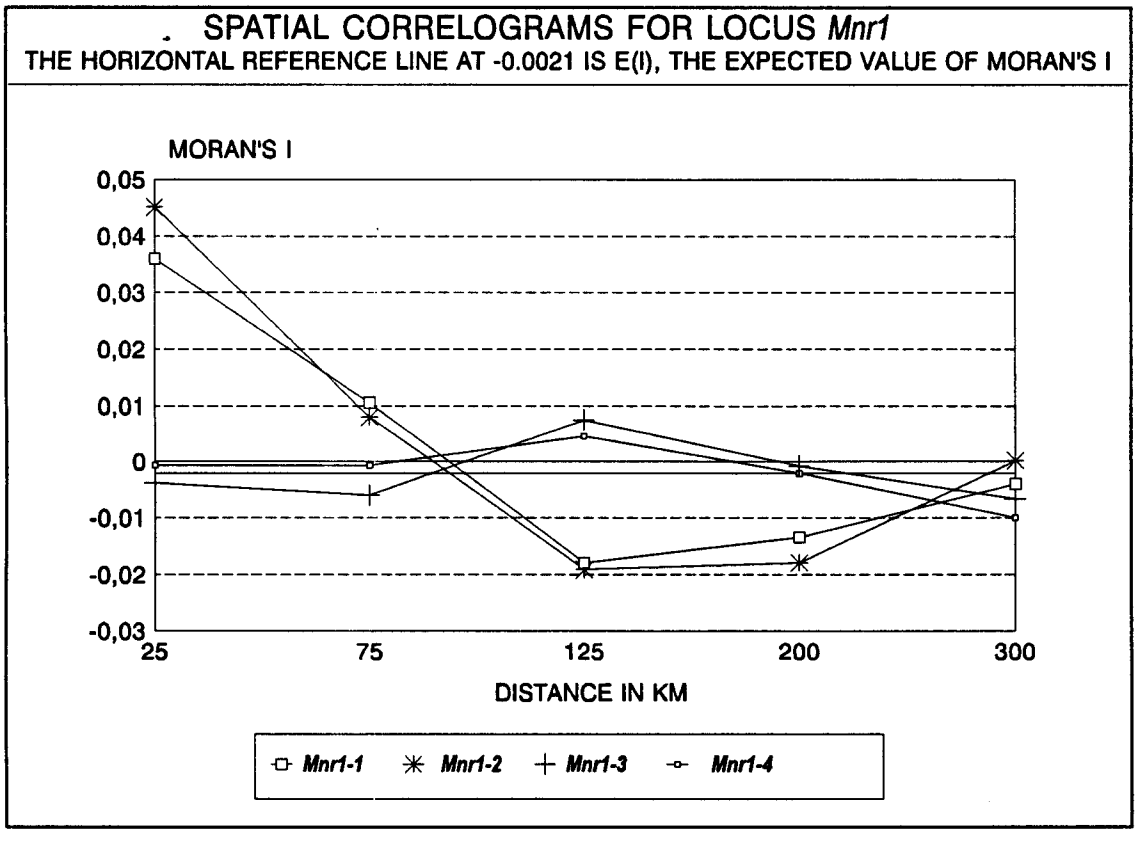


***Figure 9**: Examples of spatial correlograms: Locus G6p2 as an example for a clinal pattern and Mnr1 for a depression and a crazy quilt pattern*

The effect of other variables on the variability of Y_i is expressed as a "linear predictor" (η_i) of the form

$\eta_i = \sum_k \beta_k X_{ik}$ where X_{ik} are known explanatory variables;

β_k are parameters which are estimated from the data.

This linear predictor is related to the expected value of Y_i by a "link function" (g) of the form

$$\eta_i = g(\mu_i)$$

The most important link functions used are logit, cumulative logit, probit and complementary log-log functions. Excellent overviews of generalized linear models and the different multinomial response models (logistic regression, logit regression, probit regression, log-linear models) are given by MCCULLAGH and NELDER (1983) and WRIGLEY (1985).

Multinomial response models were used to study associations between genotypes and climate variables. Genotypes of sample trees were used as response variables, climate principal components as explanatory variables. The genetic response variables were polytomous with three possible values: a value of 0 for genotypes with no copies of a specified allele, of 0.5 for individuals with one copy, and of 1 for an individual homozygous for the allele in question. Explanatory variables were continuous.

In the polytomous case with three possible response categories, three distinct sets of odds for choosing one particular response category can be defined as

$P_{1/i} / P_{3/i}$, $P_{2/i} / P_{3/i}$ and $P_{1/i} / P_{2/i}$ with $P_{1/i}+P_{2/i}+P_{3/i} = 1$

which can be expressed in two functions (one of three functions is redundant) of an explanatory variable X_i based on the general form for multiple response categories

$$L_{rR/i} = \log_e \frac{P_{r/i}}{P_{R/i}} = \alpha_r + \beta_r X_r \qquad r=1, 2, 3, \ldots R\text{-}1$$

for the case of three response categories as the two functions:

$$L_{13/i} \log_e \frac{P_{1/i}}{P_{3/i}} = \alpha_1 + \beta_1 X_i$$

$$L_{23/i} = \log_e \frac{P_{2/i}}{P_{3/i}} = \alpha_2 + \beta_2 X_i$$

For several explanatory variables the logit equations can be written as

$$L_{rR/i} = \log_e \frac{P_{r/i}}{P_{R/i}} = \chi'_{ri} \beta_r$$

The logistic model (which is simply a re-expression of the linear logit model) takes the general form

$$P_{r/i} = \frac{e^{\alpha_r+\beta_r X_i}}{\sum_{s=1}^{R} e^{\alpha_s+\beta_s X_i}} \qquad r=1,2,3, \ldots R$$

with the three equations for the case of three response categories

$$P_{1/i} = \frac{e^{\alpha_1+\beta_1 X_i}}{1 + e^{\alpha_1+\beta_1 X_i} + e^{\alpha_2+\beta_2 X_i}}$$

$$P_{2/i} = \frac{e^{\alpha_2+\beta_2 X_i}}{1 + e^{\alpha_1+\beta_1 X_i} + e^{\alpha_2+\beta_2 X_i}}$$

$$P_{3/i} = \frac{1}{1 + e^{\alpha_1+\beta_1 X_i} + e^{\alpha_2+\beta_2 X_i}}$$

Maximum likelihood estimates are used to estimate the parameters of the linear model in order to maximize the value of the joint likelihood function of the responses. All computations were carried out using the SAS CATMOD procedure (CATegorical data MODeling; © SAS Institute Inc., Cary, NC, USA).

Different models were specified in order to evaluate the probabilities of occurrence of certain genotypes in relation to climate (and climate within region) predictor variables (PCA's).

Model 1: Saturated model with climate *Factor 1*, *Factor 2* and the interaction term *Factor 1* times *Factor 2.*

Model 2: Reduced model with either climate *Factor 1* or *Factor 2*, keeping the more significant predictor variable only.

Model 3: Reduced model with either climate *Factor 1* or *Factor 2*, keeping the more significant predictor variable only and adding a second order term of the one predictor variable. The square of the predictor variable is used to test linearity assumptions (fit of the model) since the likelihood-ratio test is not appropriate for continuous predictor variables.

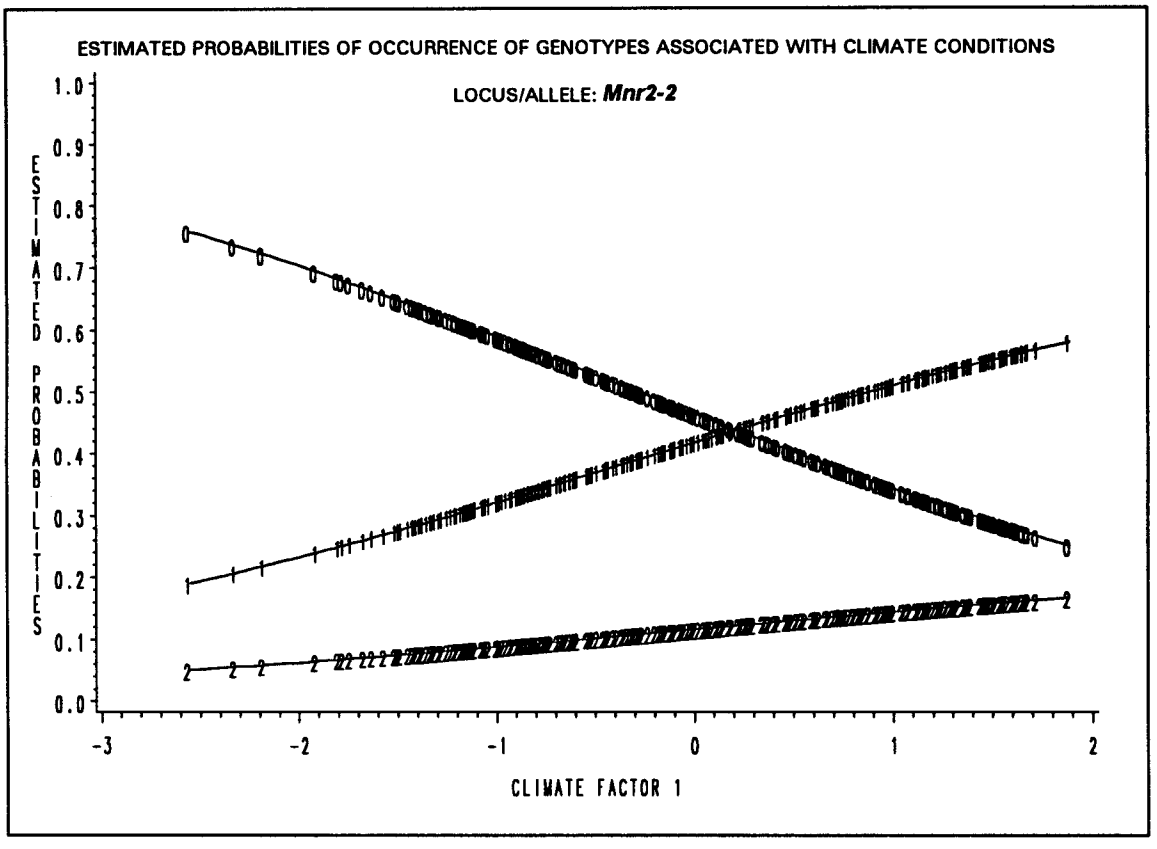
Model 4: Nested model with region and climate *Factor 1* within region as predictor variables in order to separate climate effects from regional (historic) effects.

Model 5: Nested-by-value effect model with region and climate *Factor 1* within each of the two regions separately.

Plots of probabilities versus climate predictor variables were generated for the total sampling area (*for an example see Figure 10, p. 52*) as well as for each region separately and congruence of associations in the two regions were judged from these plots. A direct influence of climate upon genotypes may be reasonably assumed if associations are congruent in the two areas. Since climate is expected to have the same effect on genotypes in both regions, parameters of functions should exhibit the same signs within both regions. Two examples of genotypic frequencies, significantly associated with climate *Factor 1* for the total sampling area, are shown in *Figure 11, p. 53*. For allele *Mnr2-2* a direct association with *Factor 1* may be assumed while *Mnr1-1* shows non-congruent, inverse associations in the two regions. Hence, a climate influence is not likely for *Mnr1-1*; other effects indirectly mimic an association with temperature.

3.6.3 Local regression models

Local regression models provide a non-parametric approach to fitting regression functions, or regression surfaces, to data. The regression surface at x is estimated by local fitting of a linear or quadratic function of the independent variables. For each x in the space of the predictors, it is assumed that in a certain neighborhood of x the regression surface is well approximated by the linear or quadratic function. The size of the neighborhood in the space of the independent variables is an adjustable parameter that determines how local the fitting is. As the neighborhood size increases, the estimate becomes smoother. An important property of local regression



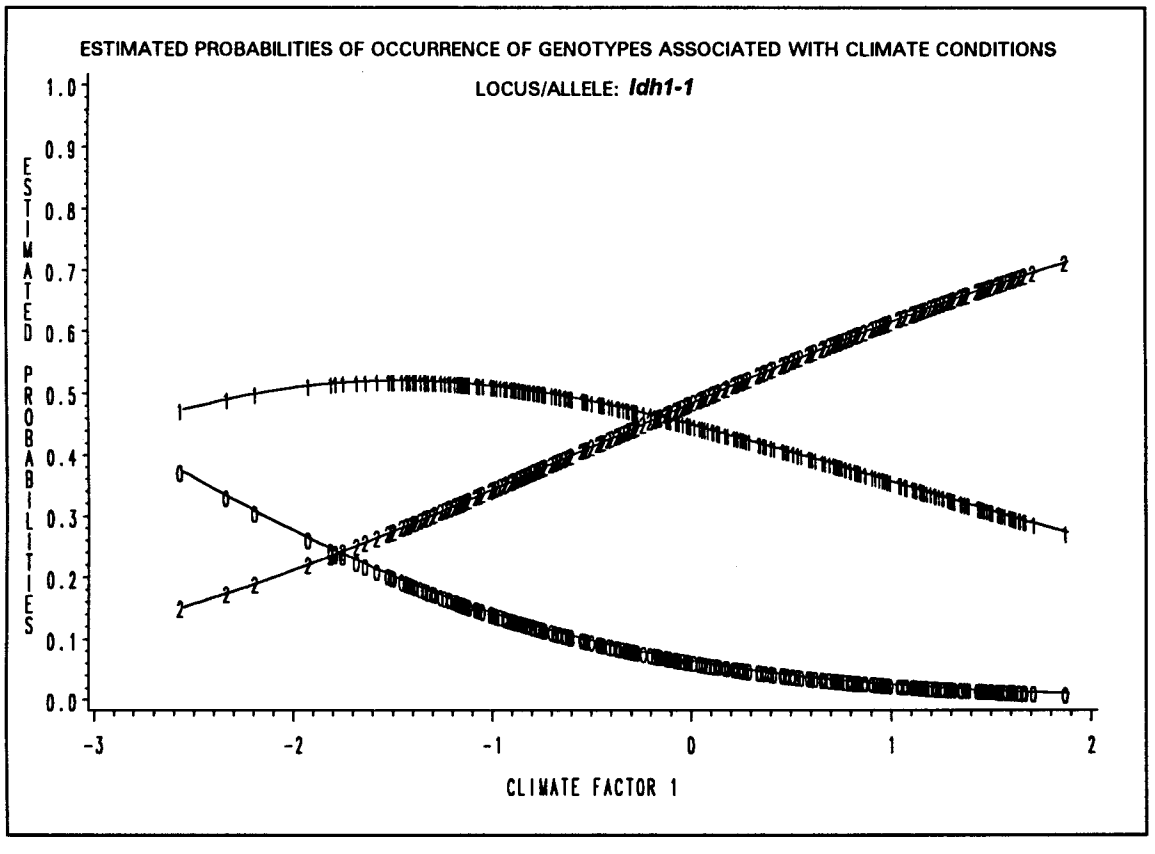
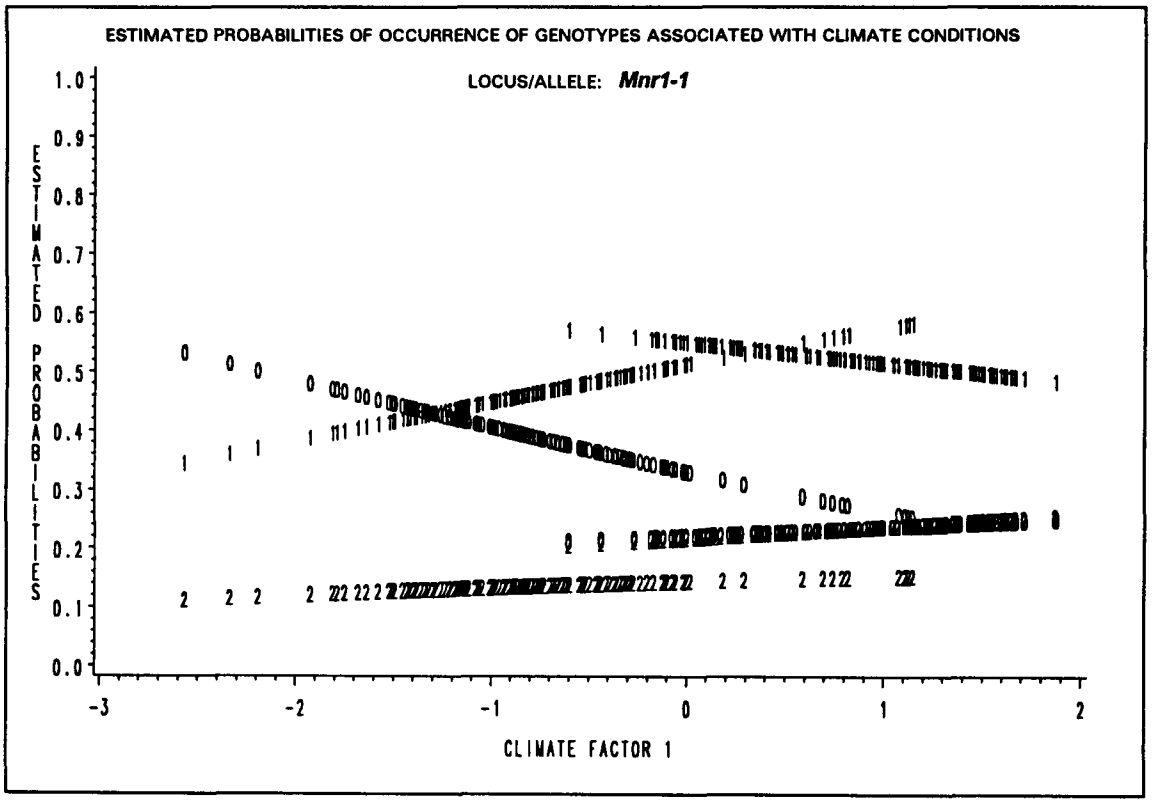


***Figure 10**: Examples of estimated probabilities of occurrence of genotypes associated with climate conditions. Models with climate Factor 1: Total sampling area. Examples of Mnr2-2 and Idh1-1. Legend: 2: 2 copies, 1: 1 copy, 0: no copy of allele*



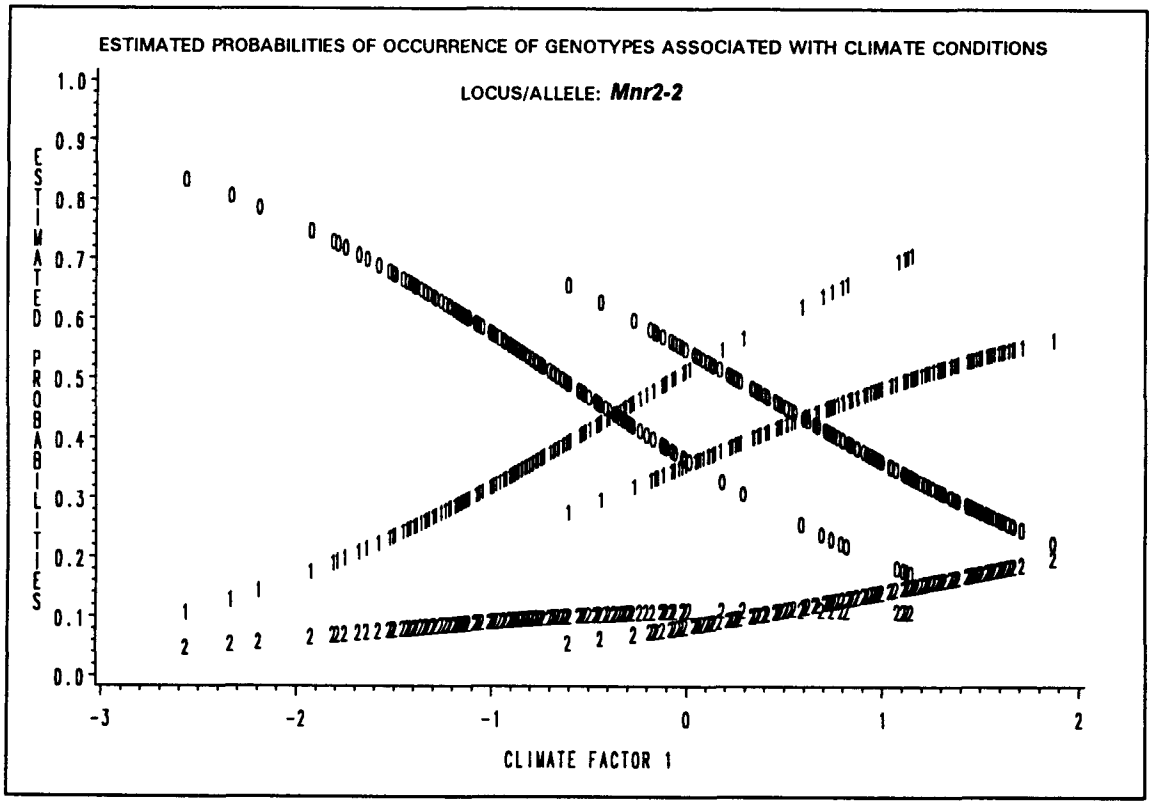


***Figure 11**: Estimated probabilities of occurrence of genotypes associated with climate conditions, shown for the two regions separately. Models with Factor 1 within regions (Central Oregon and Southwest Oregon). Examples of Mnr1-1 and Mnr2-2. Legend: 2: 2 copies, 1: 1 copy, 0: no copy of allele. Central Oregon: Values of Factor 1 between -2.6 and +1.1, Southwest Oregon between -0.6 and +1.9.*

is that the estimate at x is a linear combination of the values of the dependent variables. Consequently, the distributional properties of the least-square fitting of parametric functions are more or less conserved. The advantage of using local regression in place of standard parametric surfaces is that much wider classes of regression surfaces may be estimated without major distortion. An example of the excellent model fit for a complex surface is shown in *Figure 6, p. 40.*

In a local regression model the response and predictors are assumed to be related by

$$y_i = g(x_i) + \varepsilon_i$$

for $i = 1$ to n, where y_i is the i th observation of the response, x_i is the i th observation of p predictors, g is the regression surface, and ε_i is a random error. If x is any point in the space of the predictors, $g(x)$ is the value of the surface at x.

Local regression is conceptually quite simple. A neighborhood of q points whose x_i are closest to x is selected, the points are weighted according to the distance of x from x_i, a line or a quadratic is fit by weighted least-squares, and $g(x)$ is the fitted value at x.

For p independent variables, x_i is a vector of p values and x is a value in the p dimensional space of the independent variables. If η is a distance function in this space and $d(x)$ is the distance of x to the q th nearest x_i, then a set of weights is calculated for the points (x_i, y_i) as

$$w_i(x) = W\left(\frac{\eta(x_i - x)}{d(x)}\right) \qquad \text{with } W \text{ being the } \textit{tricube weight function}$$

A linear or quadratic function of the independent variables is fitted to y_i and evaluated at x to obtain an estimate of $g(x)$. In the linear case the fitting variables are only the independent variables; in the quadratic case the fitting variables are the independent variables, their squares, and their cross-products.

For details on local regression models, for example the possible choices of W, η and neighborhood size, the reader is referred to the detailed overviews by CLEVELAND et al. (1988) and CLEVELAND and GROSSE (1991).

All computations for local regression surfaces were made using S-PLUS statistical software (© StatSci, Statistical Sciences Inc., Seattle, WA, USA) running on a Personal Computer. Local regression fitting using S-PLUS is described in detail by CHAMBERS and HASTIE (1992).

Local regression surfaces were used to model and estimate climate values for tree locations and to model multivariate response surfaces from canonical correlation analyses using sample coordinates as predictors. Moreover, contour maps of seed transfer risks were based on local regression surfaces of response variables from multivariate analyses.

3.6.4 Multivariate statistical procedures

Multivariate statistics refers to an assortment of descriptive or inferential techniques that have been developed to handle situations where "sets" of variables are involved either as predictors or as measures of performance. Multivariate techniques offer several advantages for multidimensional data sets. For example, they provide a way of summarizing redundancies in large data sets, reflect the true multivariate nature of complex data sets, provide a way of handling large data sets with large numbers of variables and offer a means of detecting and quantifying truly multivariate patterns that arise from the correlation structure of a variable set.

3.6.4.1 Principal component analysis

Principal component analysis (PCA) is a statistical technique used to analyze the interrelationships among a large number of variables and to explain those variables in terms of a smaller

set of underlying dimensions (components). The technique involves condensing the information contained in the original variables into a smaller set of uncorrelated dimensions such that there is minimal loss of information. PCA creates weighted linear combinations of the original variables (*i.e.* principal components) which are oriented in directions that describe maximal variation among sample entities. Consequently, principal components represent gradients of maximum variation among entities. PCA summarizes data redundancy *i.e.* identical information or high correlations among variables and generates new, fully uncorrelated variables. PCA may also be used to explore and describe covariance patterns in complex multivariate data sets and to screen out various important variables. Principal components may be derived either from the sample covariance or correlation matrix. A principal component analysis of the covariance matrix gives more weight to variables with larger variances while the analysis of the correlation matrix treats all variables equally. If there is no a priori basis for deciding that one variable is more important than another, PCA of the correlation matrix is the preferred method.

The explanatory power of a principal component is measured by its eigenvalue. An eigenvalue represents the extent of variation among sample entities along the dimension specified by the principal component. The first PCA defines the dimension (or gradient) with the single greatest variance. The second component provides the greatest explanation of sample variance after the first has done its best and so forth. Principal components therefore are uncorrelated with each other. The product-moment correlations between a principal component and the original variables, *i.e.* the factor loadings, indicate how closely the component represents the original variables. These loadings (factor patterns) are used to interpret the component *i.e.* the information it represents. A good principal component solution is characterized by a few principal components extracting most of the original variance, a factor pattern with each variable loading high only on one component, a straightforward and meaningful interpretation of the components, and high final communalities of the original variables. Final communalities are equal to squared multiple correlations for predicting the variable from the principal components. Consequently, final communalities indicate how well the original variables are accounted for by the retained principal components. Principal component solutions may be improved by rotating the components. Rotating the axis normally enhances the interpretation of the components by increasing the loadings of important and decreasing the loadings of unimportant variables. A number of different rotation procedures exist with "varimax" rotation being the most effective under most circumstances (HAIR et al. 1987).

Rules for adequate sample sizes were given by HAIR et al. (1987) and JOHNSON (1981). HAIR suggests a sample size of four to five times the number of variables, whereas JOHNSON postulates a required sample size of twenty plus 3 to 5 times the number of variables.

Each sample entity in the data set has a score for each of the components, derived by multiplying the observed values for each variable (in standardized form) by the corresponding standardized scoring coefficients and summing the products for each component. Principal component scores represent the values of the new uncorrelated variables (components).

Principal component analysis was used to reduce the high redundancies present in the climate data set and in the seedling quantitative data set by extracting a few uncorrelated components for further analyses. SAS FACTOR procedure (SAS Institute Inc., Cary, NC, USA) was used to perform the analyses.

3.6.4.2 Cluster analysis

Cluster analysis refers to a large family of techniques that share a similar goal of organizing entities into discrete classes or groups such that within-group similarity is maximized and among group similarity is minimized. There are many different clustering techniques depending on whether a hierarchical or non-hierarchical, an exclusive or non-exclusive or an agglomerative or divisive method is used (SNEATH and SOKAL, 1973, GAUCH, 1982). Hierarchical clustering techniques combine similar entities into groups and arrange these groups into a hierarchy which helps reveal relationships in the data. Most widely used are agglomerative hierarchical clustering techniques, which begin with each entry in a class of its own and then fuse the classes into larger, hierarchical clusters. Entities are assigned to groups based on distance coefficients among the entities in multivariate variable space. Different distance measures may

be used; the most familiar distance being Euclidean distance. In addition, different fusion strategies may be applied to group the entities into classes on the basis of their respective distances in variable space. Single linkage, complete linkage, centroid linkage, median linkage, average linkage and Ward's minimum variance linkage are the most widely used techniques to agglomerate entities into clusters. Some of these techniques are space conserving *i.e.*, they conserve the distance in the variable space, others are space contracting or space dilating, moving entities closer or further away from the remaining entities. The most widely used fusion strategies are complete linkage and average linkage. Complete linkage is a space dilating strategy which defines the distance of an entity to a cluster to be equal to its distance to the furthest entity in that cluster. The objective of this strategy is to minimize the dissimilarity between neighbors and to produce clusters of very similar samples. Consequently, a complete linkage fusion strategy produces very distinct groups of homogeneous entities which are easily recognized. Average linkage, the most common fusion strategy, on the other hand is space conserving. An entities' dissimilarity to a cluster is defined to be equal to the average of the distances between the entity and each point in the cluster (unweighted pair-group method, UPGMA). Clusters are less distinct and less homogeneous than with complete linkage, but clusters portray the real, unchanged structure of the multidimensional data cloud.

Cluster analysis was used to test for ecotypic structures in genetic data, *i.e.* structures that are organized in patches or groups. Cluster analysis was also applied to test similarities in genetic structure of adjoining seed zones. Furthermore, areas of similar climate conditions were formed based on cluster analysis of climate data and were used in subsequent discriminant analysis. In the same way, groups of similar genotypes were formed based on cluster analysis of scores from 31 loci, and differences in climate among these groups were then described using discriminant analysis. All cluster analyses were performed with the SAS CLUSTER procedure (© SAS Institute Inc., Cary, NC, USA).

3.6.4.3 Discriminant analysis

Discriminant analysis (DA) is a technique used to objectively discriminate among pre-specified groups of sampling entities. Discriminant analysis seeks to find gradients of variation among groups of entities, such that variation among groups is maximized and variation within groups is minimized along these gradients. Discriminant analyses has two main objectives. One objective may be to exhibit optimal separation of groups, based on certain linear transformations of the discriminating variables, and to evaluate which variables are most related to the separation of the groups. This discriminant analysis is called descriptive DA and the associated linear functions are called canonical functions (or variates); it may also be referred to as canonical analysis of discriminance (CAD). The other main objective of DA may be to predict group membership for an entity of unknown origin based on its measured values of the discriminating variables. Such a formulation is referred to as predictive discriminant analysis or as classification. CAD seeks to describe the relationship among two or more groups of entities based on a set of two or more discriminating variables. Specifically, CAD involves deriving linear functions of two or more discriminating variables that will discriminate best among the a priori defined groups. The canonical functions are weighted linear combinations of the original variables, where each variable is weighted according to its ability to discriminate among the groups. The best linear combination of variables is achieved by maximizing the ratio of among-group to within-group variance in canonical scores. In a N by P data set with G groups (N being the sample size, P the number of variables) there are Q (equal to G-1 or P, whichever is smaller) possible canonical functions. Multiple CAD has the added advantage that most of the variation among groups can be explained in fewer than the original Q canonical functions thus dimensionality may be greatly reduced and interpretation of group differences may be facilitated. Classification is based on separately derived classification functions. When group dispersions are equal, linear discriminant functions, using the pooled within-group sums-of-squares and cross products matrix are used to derive the functions, otherwise quadratic classification functions are derived from the individual within-group sums-of-squares and cross products matrix. Alternatively, classification of entities may be based on the Mahalanobis distance in multidimensional space from each entity to the group centroids by classifying each entity into the nearest group. Following WILLIAMS and TITUS (1988), sample size should be at least three times the number of discriminating variables; alternatively, stepwise selection of variables may be used to reduce the number of variables. Eigenvalues in CAD represent the variance of the

corresponding canonical functions; hence, they measure the extent of group differentiation along the dimension specified by the canonical function. The eigenvalue of the first function is always the largest; this function defines the gradient with the single most discriminatory power (maximum among group variance). The second function is orthogonal to the first; it explains the greatest and independent discrimination after the first has done its best. The squared canonical correlation (R^2) is equal to the ratio of among-group to pooled within-group variation. It represents the proportion of total variation in the corresponding function explained by differences in group means *i.e.*, how much of the canonical variation is due to group differences. The product moment correlations between the functions and the original variables (structure coefficients) indicate how closely a variable and a canonical function are related. Thus, the structure coefficients provide a means for interpretation of each function, indicating which variables are most efficient in separating groups.

Each entity has canonical scores on each of the canonical functions. Non-standardized canonical weights multiplied with the raw variables produce standardized canonical scores. In other words, each axis is stretched or shrunk such that the score for an entity represents the number of standard deviations it is from the grand centroid. In this manner, the score for an entity represents the distance from the average score which is zero.

Discriminant analysis and classification were used to describe and evaluate genetic differences among seed zones, to assess genetic differences among trees within areas of similar climate or to describe differences in climate among groups of similar genotypes. Classification was used as a means of evaluating the descriptive canonical functions and to evaluate the practical significance of grouping and classifying groups of genotypes. All analyses were carried out using SAS DISCRIM or STEPDISC procedures (© SAS Institute Inc., Cary, NC, USA).

3.6.4.4 Canonical correlation analysis

Canonical correlation analysis (CCA) is a technique for analyzing the relationship between two or more sets of variables. CCA seeks to find gradients of variation within one set of variables that are maximally correlated with gradients in a second set of variables. Instead of deriving linear combinations of variables within a set (as in PCA), canonical correlation analysis derives two linear functions which explain the maximum correlation between the two sets of variables. As ordination and CAD, canonical correlation attempts to reduce the dimensionality of a multivariate data set by condensing a large number of original variables into a smaller set of new composite dimensions (canonical variates) with a minimum loss of information. CCA examines the relationship between a set of multiple independent variables and another set of multiple dependent variables. Although it is not necessary to make a distinction between dependent and independent variables for canonical analyses to be valid, ecological applications often treat one set of variables as the dependent (response) and the other set as independent (predictor) variables. Canonical correlation constructs linear combinations of variables from each data set, such that the correlation between the two canonical variates is maximized. The canonical variates are defined as weighted linear combinations of the original variables. The canonical correlation coefficient is a measure of the association between pairs of canonical variates in the two data sets. The correlation, however, does not represent the relationship between sets of original variables which makes interpretation somewhat difficult. A strong correlation may be obtained between two canonical variates even though these variates do not extract large (or meaningful) portions of variance from the respective sets of original variables. To overcome this problem, a measure, called redundancy, has been developed, whereby the amount of variance in one set of original variables that is explained by the canonical variate of the other data set can be measured (see later). The squared canonical correlation equals the proportion of variance in the canonical variate of one data domain that is accounted for by the variate of the other data domain. Structure coefficients (loadings) are the simple bivariate correlations between each original variable and the canonical variate from the same data domain; they represent the true relationships among original variables and respective functions. Thus, squared structure coefficients measure the proportion of variance in a variable accounted for by the respective canonical variate of the same data domain. Cross-loadings on the other hand are the simple bivariate correlations between each original variable of one data domain and canonical variates of the opposite data domain. Cross-loadings are obtained by multiplying the canonical structure coefficients by the canonical correlation coefficient. Again, squared canonical cross-loadings

measure the proportion of variance in a variable accounted for by the opposing canonical variates (the variables redundancy). It is tempting to interpret cross-loadings as multiple correlations between variables. However, it must be kept in mind that the relationship is between variables and a composite variate of the opposite domain which was evaluated to maximally correlate with another variate, not with single variables.

Redundancies represent the amount of variance in one variable or in one set of original variables that is explained by a canonical variate from the other data set. The redundancy of a single variable is equivalent to the square of the cross-loading. Redundancy however is mostly used as a summary measure of the amount of variance in one set of variables that is redundant or shared with the variance in the other set of variables. In this sense, redundancy is the proportion of the total variance of a measurement domain predictable from a linear composite of the other domain. The mean sum of squared structure coefficients with a variate measures the variance extracted by the variate. If this quantity is multiplied by the squared canonical correlation coefficient, the proportion of the variance of one domain that is accounted for by the respective variate of the other domain, *i.e.*, the redundancy, is obtained. Hence, in order to have a high redundancy, not only a high canonical correlation coefficient is needed but also a high degree of variance explained by the respective canonical variates.

Canonical correlation analysis was used to analyze and describe relationships between climate variables and isozyme or quantitative trait variables. Scores from canonical analyses were used to model genetic or quantitative response surfaces over the sampling areas as well as to calculate seed transfer risks among existing seed zones and among given locations on the respective response surfaces.

According to R.D. WESTFALL (USFS, Inst. of Forest Genetics, Albany, OR) (personal communication), the redundancy approach, applied to evaluate the amount of variation in genetic data which is explained by some independent variables, has serious limitations. Each gene or allele in the genome has only a small effect on the expression of a phenotypic trait. Single allele redundancies, therefore, are expected to be small. From the additivity of gene expression it follows, however, that even low levels of linkage disequilibium may produce high correlations with some independent variables, if the single small effects are oriented in the same direction in space. In other words, any relationship at an individual locus will be low, whereas the cumulative effects of many loci may be high. The concept of redundancy is unable to portray these specific characteristics of genetic data. Moreover, summarizing structure coefficients from genetic variables (to obtain redundancy) is not valid because the variables, especially at the same locus, are correlated with each other. If the relationships between dependent and independent data are nonlinear, the structure coefficients will be reduced and redundancy will be affected.

Accordingly, an alternative approach for estimating the aggregate variance in the dependent data set that is explained by the independent model was used in the canonical correlation analyses. This approach, suggested by WESTFALL (personal communication), respects the characteristics of genetic data. In a manner similar to regression analysis, the importance of the independent model is evaluated based on the comparison of the model variance to the error variance. The Hotelling-Lawley trace (T_r) is the multivariate equivalent of the ratio of the variance due to the model to the error variance, adjusted for covariances among the variables. It is therefore a measure of the importance of the model. The Hotelling-Lawley trace is calculated as the sum of all eigenvalues; hence, it represents the sum of variances of all variables in the raw covariance matrix. The total variance of the canonical model is therefore given as

$$V_m = \frac{T_r}{1 + T_r}$$ with V_m variance due to the model

T_r Hotelling-Lawley trace

Not all of this model variance however is statistically significant. Accordingly, V_m must be multiplied by the proportion of the trace (p_t) described by the significant vectors

$$V_{ms} = \frac{p_t T_r}{1 + T_r}$$

with V_{ms} proportion of total variance described by the significant variates of the model

P_t proportion of total trace described by the significant variates

SMITH (1981) suggests a required sample size of three times the sum of dependent and independent variables. The problem inherent in using a smaller sample size is that as sample size approaches the number of variables, the canonical correlation approaches 1.00 and statistical significance of the first variate is assured. Biological significance however must be questioned in such cases. According to WESTFALL (personal communication), the expected R^2, when the true correlation is zero, is given as

$$R^2 = \frac{(p+q)}{(N-p-q-1}$$

where p and q are the number of dependent and independent variables, respectively, and N is the sample size.

This expected value will give the lower limit for statistically significant and biologically meaningful correlation coefficients.

Each sample entity in the data set has a score on each canonical variate (location on each canonical axis), derived by multiplying the standardized values of the original variables with the standardized canonical weights and adding the products. Samples in close proximity on the canonical axis occupy the same relative position in the joint space defined by the two sets of variables.

Canonical scores of the significant dependent variates (made up either of isozyme or quantitative variable sets) were used to map the environmentally associated variation over the sampling area. To achieve this, the original canonical scores were related to the geographic variables latitude and longitude using local regression models. Response surfaces as well as contour plots of the variation patterns were established in this way.

Canonical scores were also used to calculate seed transfer risks and to scale contour plots of response surfaces to given transfer-risks of 10, 20 and 30% (*see section 3.6.5, p. 59*).

Partial canonical correlation analysis was used in order to partial out indirect effects which may have contributed to the pattern of variation in question. For our problem, primarily historical events, caused by possible differences in the evolutionary background of the populations in the two regions, may have had an effect on the patterns of variation which we observe today. Partial canonical correlation eliminates such indirect effects by removing known or undesirable variables (by multiple linear regression) before the analysis is performed (*see section 6.3*). All analyses were carried out using the SAS CANCOR procedure (© SAS Institute Inc., Cary, NC, USA).

3.6.5 Estimating seed transfer-risk

The method for estimating seed transfer-risk, used in the present study, was developed and described by CAMPBELL (1986). As already outlined in *section 3.1, p. 25*, mapping genetic variation is based on 3 assumptions: 1) that variation associated with the environment is the result of adaptation, 2) that the map of adaptive genetic variation is also a map of the environmental complex active in selection, and 3) that within-source genetic variability is equal on all sites. Hence, the mean genotypic value, predicted for a certain location (for example expressed as a canonical score from canonical correlation analysis) and the additive genetic variation among individuals at this location (σ_A^2) are assumed to estimate the mean environment and variation in space and time of micro-environmental elements at the respective location. When seed is moved to new locations, the mean of the plantation environment may differ from the mean of the environment of parent tree location and a certain degree of mismatch may occur between genotypes and micro-environment. This mismatch can be estimated from the difference in frequency distributions of genotypes at the seed source and at the plantation site and a

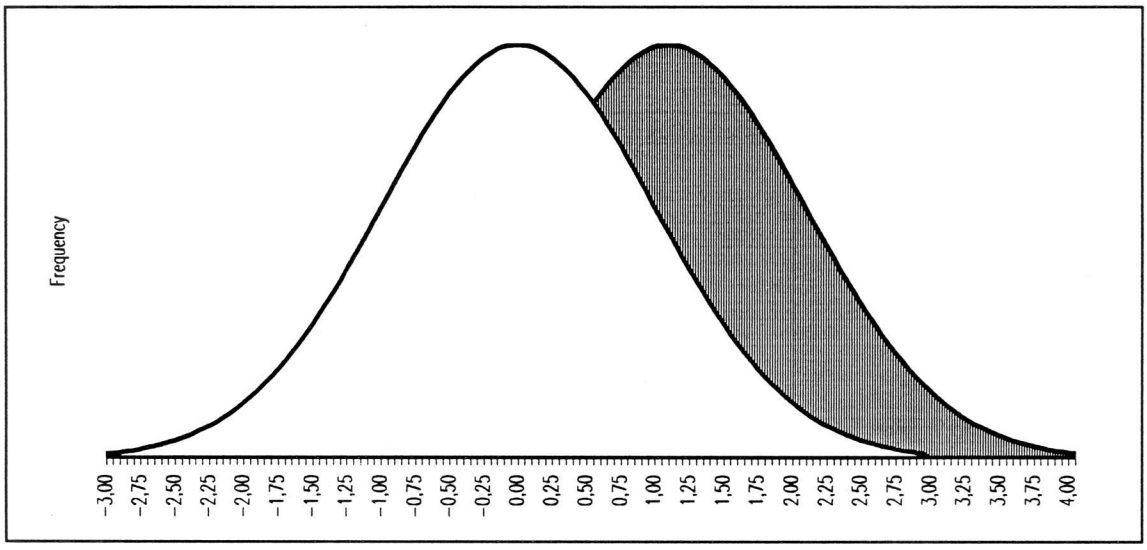


***Figure 12**: Illustration of relative transfer risk (adapted from CAMPBELL, 1986). Hypothetical frequency distributions of genetic scores at source location (left-hand curve) and at plantation site (right-hand curve). Relative seed transfer risk is calculated as the proportion of the right-hand curve not congruent with the curve at the left (hatched area)*

mismatch index can be calculated. Specifically, the mismatch index represents the proportion of the curve of the genotypic lot at a seed origin which does not overlap the curve for a plantation site (*Figure 12*). Since it is assumed, that, for a given seed source, microenvironmental and genotypic frequencies are described by the same curve, the degree of noncongruence represents the frequency of genotypes not matching available environments at the new location. Consequently, the degree of mismatch between genotypes and mircohabitats indicates the relative seed transfer risk (*R*). Under the assumption that the within-source genetic variability is equal for both sites, the proportion of overlap in the two frequency distributions may be calculated as

$$V = 1 - 2\alpha$$

where $\alpha = P\,[0 < z < \frac{0.5x}{\sigma_A}]$ is the area under the standard normal curve from 0 to *z* and

$$z = \frac{x}{2\sigma_A}$$

with

- x difference between mean score at source location and plantation site;
- σ_A standard deviation of the additive genetic variation of scores within sources;
- z standardized score.

If the relative risk of seed transfer (*R*) between the two sites is assumed to be equal to the mismatch index, then transfer risk is given as $R = 1\text{-}V = 2\alpha$. The x corresponding to a given transfer risk *R* is calculated by evaluating *z* at which $\alpha = \frac{R}{2}$, then *z* is calculated as $z = \frac{0.5x}{\sigma_A}$ and x as $x = 2z\sigma_A$

Relative transfer risks were estimated based on canonical scores from canonical correlation analysis relating allozyme data and seedling trait data to climate factors. Relative transfer risks between all pairs of seed zones were calculated using the means and variances of canonical

scores, calculated for each seed zone. For these calculations, parent trees were assumed to be unrelated within seed zones, hence the actual within-zone variance was used as $\sigma_{A(Z)}^2$ for the allozyme scores. For the seedling progenies, however, a genetic correlation of 0.30 among offspring of open-pollinated parents was assumed, and thus 3.33 times the actual within-zone variance was used as $\sigma_{A(Z)}^2$. The genetic correlation of 0.30 reflects the mating among trees within zones which are inbred to a certain degree but unrelated with each other and it assumes an average self fertilization rate of 5% to 10% (SQUILLACE, 1973, table 7, p. 153).

In addition, differences in scores between source location and plantation site (x), corresponding to transfer risks R of 10%, 20% and 30%, were calculated. These distances x, each representing a given transfer risk, were then used to scale distances between contour lines in maps of the canonical response surfaces. Within-source variance was calculated as $\sigma_{f(s)}^2$ (variance of families within sources) for the allozyme scores of parent trees, using analysis of variance procedures taking advantage of all locations with two sample trees per location, and as $3\sigma_{f(s)}^2$ for seedling quantitative traits. The multiplier assumes a genetic correlation of 0.33 among offspring of open-pollinated parents. This correlation of 0.33 reflects the greater likelihood of pollination by adjacent related trees within locations and assumes an average of 5% to 10% self fertilization (SQUILLACE, 1973, Table 6, p. 153). Within-source variance was calculated using analysis of variance procedures taking advantage of all locations with two sample trees per location.

4. Genetic structure - measures of diversity and differentiation

4.1 Univariate measures of diversity and differentiation

The genetic structure of ponderosa pine in the sampled areas of Oregon may be described with standard measures of variation, diversity and differentiation calculated from allozyme data using the total of 488 trees. It must be emphasized, however, that all these measures, which are commonly used to characterize genetic structure of populations or demes, are largely influenced by the sampling design as well as by the selection of the gene loci used in the investigation.

Instead of sampling a certain number of individuals from selected populations, individual trees were sampled on a grid, with a spacing of 8 to 30 kilometers between locations. Samples, therefore, originated from different demes which may be differentiated from each other due to adaptation, selection or random genetic processes. Consequently, variation within demes and differentiation among demes are confounded due to the sampling design. Moreover, calculations of differentiation are limited due to the lack of population samples. Corresponding to the main objective of the study (assessement of associations between multilocus allozyme frequencies and climate conditions), only polymorphic enzyme loci were included in the electrophoretic analysis. Thus, genetic diversity and heterozygosity will likely be overestimated compared to other studies. For these reasons comparisons with other investigations or species are somewhat limited.

Allelic and genotypic structures were analyzed either for the total sampling area or for the two regions separately. Genetic differentiation of ponderosa pine between the east and the west side of the Cascade crest is of special interest because results may shed light on the existence of races. Consequently, the analysis in this section focuses primarily on the comparison of the two regions. Amount and structuring of genetic differentiation within regions, on the other hand, may yield information on adaptation to environments. Since population samples were not available, description of differentiation within regions is limited. Differentiation within regions may be estimated only very crudely on the basis of artificially formed populations, taking all trees of a seed zone as a population. Sample size for these "populations" varied considerably, however, and ranged from 3 to 70 with a mean value of only 13.53 individuals per zone. Moreover, due to the different numbers of seed zones, mean sample size differed between the two regions (23.6 in Southwest Oregon vs. 9.3 in Central Oregon), most likely influencing the within region diversity estimates. Bulking adjacent zones was considered but rejected since it would have had other problems attached to it. Consequently, results regarding diversity and differentiation estimates within regions are most likely associated with large errors; they are at best descriptive, showing only tendencies.

Calculations of the following genetic parameters were carried out using the computer programs BIOSYS-1 (SWOFFORD and SELANDER, 1981), GSED (GILLET, 1994) or especially developed SAS-IML programs (SAS Institute Inc., Cary, NC, USA). Allelic and genotypic structures are described separately, using the following measures:

Allelic multiplicity at a gene locus k is characterized as the number of alleles n_k encountered at locus k; this measure of multiplicity is also called ***observed number of alleles*** (n_a). The average number of alleles per locus is simply the average of the allelic multiplicity over all sampled loci. Multiplicity measures only use numbers of genetic variants, disregarding their frequencies and their system of recombination. ***Genetic diversity*** measures, on the other hand, characterize the heterogeneity of the distribution of genetic variants in a deme. The ***proportion of polymorphic loci P*** is calculated based on allele frequencies. A locus was considered polymorphic if the most common allele had a frequency of less than 0.95. ***Allelic diversity v***, also called the ***effective number of alleles*** n_e, measures allelic evenness. The effective number of alleles n_e will equal the absolute number of alleles n_a when all alleles are at equal frequencies but will be close to 1 when single alleles occur at high frequencies accompanied by other alleles in very low frequencies.

GREGORIUS (1978) and MÜLLER-STARCK and GREGORIUS (1986) define ***Allelic diversity*** v_k of a deme at the k - th gene locus as

$$1 \le v_k = \frac{1}{\sum_{i=1}^{n_k} (p_i^k)^2} \le n_k \qquad \text{if } n_k \text{ alleles occur at locus } k \text{ with frequencies } p_i^k \, (i = 1,\ldots,n_k)$$

Genetic differentiation among two demes X and Y at the k - th locus may be measured as the proportion of alleles not shared and may be expressed as the ***allelic distance*** d_{xy} between the two demes (GREGORIUS, 1974a, 1974b; GREGORIUS and ROBERDS, 1986) in the form

$$0 \le d_{xy} = 0.5 \sum_{i=1}^{n_k} |x_i - y_i| \le 1$$

with x_i and y_i denoting allele frequencies of demes X and Y

Allelic subpopulation differentiation D among several demes may then be expressed as the allelic distance between every deme and its complement *i.e.* the union of all other demes. Hence, the differentiation of the j - th deme at the k - th locus is calculated as

$$D_j = 0.5 \sum_{i=1}^{n_j} \left| p_i^{(j)} - \overline{p_i}^{(j)} \right|$$

where the $p_i^{(j)}$ are the allele frequencies in the j - th deme and the $\overline{p_i}^{(j)}$ are those in the complement. The ***average subpopulation differentiation δ among L demes*** is measured by the weighted mean of the D_j *i.e.*

$$\delta = \sum_{j=1}^{L} c_j D_j$$

where c_j denotes the relative size of the j - th deme such that $\sum_{j=1}^{L} c_j = 1$

Another, less appealing but most widely used measure of differentiation, is NEI's unbiased genetic distance (NEI, 1978) among demes as defined in *section 3.6.1.1.*

Genotypic structures can be characterized by measures of heterozygosity, tests of Hardy-Weinberg structures and WRIGHT's F - statistic. These tests of single locus structures investigate the association of gametes in observed zygotic structures by comparing observed to corresponding expected structures under certain models of association.

The ***degree of heterozygosity*** is defined for an individual with respect to a specified number of gene loci, and is identical to the proportion of loci at which this individual carries two different alleles. The ***average degree of heterozygosity***, also called ***observed heterozygosity H_O***, refers to the distribution of this degree in a collection of individuals (demes). Hence, the observed heterozygosity equals the mean proportion of heterozygous individuals at the single locus. The concept of expected heterozygosity under Hardy-Weinberg conditions (random mating, panmixia) is used as an index of overall polymorphism for comparing species or populations, independent of the mating system. ***Expected heterozygosity H_e***, also called the ***polymorphic index (PI)*** or ***NEI's gene diversity index*** (NEI, 1973) is estimated as

$$H_e = 1 - \sum_i p_i^2$$

or with a bias correction for small sample size according to NEI (1978) as

$$H_e = \frac{2n(1-\sum_i p_i^2)}{2n-1}$$

Allelic and genotypic structure are related to each other by the ***Hardy-Weinberg principle***. After only one generation of random mating (assuming that no other evolutionary forces such as mutation, selection or drift are acting) and regardless of the initial allele frequencies, genotypic frequencies of the *F1* generation may be described by a binomial or multinomial function of the initial allele frequencies. Theoretically, expected proportions of genotypes may be calculated from observed allele frequencies using the Hardy-Weinberg principle. Deviations of the observed genotypic structure from expectation may then be tested using goodness-of-fit tests such as a χ^2- *test* of the form

$$\chi^2 = \frac{\sum_{i=1}^{k}(N_{obs} - N_{exp})^2}{N_{exp}}$$

or similar tests, developed for small sample sizes (LEVENE, 1949; HALDANE, 1954, ELSTON and FORTHOFER, 1977).

Deviations from Hardy-Weinberg-expectations may be characterized by the ***fixation index F***, calculated as

$$F = 1 - \frac{H_o}{H_e} \quad or \quad 1 - \frac{H_o}{2pq}$$

with p and q being the allele frequencies

A positive value of F indicates a deficiency of heterozygotes while a negative value indicates an excess.

WRIGHT's (1965, 1978) method of ***F - statistics*** may be used to describe the apportionment of variation to the different levels within and among demes. In the same way that the inbreeding coefficient measures the reduction of heterozygosity due to inbreeding within demes, F -statistics may be used to estimate differentiation among demes by comparing the reduction in expected heterozygosity H_e within demes to what would be obtained were the group of demes a panmictic unit. WRIGHT's approach consists of three different F - coefficients used to allocate the genetic variability to three hierarchical levels *i.e.* the total population (F_{IT}), the subpopulations (F_{ST}), and individuals (F_{IS}). These three coefficients are interrelated so that

$$F_{ST} = \frac{F_{IT} - F_{IS}}{1 - F_{IS}}$$

F_{ST} is a measure of the genetic differentiation of sub-populations and is always positive. WRIGHT's F_{ST} is identical to G_{ST} as proposed by NEI (1973). Both measures, however, posses several shortcomings when compared with the subpopulation differentiation δ by GREGORIUS (1974). As has been noticed by WRIGHT himself and shown by GREGORIUS and ROBERDS (1986), F_{ST} assumes its maximum value only if all demes are monomorphic. Moreover, F_{ST} cannot be interpreted in terms of proportions of shared alleles and thus is less appealing than δ. F_{ST} and G_{ST} are not directly comparable to absolute measures of genetic differentiation such as δ. F_{ST} measures the proportion of total variation due to differences among subpopulations and its magnitude is thus very dependent on total levels of genetic diversity present. Mean subpopulations differentiation D could be identical between 2 species, but F_{ST} differ greatly depending on the total magnitude of genetic diversity.

F_{IS} and F_{IT} are measures of the deviation from Hardy-Weinberg expectations within subpopulations and within the total population, respectively. Positive values indicate a deficiency of heterozygotes while negative values indicate an excess.

4.2 Allelic structure

Allele frequencies, the observed number of alleles and the effective number of alleles for the total sampling area (both regions combined) are given in *Table 9*. The mean number of alleles per locus was 3.29 with a range between 2 and 5 over loci, while the mean effective number of alleles was 1.478 with a range between 1.028 and 3.102. The considerable difference of 1.812 between N_a and N_e implies uneven frequency distributions of alleles at many loci. Of the 31 loci examined, 83.87% were polymorphic at the 0.95 criterion. Allele frequency distributions differed greatly at the different loci. Distributions where one allele predominates and all others occur at low frequencies (< 10 %), so called "minor polymorphisms" (LEWONTIN, 1985), were observed at 17 loci. The uneven allele distribution at these loci was reflected in small values of N_e close to 1. Rare alleles with frequencies below 5% were found at 24 loci while extremely rare alleles with frequencies below 1% were observed at 11 loci. "Major polymorphisms" having two or more variants per locus at moderate to medium frequencies (> 10%) were found at 14

LOCUS	ALLELE FREQUENCIES OF ALLELES NUMBER					NUMBER OF ALLELES		HETEROZYGOSITY	
ALLELE	1	2	3	4	5	Na	Ne	Ho	He
MNR-1	0.440	0.497	0.011	0.052	-	4	2.255	0.547	0.558
MNR-2	0.411	0.325	0.264	-	-	3	2.905	0.600	0.656
LAP-2	0.954	0.012	0.019	0.014	-	4	1.098	0.088	0.089
LAP-3	0.847	0.098	0.054	-	-	3	1.370	0.270	0.270
PEP-1	0.943	0.046	0.011	-	-	3	1.122	0.115	0.109
PEP-2	0.931	0.048	0.020	-	-	3	1.150	0.127	0.130
PEP-3	0.830	0.083	0.087	-	-	3	1.422	0.318	0.297
PEP-4	0.959	0.041	-	-	-	2	1.085	0.074	0.079
MPI-1	0.817	0.026	0.158	-	-	3	1.443	0.328	0.308
MPI-2	0.833	0.054	0.098	0.014	-	4	1.262	0.328	0.294
GOT-1	0.892	0.012	0.089	0.006	-	4	1.249	0.191	0.196
GOT-2	0.960	0.039	0.001	-	-	3	1.083	0.080	0.077
GOT-3	0.907	0.053	0.019	0.011	0.009	5	1.210	0.176	0.175
G6P-2	0.950	0.018	0.031	0.001	-	4	1.106	0.061	0.097
ACP-1	0.790	0.159	0.017	0.025	0.009	5	1.537	0.336	0.350
GDH-1	0.942	0.056	0.002	-	-	3	1.123	0.072	0.110
IDH-1	0.720	0.298	-	-	-	2	1.719	0.424	0.419
PGM-1	0.838	0.024	0.001	0.137	-	4	1.385	0.277	0.278
ACO-1	0.346	0.273	0.025	0.357	-	4	3.102	0.648	0.679
SKD-1	0.860	0.126	0.014	-	-	3	1.323	0.232	0.245
SKD-2	0.656	0.344	-	-	-	2	1.823	0.410	0.452
FDP-2	0.956	0.044	-	-	-	2	1.092	0.088	0.084
UGP-1	0.793	0.077	0.130	-	-	3	1.534	0.385	0.349
UGP-2	0.630	0.232	0.133	0.005	-	4	2.135	0.641	0.532
UGP-3	0.879	0.121	-	-	-	2	1.270	0.234	0.213
FUM-2	0.762	0.223	0.014	-	-	3	1.586	0.416	0.369
ADH-2	0.647	0.350	0.003	-	-	3	1.848	0.430	0.460
PGI-2	0.920	0.037	0.033	0.010	-	4	1.178	0.152	0.151
MDH-1	0.986	0.011	0.003	-	-	3	1.028	0.027	0.028
MDH-3	0.883	0.074	0.022	0.022	-	4	1.272	0.189	0.214
MDH-4	0.947	0.002	0.051	-	-	3	1.112	0.098	0.101
MEAN						3.29	1.478	0.270	0.270

Table 9: *Estimated allele-frequencies, observed and effective number of alleles, and observed and expected heterozygosity over the entire sampling area (N = 488). [Na: observed number of alleles, Ne: effective number of alleles (allelic diversity), He: unbiased expected heterozygosity (NEI's diversity), Ho: observed heterozygosity (direct count)*

loci *i.e.* at *Mnr-1, Mnr-2, Mpi-1, Acp-1, Idh -1, Pgm-1, Aco-1, Skd-1, Skd-2, Ugp-1, Ugp-2, Ugp-3, Fum-2* and *Adh-2*. Effective number of alleles for these loci commonly was above 1.5 and was as high as 3.1 for *Aco-1* which had 3 of its 4 alleles at roughly equal frequencies.

Table 10 shows a comparison of allelic diversity between the two regions, Central Oregon and Southwest Oregon. Southwest Oregon had a slightly higher mean number of alleles per locus and showed 5 private rare alleles not found in the Central Oregon sample *i.e.* at *Got-2, G6p-2, Gdh-1, Adh-2 and Mdh-1,* while the Central Oregon sample only had 2 region specific alleles (*Pgm-1* and *Mdh-4*). 87.1% of the loci were polymorphic in Southwest Oregon compared to 74.2% in Central Oregon (*Table 11, p. 69*). Although ponderosa pine in Southwest Oregon had a slightly higher proportion of polymorphic loci and a higher allelic diversity at 18 loci (with 16 expected by chance alone), genetic diversity was not significantly different in the two regions.

When pooled over all 31 loci, allelic frequencies were significantly heterogeneous ($p < 0.0001$) between the two regions (*Table 11, p. 69*). This heterogeneity was also reflected in an overall average subpopulation differentiation of $\delta = 0.061$ which may be interpreted as an average difference in allele frequencies of 6.1 %. Highly significant differences ($p < 0.001$) were found at

LOCUS	OBSERVED NUMBER OF ALLELES [Na]		EFFECTIVE NUMBER OF ALLELES [Ne]			OBSERVED HETEROZYGOSITY [Ho]		
	SOUTHWEST	CENTRAL	SOUTHWEST	CENTRAL	DIFFERENCE	SOUTHWEST	CENTRAL	DIFFERENCE
MNR-1	4	4	2.258	2.201	0.057	0.549	0.544	0.005
MNR-2	3	3	2.955	2.808	0.147	0.643	0.560	0.083
LAP-2	4	4	1.169	1.037	0.132	0.145	0.036	0.109
LAP-3	3	3	1.344	1.398	-0.054	0.243	0.298	-0.055
PEP-1	3	3	1.172	1.078	0.094	0.157	0.075	0.082
PEP-2	3	3	1.168	1.132	0.036	0.132	0.123	0.009
PEP-3	3	3	1.327	1.516	-0.189	0.268	0.365	-0.097
PEP-4	2	2	1.102	1.070	0.032	0.081	0.067	0.014
MPI-1	3	3	1.408	1.480	-0.072	0.328	0.329	-0.001
MPI-2	4	4	1.364	1.471	-0.107	0.298	0.353	-0.055
GOT-1	4	4	1.264	1.223	0.041	0.183	0.198	-0.015
GOT-2	3	2	1.084	1.083	0.001	0.077	0.079	-0.002
GOT-3	5	5	1.254	1.172	0.082	0.213	0.143	0.070
G6P-2	4	3	1.204	1.024	0.180	0.102	0.024	0.078
ACP-1	5	5	1.652	1.430	0.222	0.362	0.313	0.049
GDH-1	3	2	1.245	1.020	0.225	0.128	0.020	0.108
IDH-1	2	2	1.441	1.921	-0.480	0.328	0.516	-0.188
PGM-1	3	4	1.252	1.522	-0.270	0.209	0.341	-0.132
ACO-1	4	4	2.992	3.145	-0.153	0.660	0.639	0.021
SKD-1	3	3	1.270	1.366	-0.096	0.204	0.258	-0.054
SKD-2	2	2	1.589	1.966	-0.377	0.323	0.492	-0.169
FDP-2	2	2	1.021	1.157	-0.136	0.021	0.151	-0.130
UGP-1	3	3	1.470	1.590	-0.120	0.362	0.409	-0.047
UGP-2	4	4	2.224	2.043	0.181	0.698	0.591	0.107
UGP-3	2	2	1.285	1.256	0.029	0.247	0.222	0.025
FUM-2	3	3	1.623	1.549	0.074	0.438	0.397	0.041
ADH-2	3	2	1.717	1.943	-0.226	0.370	0.488	-0.118
PGI-2	4	4	1.233	1.124	0.109	0.196	0.111	0.085
MDH-1	3	2	1.026	1.032	-0.006	0.021	0.032	-0.011
MDH-3	4	4	1.329	1.221	0.108	0.230	0.151	0.079
MDH-4	2	3	1.135	1.087	0.048	0.111	0.087	0.024
MEAN	3.20	3.10	1.470	1.480		0.268	0.271	

Table 10: *Observed and expected number of alleles and observed heterozygosity by regions. [Na: observed number of alleles, Ne: effective number of alleles (allelic diversity v), Ho: observed heterozygosity (direct count)]. Sample size: Southwest N = 236, Central Oregon N = 252*

LOCUS ALELE	SOUTHWEST FREQUNCIES	CASCADES FREQUENCIES	F(IS)	F(IT)	F(ST)	HETEROGENEITY χ^2	HETEROGENEITY PROB	DIFFERENTIATION GREGORIUS (1986)
MNR-1								
1	0.494	0.389	0.008	0.018	0.010	11.930	0.0076	0.110
2	0.445	0.548					**	
3	0.009	0.014						
4	0.053	0.050						
MNR-2								
1	0.370	0.450	0.080	0.084	0.004	6.856	0.0324	0.082
2	0.355	0.296					*	
3	0.274	0.254						
LAP-2								
1	0.923	0.982	0.001	0.014	0.013	21.042	0.0001	0.058
2	0.021	0.004					***	
3	0.028	0.012						
4	0.028	0.002						
LAP-3								
1	0.855	0.839	-0.003	-0.002	0.000	0.487	0.7840	0.017
2	0.094	0.103					n.s.	
3	0.051	0.058						
PEP-1								
1	0.921	0.962	-0.057	-0.051	0.006	10.936	0.0042	0.041
2	0.057	0.036					**	
3	0.021	0.002						
PEP-2								
1	0.923	0.938	0.023	0.024	0.001	1.313	0.5186	0.015
2	0.051	0.046					n.s.	
3	0.026	0.016						
PEP-3								
1	0.862	0.800	-0.077	-0.071	0.006	9.789	0.0074	0.063
2	0.055	0.109					**	
3	0.083	0.091						
PEP-4								
1	0.951	0.966	0.063	0.064	0.001	1.428	0.2321	0.015
2	0.049	0.034					n.s.	
MPI-1								
1	0.828	0.806	-0.069	-0.068	0.001	2.836	0.2422	0.023
2	0.017	0.034					n.s.	
3	0.155	0.161						
MPI-2								
1	0.851	0.817	-0.119	-0.117	0.001	5.008	0.1711	0.032
2	0.049	0.060					n.s.	
3	0.094	0.101						
4	0.006	0.022						
GOT-1								
1	0.883	0.901	0.027	0.028	0.002	3.959	0.2659	0.028
2	0.009	0.016					n.s.	
3	0.104	0.075						
4	0.004	0.008						

Table 11: *Allele frequencies for each of two regions of ponderosa pine in Oregon, F - statistic, test of allele frequency heterogeneity (χ^2 test according to WORKMAN and NISWANDER (1970), and average subpopulation differentiation δ (GREGORIUS and ROBERDS (1986)) among the two regions*

Table 11 (continued)

LOCUS ALELE	SOUTHWEST FREQUNCIES	CASCADES FREQUENCIES	F(IS)	F(IT)	F(ST)	HETEROGENEITY χ^2	HETEROGENEITY PROB	DIFFERENTIATION GREGORIUS (1986)
GOT-2								
1	0.962	0.960	-0.04	-0.03	0.000	1.152	0.5622	1.468	0.4800	0.002
2	0.036	0.040					n.s.		n.s.	
3	0.002	0.000								
GOT-3										
1	0.889	0.923	-0.01	-0.01	0.004	7.390	0.1166	7.407	0.1159	0.040
2	0.072	0.036					n.s.		n.s.	
3	0.021	0.018								
4	0.011	0.012								
5	0.006	0.012								
G6P-2										
1	0.909	0.988	0.350	0.366	0.025	33.392	0.0000	38.255	0.0000	0.079
2	0.030	0.008					***		***	
3	0.060	0.004								
4	0.002	0.000								
ACP-1										
1	0.753	0.827	0.022	0.032	0.010	20.866	0.0003	18.683	0.0009	0.097
2	0.206	0.115					***		***	
3	0.006	0.024								
4	0.021	0.028								
5	0.013	0.006								
GDH-1										
1	0.889	0.990	0.321	0.350	0.043	44.841	0.0000	51.259	0.0000	0.100
2	0.106	0.010					***		***	
3	0.004	0.000								
IDH-1										
1	0.811	0.599	-0.07	-0.01	0.054	51.894	0.0000	53.606	0.0000	0.212
2	0.189	0.401					***		***	
PGM-1										
1	0.887	0.792	-0.00	0.005	0.013	18.357	0.0003	19.551	0.0002	0.096
2	0.011	0.036					***		***	
3	0.000	0.002								
4	0.102	0.171								
ACO-1										
1	0.309	0.383	0.037	0.042	0.006	14.101	0.0027	14.768	0.0020	0.101
2	0.270	0.276					**		**	
3	0.015	0.034								
4	0.406	0.308								
SKD-1										
1	0.881	0.839	0.050	0.054	0.004	5.129	0.0769	5.271	0.0717	0.047
2	0.102	0.149					n.s.		n.s.	
3	0.017	0.012								
SKD-2										
1	0.753	0.563	0.056	0.094	0.040	38.731	0.0000	39.793	0.0000	0.191
2	0.247	0.437					***		***	
FDP-2										
1	0.989	0.925	-0.07	-0.04	0.025	24.169	0.0000	27.650	0.0000	0.065
2	0.011	0.075					***		***	

Table 11 (continued)

LOCUS ALELE	SOUTHWEST FREQUNCIES	CASCADES FREQUENCIES	F(IS)	F(IT)	F(ST)	HETEROGENEITY χ^2	HETEROGENEITY PROB	DIFFERENTIATION GREGORIUS (1986)
UGP-1								
1	0.813	0.774	-0.109	-0.107	0.002	2.254	0.3239	0.040
2	0.070	0.083					n.s.	
3	0.117	0.143						
UGP-2								
1	0.606	0.651	-0.211	-0.210	0.001	2.222	0.5275	0.043
2	0.247	0.218					n.s.	
3	0.140	0.127						
4	0.006	0.004						
UGP-3								
1	0.872	0.885	-0.100	-0.100	0.000	0.362	0.5476	0.012
2	0.128	0.115					n.s.	
FUM-2								
1	0.749	0.774	-0.130	-0.129	0.001	1.904	0.3858	0.024
2	0.232	0.216					n.s.	
3	0.019	0.010						
ADH-2								
1	0.706	0.589	0.049	0.064	0.016	18.942	0.0000	0.125
2	0.287	0.441					***	
3	0.006	0.000						
PGI-2								
1	0.898	0.940	-0.011	-0.006	0.005	9.954	0.0189	0.042
2	0.040	0.034					*	
3	0.051	0.016						
4	0.011	0.010						
MDH-1								
1	0.987	0.984	0.062	0.063	0.001	5.159	0.0758	0.010
2	0.006	0.016					n.s.	
3	0.006	0.000						
MDH-3								
1	0.866	0.903	0.099	0.101	0.003	5.683	0.1280	0.045
2	0.085	0.060					n.s.	
3	0.030	0.014						
4	0.019	0.024						
MDH-4								
1	0.936	0.956	0.026	0.029	0.002	4.732	0.0938	0.024
2	0.000	0.004					n.s.	
3	0.064	0.040						
TOTAL						386.81	0.0000	
MEAN			-0.011	-0.001	0.010		***	0.061
A (SE)	3.2 (0.20)	3.1 (0.20)						
P	87.1	74.2						
He (SE)	0.269 (0.030)	0.266 (0.035)						
Ho (SE)	0.268 (0.032)	0.271 (0.034)						

Subpopulations for F(ST): 2 regions [Central - Southwest]
Tests of heterogeneity of allele-frequencies: Chi square test according to WORKMANN and NISWANDER (1970)
Subpopulation differentiation [2 regions] according to GREGORIUS and ROBERDS (1986)

9 loci (*Lap-2, G6p-2, Acp-1, Gdh-1, Idh-1, Pgm-1, Skd-2, Fdp-2, Adh-2*), while an additional 4 loci were heterogeneous at $p < 0.01$ (*Mnr-1, Pep-1, Pep-3, Aco-1*) and 3 loci at $p < 0.05$ (*Mnr-2, Pgi-2, Mdh-1*) (*Table 11, p. 67 to 69*). Regional differentiation (δ) ranged from 1% to 21.2% and was especially striking for *Idh-1* (21.2), *Skd-2* (19.1), *Adh-2* (12.5), *Mnr-1* (11.0), *Aco-1* (10.1), *Gdh-1* (10.0), *Acp-1* (9.7), *Pgm-1* (9.6) and *Mnr-2* (8.2).

F - statistics, on the other hand, indicated a much smaller differentiation between the two regions with an overall F_{ST} (G_{ST}) - value of 0.01 and a range of 0.000 to 0.054 among the 31 loci (*Table 11, p. 69*). Relatively high F_{ST} - values were manifest for the loci *Idh-1* (0.054), *Gdh-1* (0.043), *Skd-2* (0.040), *G6p-2* (0.025), *Fdp-2* (0.025), *Adh-2* (0.016), *Lap-2* (0.013) and *Pgm-1* (0.013).

Variation in allelic diversity among seed zones within the two regions was, as expected, largely influenced by sample size (number of individuals per zone) which varied between 3 and 70. Therefore, results of diversity within regions are given for completedness only. The mean number of alleles per locus within seed zones was 2.205 with a range of 1.7 to 3. N_a rather reflected sample size than an actual variation in multiplicity; N_a was highly correlated with sample size *N* ($R^2 = 0.73$). The proportion of polymorphic loci was less influenced by sample size ($R^2 = 0.21$). On average, 68.27% of the loci were polymorphic in each seed zone with a range between 51.5% and 96.8%.

According to *F* - statistics, overall heterogeneity in allele frequencies among seed zones did not differ significantly within the two regions. Eight percent of the total variation was due to seed zone differences within Southwest Oregon compared to 7% within Central Oregon (*Table 12, p. 71*). In Southwest Oregon, allelic frequencies were significantly heterogeneous ($p < 0.001$) among the seed zones over all 31 loci while frequency distribution did not significantly differ among zones in Central Oregon. Highly significant ($p < 0.001$) heterogeneities in allele frequencies among zones in Southwest Oregon were manifest for *Got-1*, *G6p-2*, *Acp-1*, *Gdh-1* and *Skd-1* while two more loci (*Mnr-2* and *Got-3*) were significantly heterogeneous at $p < 0.05$. F_{ST} values for these loci ranged widely between 0.060 for *Mnr-2* and 0.443 for locus *Gdh-1*. Differentiation among seed zones within Central Oregon was less pronounced than within Southwest Oregon, as regards the overall F_{ST} - values of 0.070 as well as the number of loci showing significant heterogeneities in frequency distributions among zones. F_{ST} - values for Central Oregon were all below 0.01. Only three loci exhibited significant differences in frequency distributions among zones *i.e. Lap-2* ($p = 0.002$), *Pep-2* ($p = 0.011$) and *Aco-1* ($p = 0.011$). Since mean sample size per zone was much smaller in Central Oregon (9.3) than in Southwest Oregon (23.6), the smaller differentiation observed in Central Oregon may, however, just be a reflection of a lower precision of parameter estimates rather than a real difference.

Allelic subpopulation differentiation *D* among seed zones and average subpopulation differentiation δ within regions (GREGORIUS and ROBERDS, 1986) are given in *Table 13, p. 72 and 73*. Average subpopulation differentiation among all 39 sampled seed zones and over all 31 loci was 8.8%. Differentiation varied greatly among the loci, however, ranging from 1% to 12.6%. Overall high differentiation was manifest at *Mnr-1* (12.6), *Mnr-2* (12.6), *Aco-1* (12.3), *Skd-2* (11.8), *Idh-1* (11.7), *Adh-2* (9.8), *Ugp-2* (9.3) and *Pep-3* (8.1).

Based on average subpopulation differentiation, Southwest Oregon appeared slighly less differentiated over all 31 loci ($\delta = 5.5\%$) than Central Oregon ($\delta = 7.7\%$), the amount of differentiation being in the same order of magnitude, however. Overall, average differentiation among seed zones within Central Oregon was slighly higher, within Southwest Oregon slighly lower than differentiation between the two regions ($\delta = 6.1\%$, *see Table 11, p. 69*). Average subpopulation differentiationδ among seed zones within regions varied widely among the loci and displayed different patterns of differentiation in the two regions. Southwest Oregon was mostly differentiated at *Mnr-1* (13.0), *G6p-2* (9.8), *Aco-1* (9.3), *Mnr-2* (9.2), *Got-3* (8.6), *Got-1* (7.7), *Ugp-2* (7.5), *Gdh-1* (7.4), *Acp-1* (6.8), *Lap-2* (6.6) and *Mpi-1* (6.1). Central Oregon showed high differentiation at most loci, 6 of which exceeding 10% *i.e. Aco-1* (15.0), *Mnr-2* (14.9), *Mnr-1* (12.4), *Ugp-2* (11.6), *Adh-2* (11.5) and *Mpi-2* (10.5). In addition, above average differentiation was manifest for *Pep-3* (9.7), *Acp-1* (9.6), *Idh-1* (9.2), *Pgm-1* (9.2), *Skd-2* (9.1), *Fum-2* (9.0),

LOCUS		SOUTHWEST OREGON					CENTRAL OREGON						LOCUS
	F (IS)	F(IT)	F(ST)	Chi-sq.	df	Prob.	F (IS)	F(IT)	F(ST)	Chi-sq.	df	Prob.	
MNR-1	0.022	0.119	0.099	39.848	27	0.053	-0.054	0.030	0.080	85.657	78	0.259	MNR-1
MNR-2	-0.027	0.044	0.069	31.221	18	0.0271 *	0.094	0.163	0.076	66.592	52	0.084	MNR-2
LAP-2	-0.074	0.003	0.071	30.668	27	0.285	-0.115	-0.018	0.087	118.775	78	0.002 **	LAP-2
LAP-3	0.019	0.031	0.012	7.392	18	0.986	-0.094	-0.023	0.065	55.559	52	0.342	LAP-3
PEP-1	-0.158	-0.079	0.068	19.483	18	0.363	-0.116	-0.038	0.069	42.251	52	0.831	PEP-1
PEP-2	0.008	0.037	0.029	12.875	18	0.799	-0.177	-0.064	0.096	77.811	52	0.011 *	PEP-2
PEP-3	-0.161	-0.098	0.054	22.721	18	0.201	-0.175	-0.087	0.075	53.821	52	0.404	PEP-3
PEP-4	0.179	0.221	0.051	12.188	9	0.203	-0.105	-0.032	0.066	25.701	26	0.479	PEP-4
MPI-1	-0.183	-0.129	0.045	15.662	18	0.616	-0.086	-0.036	0.046	42.057	52	0.836	MPI-1
MPI-2	-0.152	-0.120	0.028	18.183	27	0.898	-0.201	-0.110	0.076	72.555	78	0.652	MPI-2
GOT-1	0.056	0.282	0.240	99.677	27	0.000 ***	-0.137	-0.079	0.050	59.647	78	0.939	GOT-1
GOT-2	-0.093	-0.036	0.052	15.504	18	0.627	-0.099	-0.039	0.054	20.326	26	0.776	GOT-2
GOT-3	-0.171	-0.079	0.079	61.921	36	0.004 **	-0.098	-0.026	0.066	114.182	104	0.233	GOT-3
G6P-2	0.314	0.570	0.373	170.441	27	0.000 ***	-0.050	-0.006	0.041	31.636	52	0.988	G6P-2
ACP-1	-0.068	-0.028	0.037	86.931	36	0.000 ***	-0.139	-0.055	0.073	105.613	104	0.437	ACP-1
GDH-1	0.482	0.712	0.443	98.509	18	0.000 ***	-0.062	-0.008	0.052	18.611	26	0.852	GDH-1
IDH-1	-0.177	-0.115	0.053	10.111	9	0.342	-0.132	-0.067	0.057	22.451	26	0.663	IDH-1
PGM-1	-0.038	0.043	0.078	18.301	18	0.436	-0.082	-0.025	0.053	71.291	78	0.691	PGM-1
ACO-1	0.021	0.064	0.044	25.168	27	0.565	-0.004	0.090	0.093	109.059	78	0.011 *	ACO-1
SKD-1	0.004	0.101	0.097	80.082	18	0.000 ***	0.016	0.069	0.054	39.873	52	0.890	SKD-1
SKD-2	-0.004	0.010	0.014	2.995	9	0.964	-0.073	0.009	0.077	27.772	26	0.370	SKD-2
FDP-2	-0.022	-0.006	0.015	4.167	9	0.900	-0.161	-0.081	0.069	25.029	26	0.517	FDP-2
UGP-1	-0.093	-0.055	0.035	12.738	18	0.806	-0.172	-0.098	0.063	42.772	52	0.815	UGP-1
UGP-2	-0.267	-0.196	0.056	28.376	27	0.392	-0.250	-0.174	0.060	72.853	78	0.643	UGP-2
UGP-3	-0.137	-0.100	0.033	6.842	9	0.653	-0.157	-0.048	0.095	36.384	26	0.085	UGP-3
FUM-2	-0.147	-0.128	0.017	18.613	18	0.416	-0.193	-0.125	0.057	44.956	52	0.745	FUM-2
ADH-2	0.105	0.137	0.035	24.091	18	0.152	-0.097	-0.003	0.085	34.093	26	0.133	ADH-2
PGI-2	-0.011	0.038	0.048	34.382	27	0.155	-0.062	-0.007	0.052	65.669	78	0.839	PGI-2
MDH-1	0.140	0.167	0.032	16.304	18	0.571	-0.107	-0.019	0.080	36.501	26	0.083	MDH-1
MDH-3	-0.026	0.003	0.028	20.045	27	0.828	0.111	0.169	0.065	73.729	78	0.616	MDH-3
MDH-4	0.189	0.217	0.036	8.425	9	0.492	-0.126	-0.043	0.073	56.099	52	0.324	MDH-4
MEAN/TOTAL	-0.038	0.045	0.080	1053.861	621	0.000 ***	-0.094	-0.017	0.070	1749.324	1716	0.282	MEAN/TOTAL

Table 12: *Allele frequency differentiation among seed zones within regions: estimated F - statistics and chi-square tests for homogeneity of allele-frequencies (WORKMAN AND NISWANDER (1970). [F(IS): F - value for individuals within sub-populations (seed zones), F(IT): F - value for individuals in total population (total sample: Southwest N = 236, Central Oregon N = 252), F(ST): F - value for sub-populations within total population]*

ZONE	MNR-1	MNR-2	LAP-2	LAP-3	PEP-1	PEP-2	PEP-3	PEP-4	MPI-1	MPI-2	GOT-1	GOT-2	GOT-3	G6P-2	ACP-1	GDH-1
81	0.104	0.147	0.078	0.095	0.226	0.076	0.120	0.050	0.175	0.057	0.017	0.041	0.039	0.072	0.051	0.113
90	0.444	0.379	0.164	0.096	0.081	0.079	0.142	0.050	0.030	0.155	0.629	0.041	0.114	0.792	0.273	0.774
270	0.014	0.252	0.079	0.032	0.022	0.053	0.109	0.050	0.113	0.057	0.116	0.042	0.161	0.099	0.129	0.114
321	0.207	0.135	0.078	0.033	0.027	0.033	0.200	0.050	0.098	0.073	0.227	0.041	0.246	0.033	0.075	0.062
491	0.144	0.116	0.056	0.021	0.106	0.029	0.068	0.111	0.086	0.056	0.045	0.040	0.097	0.100	0.068	0.011
492	0.186	0.086	0.068	0.033	0.017	0.045	0.063	0.007	0.016	0.056	0.014	0.023	0.105	0.073	0.056	0.051
501	0.196	0.091	0.079	0.032	0.081	0.079	0.116	0.014	0.048	0.045	0.078	0.090	0.049	0.064	0.046	0.114
502	0.123	0.012	0.085	0.020	0.031	0.015	0.007	0.019	0.047	0.031	0.093	0.024	0.067	0.069	0.043	0.059
511	0.051	0.072	0.034	0.031	0.023	0.018	0.005	0.007	0.091	0.063	0.037	0.037	0.065	0.033	0.067	0.015
512	0.084	0.218	0.033	0.028	0.031	0.029	0.041	0.001	0.078	0.077	0.050	0.040	0.086	0.208	0.100	0.103
δ SOUTHWEST	0.130	0.092	0.066	0.030	0.040	0.031	0.044	0.024	0.061	0.055	0.077	0.033	0.086	0.098	0.068	0.074
42	0.114	0.456	0.018	0.163	0.038	0.062	0.201	0.034	0.341	0.185	0.100	0.040	0.078	0.012	0.173	0.010
661	0.218	0.312	0.019	0.056	0.021	0.036	0.036	0.024	0.156	0.115	0.030	0.013	0.053	0.047	0.090	0.011
662	0.181	0.088	0.019	0.111	0.021	0.064	0.150	0.080	0.086	0.191	0.045	0.016	0.080	0.012	0.093	0.010
671	0.175	0.105	0.019	0.006	0.021	0.047	0.193	0.035	0.086	0.027	0.041	0.016	0.084	0.054	0.140	0.010
672	0.204	0.300	0.018	0.058	0.091	0.063	0.050	0.093	0.220	0.171	0.161	0.040	0.115	0.012	0.175	0.010
673	0.114	0.129	0.018	0.173	0.038	0.321	0.092	0.034	0.197	0.082	0.100	0.040	0.157	0.012	0.175	0.010
674	0.118	0.173	0.018	0.127	0.037	0.100	0.059	0.035	0.055	0.178	0.086	0.033	0.043	0.012	0.076	0.010
675	0.100	0.113	0.019	0.046	0.009	0.034	0.069	0.036	0.017	0.055	0.069	0.021	0.032	0.023	0.099	0.011
681	0.128	0.064	0.019	0.076	0.038	0.030	0.177	0.036	0.100	0.158	0.019	0.042	0.047	0.013	0.061	0.011
682	0.066	0.192	0.018	0.026	0.039	0.017	0.030	0.035	0.035	0.129	0.038	0.041	0.043	0.012	0.175	0.010
690	0.237	0.247	0.019	0.060	0.039	0.064	0.023	0.035	0.051	0.111	0.086	0.074	0.068	0.012	0.035	0.010
701	0.222	0.059	0.040	0.148	0.015	0.040	0.105	0.035	0.017	0.065	0.051	0.063	0.079	0.012	0.047	0.042
702	0.040	0.059	0.007	0.053	0.009	0.034	0.018	0.031	0.045	0.032	0.035	0.011	0.099	0.008	0.035	0.023
703	0.038	0.140	0.019	0.061	0.003	0.024	0.130	0.005	0.007	0.051	0.025	0.001	0.050	0.013	0.076	0.030
711	0.066	0.370	0.018	0.090	0.039	0.064	0.051	0.035	0.201	0.155	0.051	0.041	0.080	0.012	0.069	0.010
712	0.077	0.112	0.018	0.026	0.039	0.017	0.142	0.094	0.035	0.068	0.102	0.088	0.080	0.012	0.204	0.010
713	0.116	0.229	0.018	0.088	0.037	0.057	0.133	0.035	0.092	0.055	0.069	0.041	0.171	0.012	0.343	0.010
721	0.153	0.157	0.100	0.043	0.168	0.055	0.102	0.068	0.142	0.281	0.086	0.040	0.079	0.012	0.074	0.010
722	0.219	0.095	0.035	0.028	0.105	0.047	0.171	0.060	0.025	0.082	0.103	0.006	0.093	0.012	0.085	0.010
731	0.091	0.149	0.018	0.103	0.039	0.017	0.142	0.035	0.035	0.155	0.116	0.041	0.080	0.012	0.154	0.010
751	0.052	0.057	0.048	0.169	0.039	0.064	0.043	0.035	0.017	0.153	0.061	0.011	0.043	0.012	0.103	0.010
842	0.077	0.256	0.123	0.058	0.038	0.081	0.204	0.034	0.093	0.149	0.050	0.087	0.079	0.012	0.175	0.010
901	0.452	0.253	0.490	0.163	0.038	0.062	0.201	0.034	0.195	0.185	0.100	0.040	0.078	0.012	0.173	0.010
902	0.454	0.296	0.018	0.146	0.038	0.206	0.202	0.034	0.090	0.186	0.100	0.040	0.078	0.012	0.136	0.010
911	0.027	0.084	0.018	0.103	0.039	0.048	0.145	0.035	0.035	0.084	0.051	0.088	0.061	0.012	0.050	0.010
912	0.355	0.300	0.018	0.343	0.038	0.208	0.143	0.034	0.091	0.188	0.101	0.040	0.115	0.012	0.058	0.010
921	0.069	0.126	0.073	0.148	0.039	0.039	0.222	0.051	0.035	0.189	0.024	0.041	0.073	0.012	0.059	0.010
943	0.154	0.132	0.018	0.148	0.049	0.063	0.163	0.035	0.035	0.104	0.102	0.041	0.049	0.012	0.091	0.075
952	0.078	0.196	0.018	0.044	0.039	0.017	0.030	0.030	0.092	0.110	0.102	0.041	0.157	0.056	0.108	0.010
δ CENTRAL	0.124	0.149	0.026	0.080	0.034	0.050	0.097	0.038	0.069	0.105	0.060	0.033	0.076	0.018	0.096	0.016
δ TOTAL AREA	0.126	0.126	0.054	0.053	0.040	0.041	0.081	0.030	0.062	0.077	0.062	0.031	0.077	0.057	0.010	0.058

Table 13: *Allelic subpopulation (=seed-zone) differentiation according to GREGORIUS and ROBERDS (1986). Differentiation-index D [=genetic distance between the respective subpopulations and its complements) and δ (amount of differentiation among all sub-populations) within regions and for total area*

Table 13 (continued)

ZONE	IDH-1	PGM-1	ACO-1	SKD-1	SKD-2	FDP-2	UGP-1	UGP-2	UGP-3	FUM-2	ADH-2	PGI-2	MDH-1	MDH-3	MDH-4	MEAN
81	0.114	0.115	0.126	0.121	0.055	0.011	0.190	0.164	0.130	0.051	0.219	0.052	0.013	0.141	0.065	0.098
90	0.194	0.116	0.315	0.277	0.106	0.011	0.072	0.151	0.057	0.056	0.007	0.095	0.013	0.056	0.066	0.188
270	0.064	0.024	0.100	0.024	0.004	0.011	0.073	0.132	0.067	0.020	0.027	0.064	0.013	0.071	0.066	0.071
321	0.108	0.030	0.120	0.122	0.081	0.011	0.106	0.259	0.045	0.151	0.122	0.104	0.013	0.122	0.020	0.098
491	0.015	0.084	0.133	0.053	0.050	0.012	0.073	0.077	0.052	0.055	0.074	0.013	0.014	0.026	0.013	0.061
492	0.049	0.023	0.067	0.107	0.019	0.013	0.044	0.043	0.058	0.037	0.050	0.037	0.008	0.031	0.056	0.050
501	0.066	0.218	0.200	0.089	0.060	0.011	0.073	0.263	0.063	0.084	0.044	0.206	0.058	0.047	0.064	0.089
502	0.047	0.033	0.074	0.028	0.004	0.005	0.040	0.046	0.028	0.030	0.050	0.023	0.009	0.039	0.019	0.039
511	0.018	0.031	0.049	0.026	0.046	0.014	0.050	0.061	0.004	0.054	0.082	0.032	0.008	0.030	0.025	0.038
512	0.097	0.013	0.146	0.116	0.005	0.016	0.015	0.080	0.030	0.034	0.094	0.093	0.041	0.067	0.013	0.067
δ SOUTHWEST	0.052	0.043	0.093	0.066	0.027	0.011	0.051	0.075	0.037	0.044	0.065	0.045	0.014	0.043	0.032	0.055
42	0.400	0.209	0.622	0.157	0.436	0.074	0.359	0.349	0.116	0.227	0.416	0.058	0.016	0.442	0.042	0.192
661	0.011	0.075	0.066	0.126	0.011	0.019	0.064	0.190	0.004	0.053	0.118	0.028	0.017	0.030	0.015	0.067
662	0.183	0.078	0.084	0.093	0.220	0.097	0.175	0.278	0.062	0.179	0.146	0.043	0.016	0.101	0.019	0.098
671	0.163	0.158	0.204	0.138	0.126	0.076	0.175	0.105	0.054	0.121	0.146	0.043	0.041	0.033	0.043	0.087
672	0.024	0.171	0.248	0.032	0.060	0.052	0.407	0.099	0.117	0.034	0.040	0.058	0.016	0.113	0.089	0.108
673	0.235	0.042	0.177	0.159	0.271	0.074	0.193	0.159	0.052	0.060	0.251	0.137	0.016	0.098	0.042	0.118
674	0.043	0.104	0.120	0.014	0.067	0.076	0.063	0.082	0.118	0.076	0.059	0.057	0.016	0.039	0.043	0.069
675	0.014	0.009	0.153	0.065	0.056	0.079	0.012	0.113	0.003	0.169	0.066	0.081	0.015	0.068	0.013	0.054
681	0.072	0.100	0.056	0.023	0.105	0.007	0.045	0.080	0.051	0.114	0.051	0.061	0.017	0.040	0.044	0.061
682	0.040	0.086	0.228	0.097	0.003	0.076	0.110	0.100	0.054	0.040	0.153	0.080	0.016	0.039	0.155	0.070
690	0.105	0.136	0.377	0.023	0.126	0.019	0.033	0.187	0.004	0.119	0.031	0.027	0.016	0.097	0.019	0.080
701	0.001	0.039	0.127	0.162	0.036	0.028	0.033	0.033	0.141	0.131	0.224	0.027	0.017	0.211	0.013	0.073
702	0.075	0.065	0.082	0.027	0.012	0.003	0.049	0.016	0.024	0.065	0.081	0.006	0.018	0.060	0.009	0.036
703	0.107	0.040	0.076	0.078	0.191	0.004	0.052	0.113	0.081	0.015	0.072	0.112	0.064	0.034	0.044	0.057
711	0.089	0.082	0.134	0.033	0.003	0.053	0.091	0.165	0.119	0.105	0.024	0.059	0.016	0.053	0.026	0.078
712	0.089	0.211	0.070	0.225	0.255	0.076	0.091	0.162	0.119	0.040	0.235	0.032	0.016	0.068	0.043	0.092
713	0.031	0.118	0.237	0.014	0.006	0.071	0.084	0.155	0.118	0.149	0.014	0.063	0.016	0.100	0.043	0.088
721	0.103	0.111	0.502	0.056	0.067	0.027	0.127	0.185	0.015	0.027	0.015	0.059	0.016	0.041	0.064	0.096
722	0.132	0.059	0.148	0.086	0.074	0.029	0.139	0.069	0.165	0.107	0.137	0.060	0.017	0.081	0.051	0.082
731	0.089	0.086	0.170	0.097	0.126	0.076	0.091	0.093	0.119	0.099	0.041	0.059	0.113	0.100	0.043	0.084
751	0.103	0.098	0.060	0.059	0.016	0.076	0.060	0.049	0.016	0.079	0.119	0.060	0.017	0.094	0.065	0.061
842	0.151	0.085	0.220	0.107	0.315	0.179	0.109	0.286	0.010	0.161	0.040	0.095	0.016	0.099	0.042	0.111
901	0.102	0.331	0.622	0.157	0.066	0.074	0.359	0.283	0.116	0.227	0.416	0.058	0.016	0.098	0.042	0.176
902	0.102	0.210	0.344	0.106	0.186	0.074	0.108	0.350	0.136	0.034	0.086	0.058	0.016	0.098	0.042	0.129
911	0.169	0.049	0.106	0.033	0.197	0.053	0.045	0.097	0.075	0.040	0.282	0.048	0.016	0.076	0.043	0.072
912	0.151	0.175	0.480	0.032	0.188	0.075	0.044	0.032	0.264	0.288	0.341	0.095	0.111	0.099	0.042	0.144
921	0.189	0.215	0.179	0.075	0.152	0.010	0.112	0.161	0.053	0.035	0.173	0.059	0.069	0.100	0.132	0.094
943	0.104	0.167	0.264	0.075	0.067	0.096	0.112	0.213	0.224	0.061	0.169	0.069	0.016	0.075	0.043	0.098
952	0.234	0.092	0.064	0.097	0.132	0.011	0.084	0.097	0.010	0.035	0.088	0.054	0.016	0.100	0.060	0.074
δ CENTRAL	0.092	0.092	0.150	0.075	0.091	0.045	0.084	0.116	0.067	0.090	0.115	0.052	0.026	0.074	0.040	0.077
δ TOTAL AREA	0.117	0.076	0.123	0.066	0.118	0.044	0.063	0.093	0.050	0.065	0.098	0.053	0.020	0.060	0.034	0.088

Ugp-1 (8.4) and *Lap-3* (8.0). Again, however, observed differences in the patterns of differentiation between the two regions may be primarily reflecting differences in sample size between the two regions rather than real differences.

While NEI's unbiased genetic distance between the two regions averaged 0.007 over all 31 loci, the average genetic distance between seed zones was 0.007 (range 0.000 to 0.034) within Central Oregon and 0.018 (range 0.000 to 0.100) within Southwest Oregon. The high average genetic differentiation within Southwest Oregon was mainly due to seed zone 90, which had an average distance of 0.084 compared to all the other zones. The average genetic distance among zones was reduced to 0.0038 when zone 90 was excluded.

Average-linkage cluster analysis based on NEI's unbiased genetic distances between all pairs of the 37 sampled seed zones revealed complex patterns of differentiation (*Figure 13*). Seed zone 90, which was represented by only 7 individuals, appeared genetically distinct from the rest of the area. No clear separation between the two regions was apparent. Within Central Oregon, northern and southern zones clustered together and were separated from the central area. No readily interpretable patterns were manifest within Southwest Oregon.

Complete-linkage cluster analysis based on NEI's unbiased genetic distance between all pairs of the 37 seed zones revealed more distinct clusters than the average-linkage fusion strategy (*Figure 14, p. 75*). With few exceptions (zones 661, 673, 674 and 712 clustered with Southwest Oregon), the two regions were clearly separated. Seed zone 90 was again most distinct from the rest of the sampling area. Average genetic distance of seed zone 90 to the rest of the area was about seven times the distance between the two regions. Within Southwest Oregon no pattern of differentiation was apparent; within Central Oregon, however, northern and southern zones clustered together while zones towards the center formed a distinct cluster.

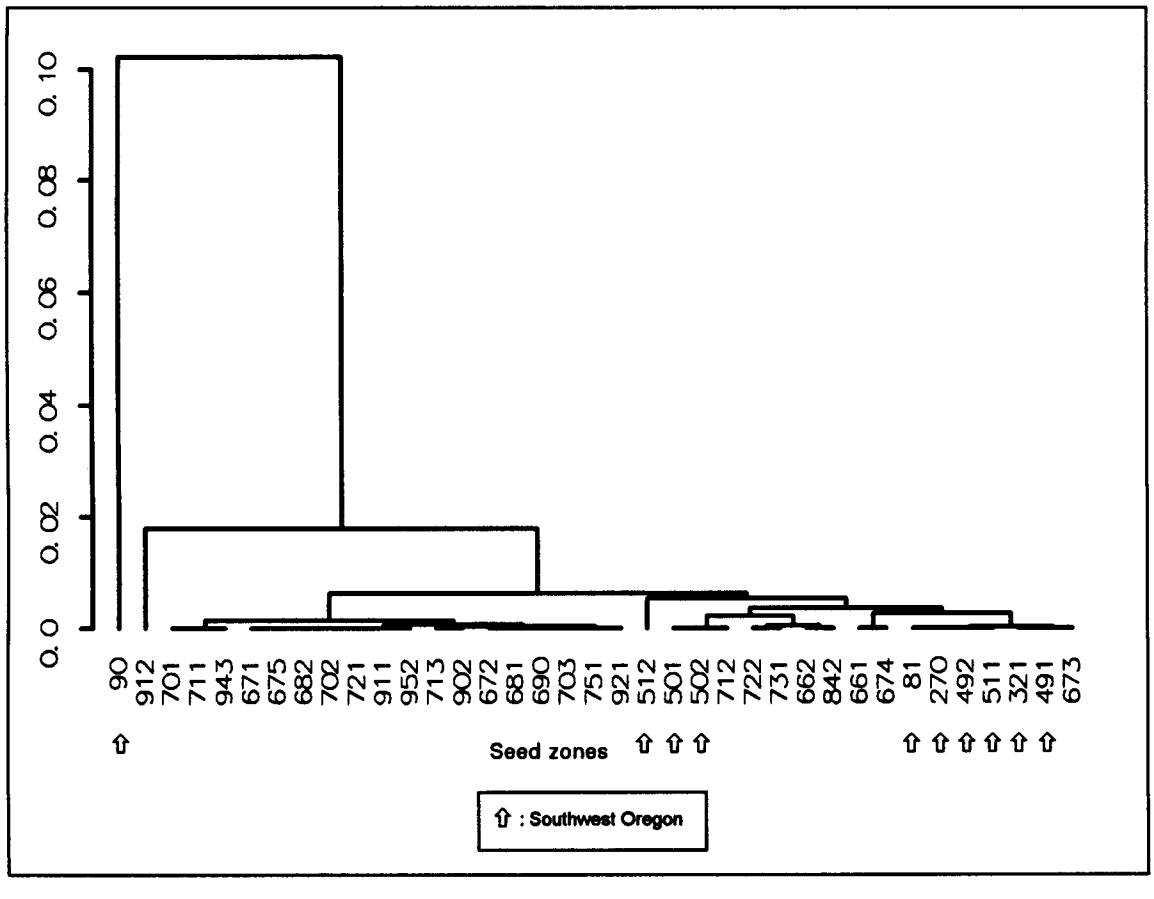


***Figure 13**: Cluster diagram of genetic distance by NEI: Total area. Average-linkage clustering of 37 seed zones, based on matrix of unbiased genetic distances (NEI, 1978)*

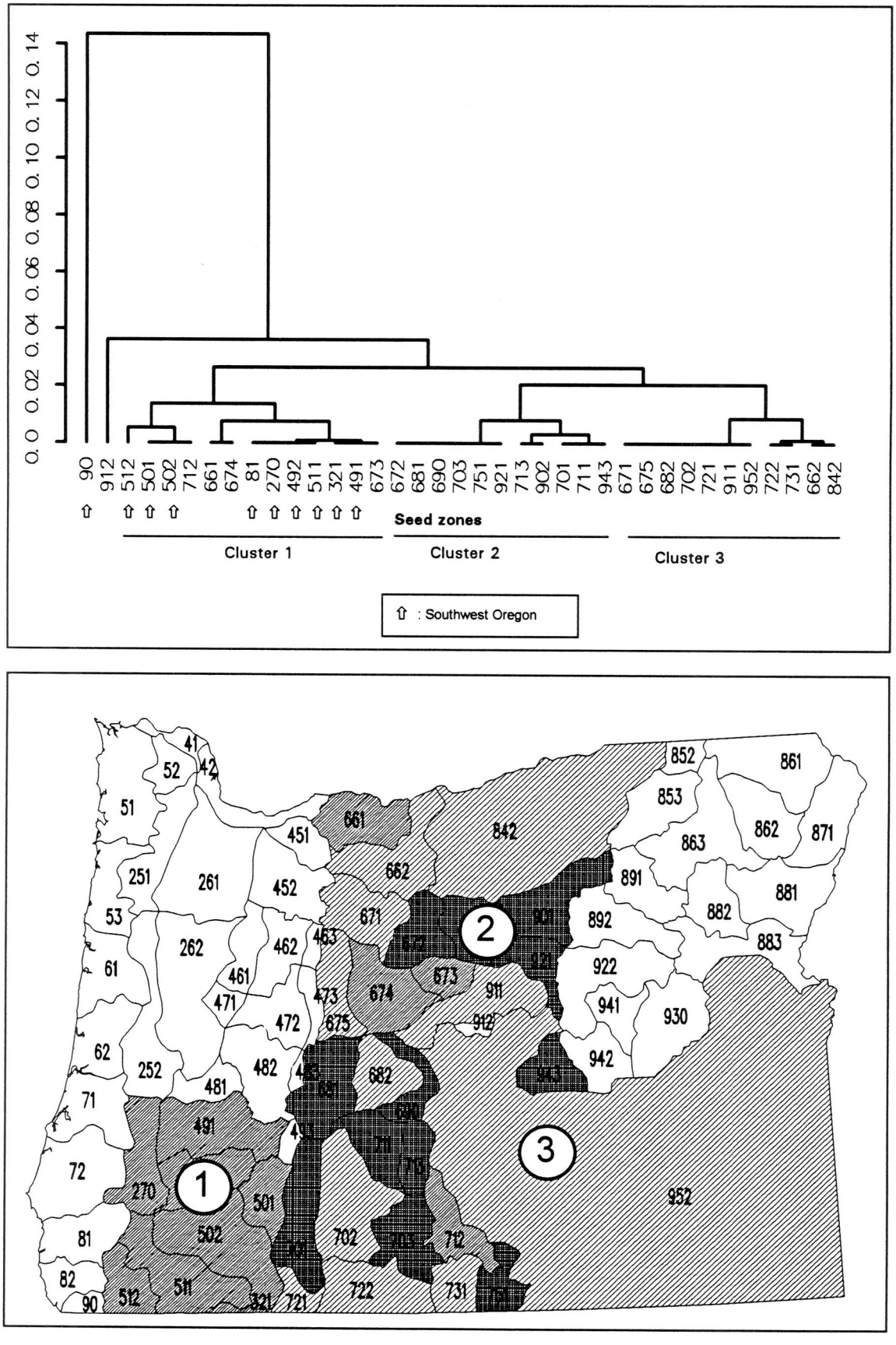


Figure 14: *Cluster diagram of genetic distance by NEI and mapped illustration of 3 main clusters: Total area. Complete-linkage clustering of 37 seed zones, based on a matrix of unbiased genetic distance (NEI, 1978)*

Average-linkage cluster analysis based on genetic distance by GREGORIUS (1974) among the 37 seed zones did not uncover clearly distinct groups of genetically similar zones (*Figure 15*). All seed zones within Southwest Oregon (except zone 90) and most of the zones situated between the southern and the central part of Central Oregon were least differentiated with a genetic distance below 10%. Seed zone 90 was clearly distant from the rest of the area with an average distance of about 22%. Again, zone 661 clustered with southwestern zones. Patterns within Central Oregon were rather complex and not readily interpretable.

Results from complete-linkage cluster analysis based on genetic distance by GREGORIUS (1974) are illustrated in *Figure 16, p. 77*. With the exception of two zones (90, 912), seed zones clearly clustered into two main groups corresponding to the two regions. Zone 90 and 912 were very distinct from the rest of the sampling area. Within Southwest Oregon, all zones except zone 501 formed a group (cluster 1) with an average genetic distance of less than 12%. Zone 661 clustered again with southwestern zones. Within Central Oregon, four distinct groups, encompassing geographically adjoining seed zones, were apparent (the two exceptions being zone 712 and 943). Hence, genetic clustering appeared to have clear associations with regions and, to a lesser extent, with geographical location within regions.

To test the significance of differentiation patterns and to describe the main clusters in terms of multilocus genetic structures, descriptive canonical discriminant analysis was performed on the 5 main groups (*Figure 16, p. 77*) which were formed by cluster analysis on genetic distance by GREGORIUS (1974). Allozyme scores from all 31 loci were used as discriminating variables in the analysis.

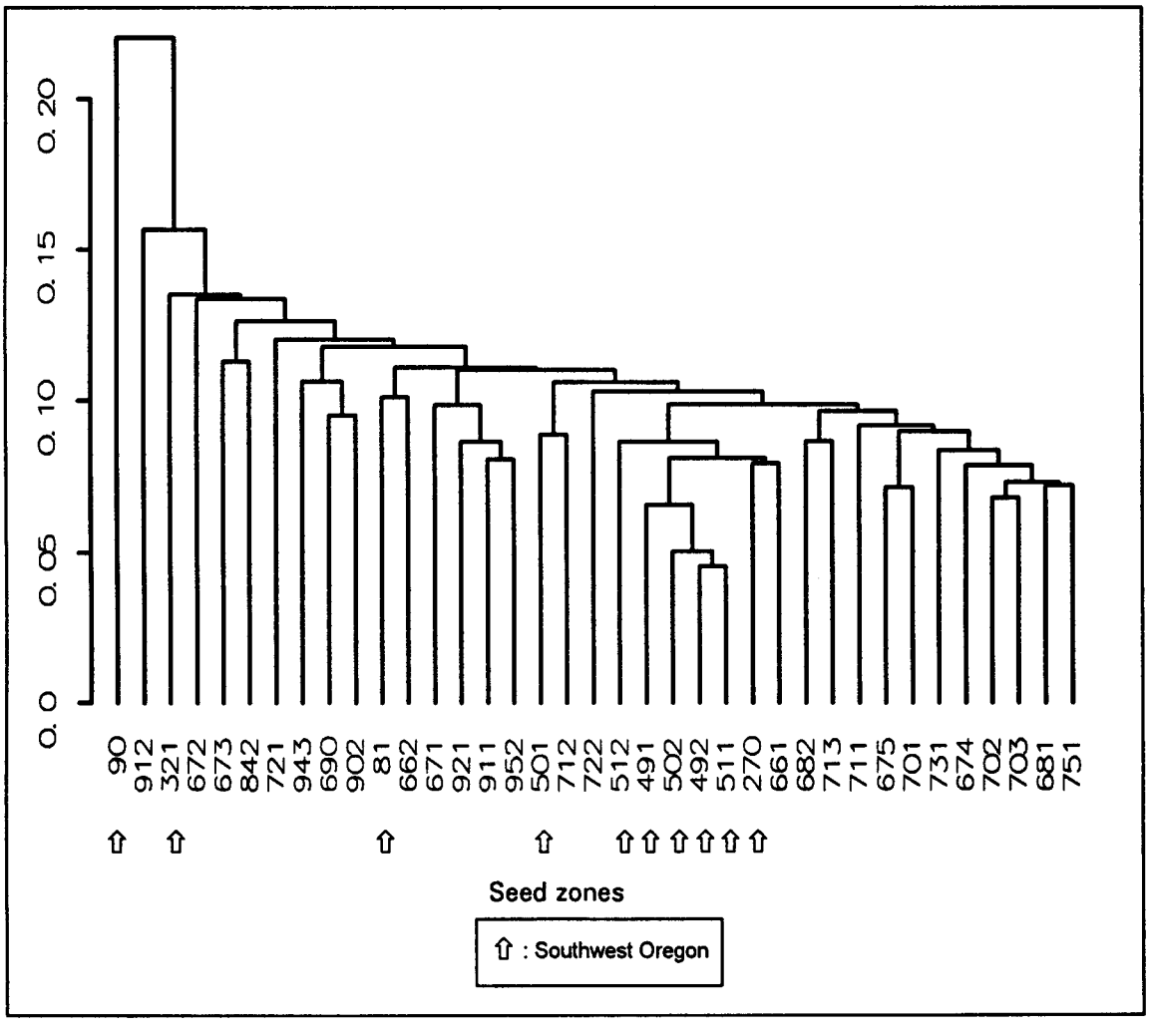
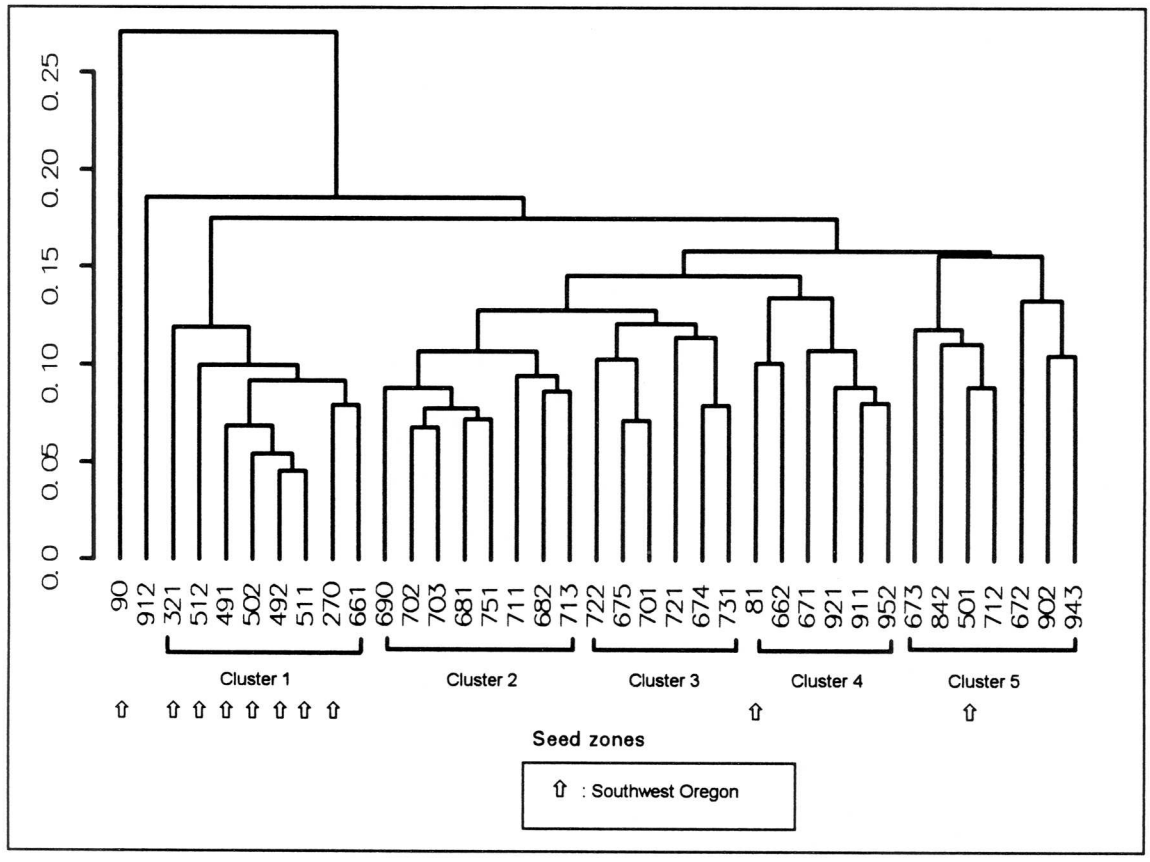


***Figure 15**: Cluster diagram of genetic distance by GREGORIUS: Total area. Average-linkage clustering of 37 seed zones, based on matrix of genetic distance (GREGORIUS, 1974)*



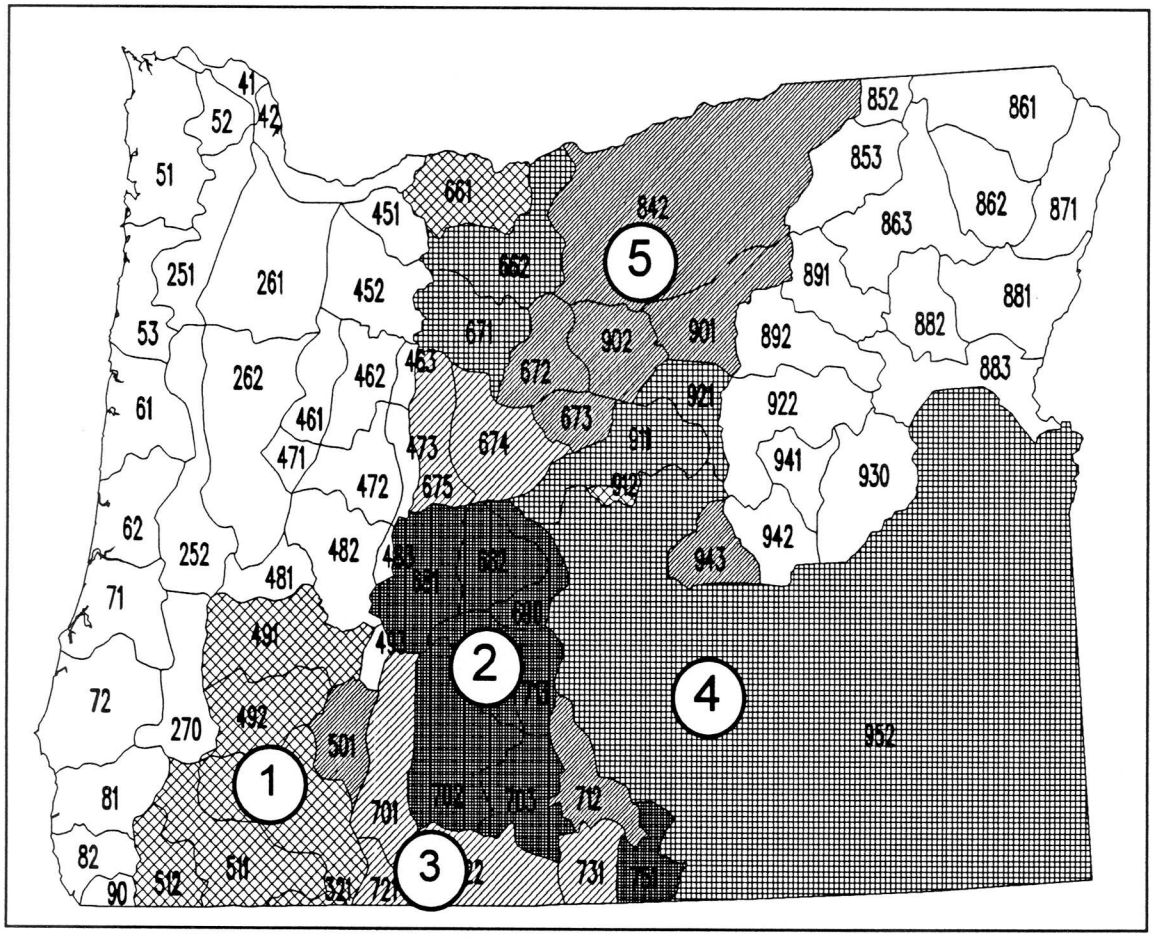


Figure 16: Cluster diagram of genetic distance by GREGORIUS and mapped illustration of main clusters: Total area. Complete-linkage clustering of 37 seed zones, based on matrix of genetic distance (GREGORIUS, 1974)

The first 2 variates generated by canonical discriminant analysis were significant ($p < 0.05$), accounting for 71% of the variation in allozyme scores (*Table 14*). Canonical R^2 indicated that the group contribution to explaining variation in allozyme scores was 31% for the first canonical variate and 9% for the second variate. The first and highly significant ($p < 0.0001$) variate was dominated by several alleles. Highest redundancies *i.e.* largest proportions of variation contributing to group difference (in brackets) were manifest for *Idh1-1* (7.2), *Skd2-1* (5.2), *Gdh1-1* (4.7), *G6p2-1* (3.3), *Fdp2-1* (3.3), *Adh2-1* (3.0), *Pgm1-1* (2.8), *Lap2-1* (2.7), *Mnr1-2* (2.6), *Acp1-2* (2.0), *Pep3-1* (1.6) and *Mnr2-1* (1.6). Hence, canonical variate 1, explaining the single maximum among group variance (31%) was dominated by the same loci which showed a highly significant heterogeneity in allele frequencies between the two regions. Plotting mean canonical scores of seed zones for the first two canonical axes also reflected the association of allozyme scores at these loci with the different regions (*Figure 17, p. 79*). Scores on canonical variate 1 clearly separated the seed zones of the two regions with only two exceptions (zones 661 and 673).

	CANONICAL FUNCTIONS								
Locus	CAN 1		CAN 2		Locus	CAN 1		CAN 2	
Allele	struc.	% red.	struc.	% red.	Allele	struc.	% red.	struc.	% red.
MNR1-1	-0.2693	2.25	-0.1621	0.23	ACP1-3	0.0905	0.25	0.0319	0.02
MNR1-2	0.2878	2.57	0.2155	0.40	ACP1-4	0.0468	0.07	-0.0387	0.02
MNR1-3	0.0194	0.01	0.0910	0.07	GDH1-1	0.3892	4.69	-0.0087	0.00
MNR2-1	0.2303	1.64	0.1539	0.21	GDH1-2	-0.3767	4.40	0.0100	0.00
MNR2-2	-0.1991	1.23	-0.3039	0.80	IDH1-1	-0.4833	7.24	0.1554	0.21
LAP2-1	0.2976	2.74	0.0244	0.01	PGM1-1	-0.3011	2.81	-0.1965	0.34
LAP2-2	-0.1703	0.90	0.0643	0.04	PGM1-2	0.1906	1.13	-0.2204	0.42
LAP2-3	-0.1207	0.45	-0.0549	0.03	PGM1-3	0.0819	0.21	0.0897	0.07
LAP3-1	-0.0219	0.01	-0.0652	0.04	ACO1-1	0.1800	1.00	0.1568	0.21
LAP3-2	0.0355	0.04	0.1039	0.09	ACO1-2	-0.0128	0.01	-0.0471	0.03
PEP1-1	0.1582	0.78	0.1582	0.21	ACO1-3	0.1188	0.44	-0.0999	0.09
PEP1-2	-0.1007	0.31	-0.1421	0.20	SKD1-1	-0.1405	0.61	0.0878	0.07
PEP2-1	0.0829	0.21	0.0211	0.01	SKD1-2	0.1724	0.92	0.0763	0.06
PEP2-2	-0.0766	0.18	0.0047	0.00	SKD2-1	-0.4098	5.20	0.1054	0.10
PEP3-1	-0.2263	1.59	0.1038	0.09	FDP2-1	-0.3245	3.27	0.0164	0.00
PEP3-2	0.2101	1.37	0.0489	0.02	UGP1-1	-0.0756	0.18	-0.3060	0.82
PEP4-1	0.0940	0.27	0.0139	0.00	UGP1-2	0.0105	0.00	0.1549	0.21
MPI1-1	0.0527	0.09	0.0549	0.03	UGP2-1	0.0088	0.00	0.2242	0.44
MPI1-2	0.1024	0.32	0.1823	0.29	UGP2-2	0.0215	0.01	-0.1361	0.16
MPI2-1	-0.1817	1.02	-0.1112	0.11	UGP2-3	-0.0155	0.00	-0.1104	0.11
MPI2-2	0.1070	0.35	0.1385	0.17	UGP3-1	0.0801	0.20	0.0663	0.04
MPI2-3	0.0858	0.23	0.0231	0.01	FUM2-1	0.0548	0.09	0.0665	0.04
GOT1-1	0.0272	0.02	-0.0270	0.01	FUM2-2	-0.0218	0.01	-0.0665	0.04
GOT1-2	0.1095	0.37	0.0125	0.00	ADH2-1	-0.3013	2.81	0.2885	0.72
GOT1-3	-0.0702	0.15	0.0315	0.02	ADH2-2	0.3155	3.08	-0.2878	0.72
GOT2-1	-0.0233	0.02	-0.0055	0.00	PGI2-1	0.1374	0.59	-0.0883	0.07
GOT2-2	0.0355	0.04	0.0066	0.00	PGI2-2	-0.0565	0.10	0.2495	0.54
GOT3-1	0.0862	0.23	0.0335	0.02	PGI2-3	-0.1723	0.92	-0.1293	0.15
GOT3-2	-0.1349	0.56	-0.0348	0.02	MDH1-1	-0.0355	0.04	0.0867	0.07
GOT3-3	0.0033	0.00	-0.0153	0.00	MDH1-2	0.0737	0.17	-0.1594	0.22
GOT3-4	-0.0349	0.04	0.1405	0.17	MDH3-1	0.1303	0.53	-0.0880	0.07
G6P2-1	0.3242	3.26	0.0960	0.08	MDH3-2	-0.1280	0.51	0.0772	0.06
G6P2-2	-0.1285	0.51	-0.0782	0.05	MDH3-3	-0.1163	0.42	0.0597	0.03
G6P2-3	-0.2949	2.70	-0.0629	0.04	MDH4-1	0.0640	0.13	-0.0004	0.00
ACP1-1	0.2006	1.25	0.0208	0.01	MDH4-2	0.1042	0.34	-0.2165	0.41
ACP1-2	-0.2592	2.08	0.0020	0.00					
Eigenvalue	0.6605		0.2497		- Groups: 5 groups formed by complete linkage cluster analysis on genetic distance by GREGORIUS (1974) among the 37 seed zones				
Percent varianc	51.72		19.55						
Can. correlatio	0.5567		0.2951						
r square	0.3099		0.0871		- Discriminating variables: 71 allozyme variables (allozyme scores)				
Significance	0.0001		0.0287						

Table 14: *Canonical discriminant analysis of allozyme scores, discriminating among the 5 groups (formed by complete-linkage cluster analysis on genetic distance by GREGORIUS (1974)) as shown in Figure 16, p. 77 . Struc: structure coefficients, % red: percent redundancies*

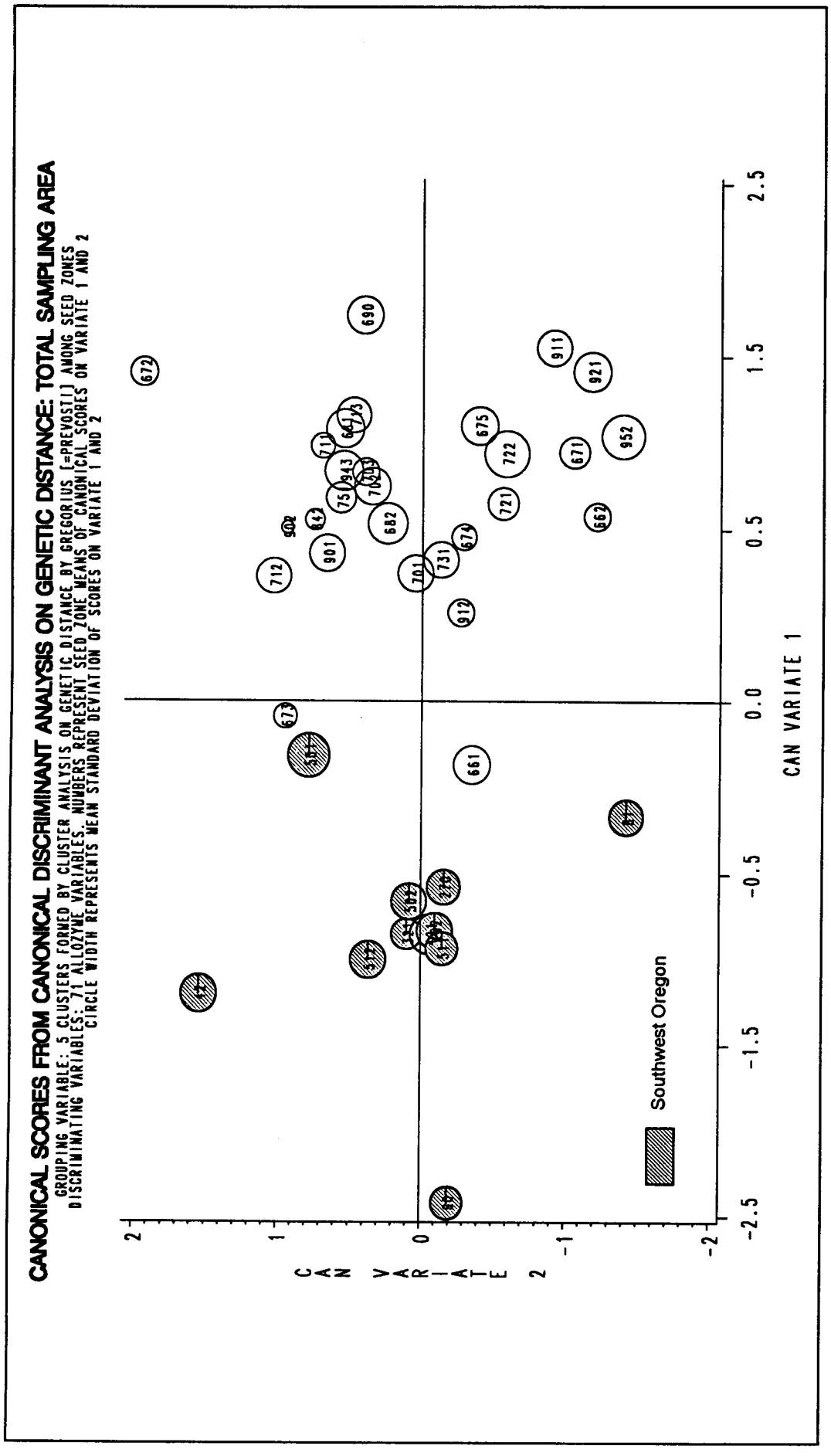


Figure 17: Canonical discriminant analysis on allozyme scores, discriminating among the clusters of genetic distance by GREGORIUS (1974): Total area. Plot of mean canonical scores on canonical variate 1 and 2 for the 37 seed zones [Zones in Southwest Oregon are in solid lines, zones in Central Oregon in normal lines]

The first canonical variate (horizontal axis) separated the zones in a west to east fashion. Regressing canonical scores of variate 1 against longitude produced a significant correlation of *R = -0.58*, suggesting a clinal variation at these loci in a west to east direction. Since correlations between canonical scores and longitude were low within both regions (*i.e.* Southwest: *R* =-0.29, Central Oregon: *R = -0.16*), however, the relatively high overall correlation coefficient was rather a consequence of the frequency distributions being different between the two regions than an indication of a general continuous clinal change in allele frequencies from west to east.

The second canonical variate, explaining 8% of the variation in allozyme scores, was associated with several alleles, having correlation coefficients of *R = 0.20* or higher: *Mnr1-2*, *Mnr2-2*, *Pgm1-2*, *Ugp1-1*, *Ugp2-1*, *Adh2-1*, *Pgi2-2* and *Mdh4-2*. Mean scores on canonical axis 2 (vertical axis) showed a complex pattern (*Figure 17, p. 79*) which gave little evidence of any relationship between the scores on this axis and geographical location.

Allele frequencies of the most discriminating alleles for the five groups are shown in *Table 15*. Cluster 1, grouping all zones from Southwest Oregon, was characterized by highest allele frequencies of the most common allele at *Mnr1-1*, *Pep3-1*, *Idh1-1*, *Pgm1-1*, *Skd2-1*, *Fdp2-1* and *Adh2-1* while frequencies were lowest at *Lap2-1*, *G6p2-1*, *Acp1-1* and *Gdh1-1*. Within Central Oregon, clusters 4 and 5 were most distinct from each other (*i.e.* at *Mnr1-1*, *Pep3-1*, *Idh1-1*, *Pgm1-1* and *Adh2-1*) while clusters 2 and 4 were little differentiated from each other, but relatively distinct from the other groups.

Locus	Cluster 1	[N=246]	Cluster 2	[N=103]	Cluster 3	[N=58]	Cluster 4	[N=45]	Cluster 5	[N=36]
Allele	Mean Freq.	Std	Mean Freq.	Std	Mean Freq.	Std	Mean Freq.	Std	Mean Freq.	Std
MNR1-1	0.502	0.348	0.378	0.331	0.379	0.400	0.422	0.336	0.291	0.301
MNR1-2	0.430	0.348	0.587	0.331	0.551	0.393	0.488	0.345	0.625	0.345
LAP2-1	0.926	0.182	0.990	0.069	0.965	0.127	0.988	0.074	0.972	0.116
PEP3-1	0.863	0.227	0.796	0.275	0.801	0.246	0.755	0.274	0.833	0.267
PEP3-2	0.059	0.167	0.126	0.229	0.094	0.197	0.100	0.202	0.083	0.188
G6P2-1	0.912	0.237	0.995	0.049	0.991	0.065	0.966	0.126	0.986	0.083
G6P2-3	0.058	0.205	0.005	0.049	0.000	0.000	0.011	0.074	0.000	0.000
ACP1-1	0.752	0.312	0.815	0.279	0.844	0.251	0.822	0.242	0.847	0.233
ACP1-2	0.201	0.298	0.111	0.220	0.086	0.190	0.133	0.223	0.152	0.233
GDH1-1	0.894	0.253	0.985	0.084	0.991	0.065	1.000	0.000	0.986	0.083
GDH1-2	0.101	0.251	0.014	0.084	0.008	0.065	0.000	0.000	0.013	0.083
IDH1-1	0.796	0.277	0.601	0.345	0.629	0.357	0.522	0.319	0.694	0.274
PGM1-1	0.890	0.212	0.781	0.302	0.818	0.259	0.811	0.287	0.708	0.324
PGM1-2	0.012	0.077	0.033	0.126	0.017	0.092	0.077	0.183	0.013	0.083
SKD2-1	0.739	0.331	0.519	0.363	0.603	0.359	0.522	0.336	0.736	0.279
FDP2-1	0.985	0.083	0.932	0.172	0.922	0.182	0.922	0.183	0.930	0.175
ADH2-1	0.703	0.340	0.572	0.359	0.629	0.344	0.433	0.312	0.736	0.304
ADH2-2	0.290	0.337	0.427	0.359	0.370	0.344	0.566	0.312	0.263	0.304
Can 1	-0.793	0.999	0.921	1.012	0.738	1.027	0.919	0.968	0.443	0.958
Can 2	-0.030	0.999	0.437	0.919	-0.320	1.050	-1.179	1.086	0.944	1.026

Table 15: Mean allele frequencies of the most discriminating alleles (evaluated by stepwise discriminant analysis) for the five clusters

Predictive discriminant analysis, based on the 22 most discriminating allozyme variables among groups (evaluated by stepwise procedure), produced 3 highly significant canonical variates ($p < 0.0001$) accounting for 88.8% of total variation in allozyme scores. The group contribution to explaining variation in these scores ranged from 34.8% in the first variate to 9.8% in the last variate. Utilization of derived linear discriminant functions to predict cluster membership of individual trees from their allozyme scores, resulted in a reasonably high overall classification accuracy of 48% (*Table 16, p. 81*). Individual trees could be assigned to their initial clusters with a precision that was on average 28% higher than what might be expected by

From actual cluster number:	Percent trees classified into cluster number:					Sample size resp. mean
	1	2	3	4	5	
1	59.30	9.80	8.10	10.20	12.60	246
2	14.50	36.70	9.70	17.50	21.40	103
3	22.40	17.20	29.30	19.00	12.00	58
4	8.90	13.30	11.10	55.60	11.10	45
5	13.90	19.40	2.80	5.60	58.30	36
Prior probabilities [%]	20.00	20.00	20.00	20.00	20.00	
Error rate [%]	40.70	63.30	70.70	44.40	41.70	52.10
Correctly classified [%]	59.30	36.70	29.30	55.60	58.30	47.90
Higher than prior prob. [%]	39.30	16.70	9.30	35.60	38.30	27.90
Crossvalidation:						
Error rate [%]	41.90	71.80	77.60	62.20	66.70	64.00

Table 16: *Percentage of individual trees classified into the initial five groups, using linear discriminant functions from predictive discriminant analysis based on 22 allozyme scores*

chance alone (prior probabilities). Classification rates differed between the clusters, however. High precision were reached for clusters 1, 4 and 5 with 55% to 60% of correctly assigned trees, emphasizing the distinct genetic structures of each group. Clusters 2 and 3 obviously were less distinct in allelic structures since only 30% to 35% of individual trees were correctly assigned to the respective groups.

While seed zones within Southwest Oregon appeared quite uniform in the overall analysis, all (except zones 81 and 501) forming one cluster, patterns were apparent from cluster analyses restricted to Southwest Oregon seed zones only. Complete-linkage clustering, based on genetic distance by NEI between all pairs of seed zones within Southwest Oregon, revealed two main groups, separating a more central area from the slopes to the east as well as to the west (*Figure 18, p. 82*). Zone 90 was again highly distinct from the rest of the area. Average linkage clustering resulted in a dendrogram very similar to that in *Figure 18*.

Patterns of differentiation from complete-linkage cluster analysis on genetic distance by GREGORIUS are illustrated by the dendrogram in *Figure 19, p. 82*. A gradually increasing differentiation from the central area (zones 511, 492, 502, 491) to the east as well as to the west was apparent, the zones to the west (exception 321) being most distinct. Zone 90 again was highly distinct from all other zones.

Patterns of differentiation within Central Oregon, as revealed by complete-linkage cluster analysis on genetic distance by NEI, restricted to Central Oregon only, were quite complex (*Figure 20, p. 83*). However, disregarding some exceptions, a tendency towards a change in genetic structure from the northern as well as from the southern zones towards the central area were manifest, northern zones being more similar to southern zones than to central ones. Zone 912 was very distinct from all the other zones which might have been caused by the small sample size of this zone ($N = 4$).

Four geographic areas with similar genetic structure, regrouping several adjoining seed zones, were apparent from cluster analysis on genetic distance by GREGORIUS among all pairs of the 27 seed zones within Central Oregon (*Figure 21, p. 84*) With some exceptions, a northern area was separated from a central-southeastern and from a central-southwestern area. Again, zone 912 was very distinct from all the other zones.

4.3 Genotypic structure

Mean observed heterozygosity H_O over all 31 loci and for the total sampling area was 0.27 (*Table 9, p. 65*). H_O varied widely among loci, ranging from 0.027 to 0.64. Mean observed and mean expected heterozygosity H_E were identical. Observed heterozygosity was higher than the expected heterozygosity at 13 loci, lower at 16 loci and identical at 2 loci. Observed heterozygosity was slightly lower in Southwest Oregon (0.268) than in Central Oregon (0.271) (*Table 10, p. 66*).

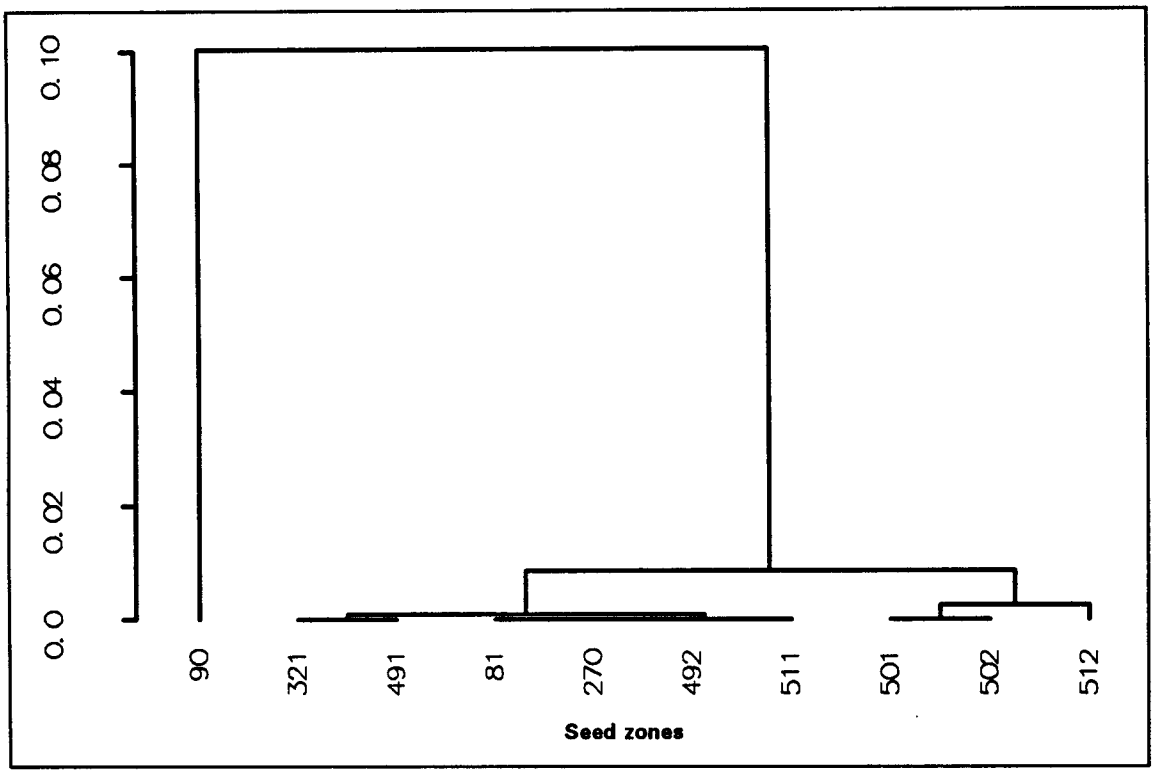


Figure 18: *Cluster diagram of genetic distance by NEI: Southwest Oregon. Complete-linkage clustering of 10 seed zones, based on matrix of genetic distance (NEI, 1978)*

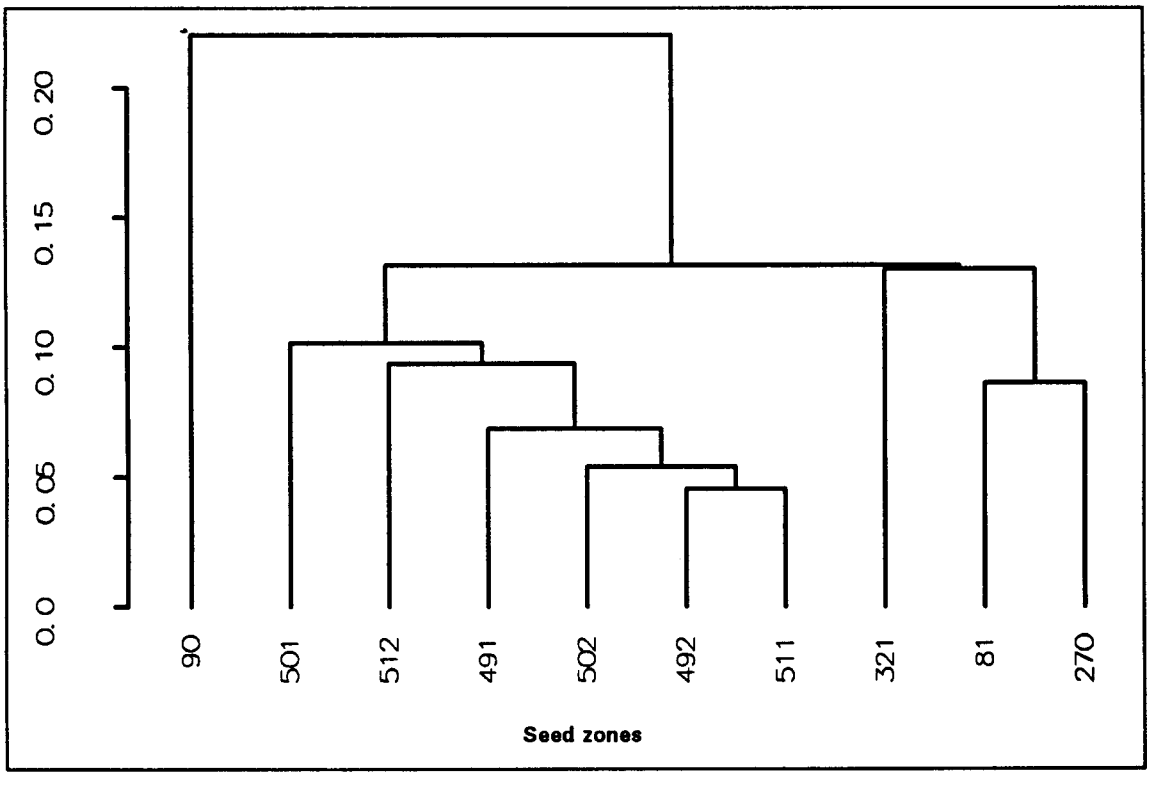


Figure 19: *Cluster diagram of genetic distance by GREGORIUS: Southwest Oregon. Complete-linkage clustering of 10 seed zones, based on matrix of genetic distance (GREGORIUS, 1974)*

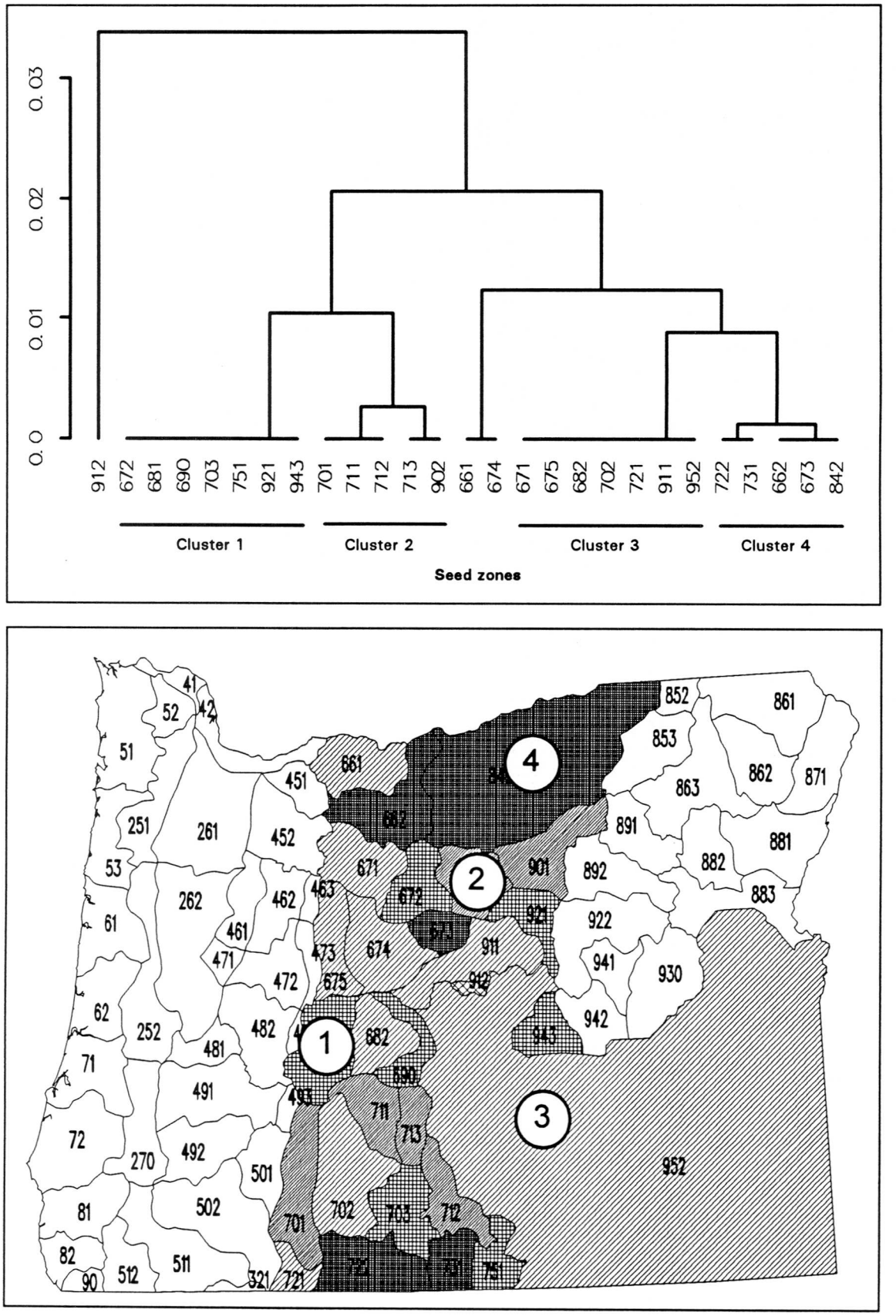
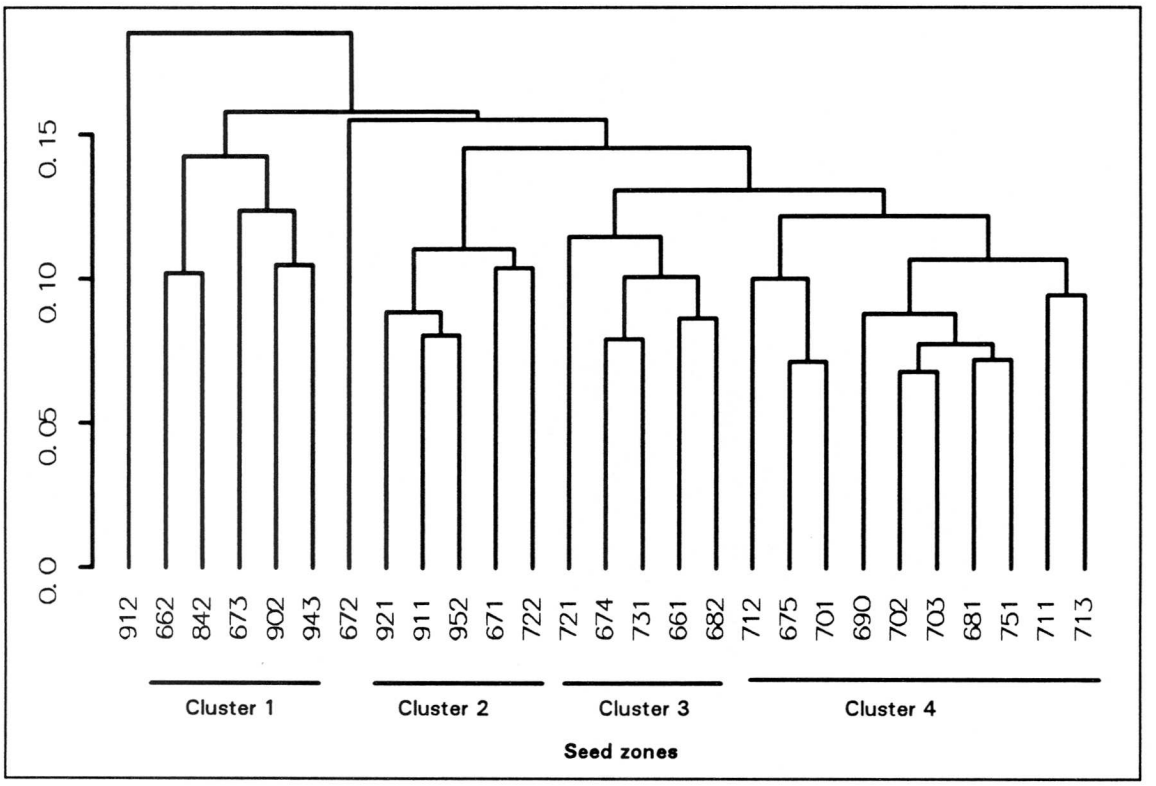


Figure 20: *Cluster diagram of genetic distance by NEI and mapped illustration of main clusters: Central Oregon. Complete-linkage clustering of 27 seed zones, based on matrix of genetic distance (NEI, 1978)*



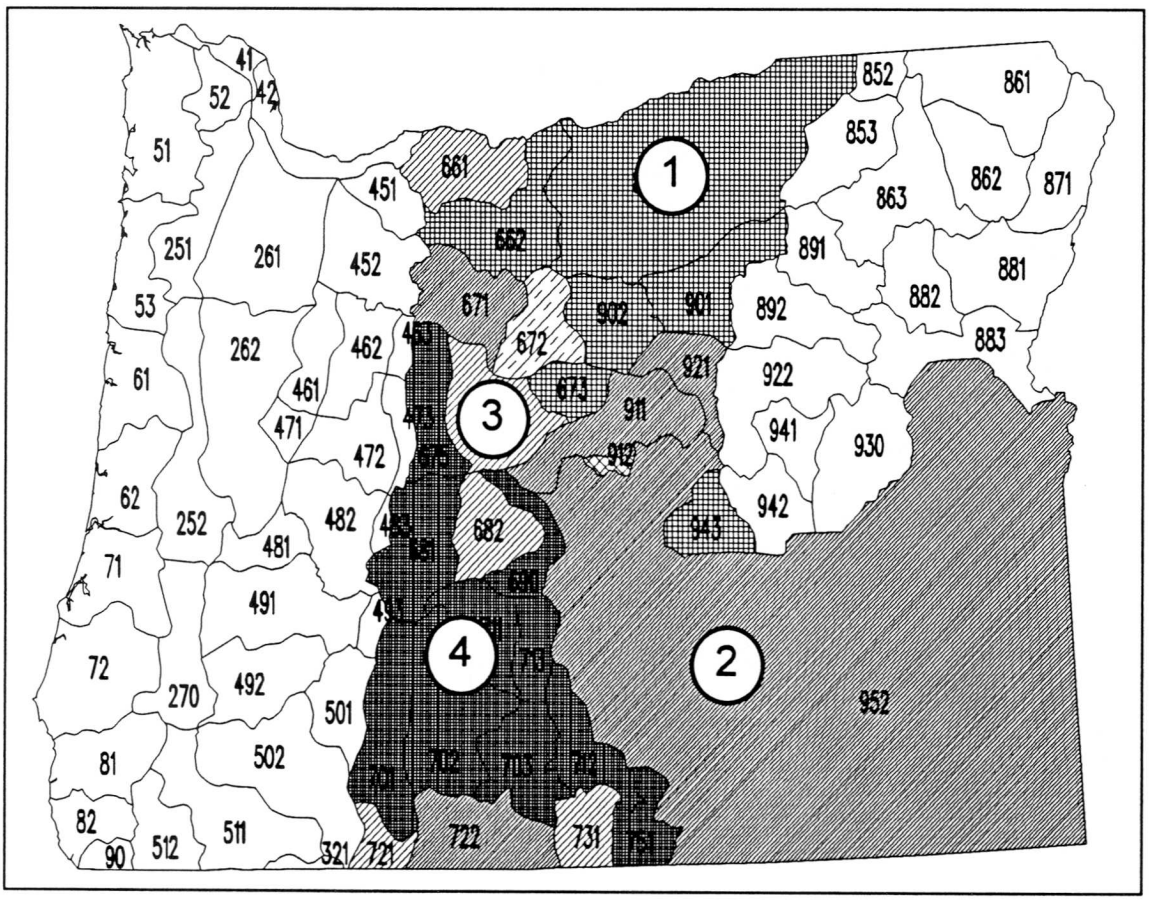


Figure 21: *Cluster diagram of genetic distance by GREGORIUS and mapped illustration of main clusters: Central Oregon. Complete-linkage clustering of 27 seed zones, based on matrix of genetic distance (GREGORIUS, 1974)*

Genotypic structures and tests for Hardy-Weinberg expectations are presented in *Table 17, p. 86*, *Table 18, p. 87 and Table 19, p. 88*. Where more than two alleles were observed at a locus, genotypes were pooled into three classes *i.e.* homozygote individuals for the most common allele, heterozygote individuals for the most common and another allele and all other genotypes. Different goodness-of-fit tests, using the observed genotypic frequencies, as well as estimates of fixation index *F*, are indicated.

Overall, *i.e.* averaged over all 31 loci, the genotypic structure of the total sampling area was in Hardy-Weinberg equilibrium (*Table 17, p. 86*). Significant deviations of genotypic structures from Hardy-Weinberg expectations were manifested at 11 out of 31 loci, however. Highly significant deviations from expectations were found for *Mpi-2*, *G6p-2*, *Gdh-1*, *Ugp-1* and *Ugp-2*, while significant differences were manifested at *Mnr-2*, *Pep-3*, *Skd-2*, *Ugp-3*, *Fum-2* and *Mdh-3*. Seventeen loci had a positive *F* - index *i.e.* exhibited a deficiency of heterozygous individuals while 14 loci had an excess of heterozygotes compared to expectation. Extremely high deficiencies of heterozygous individuals were observable at *G6p-2* and *Gdh-1* with *F* - values of 0.364 and 0.349 respectively. A moderate to high excess of heterozygous individuals were present at locus *Mpi-2*, *Ugp-1*, *Ugp-2* and *Fum-2*.

Overall genotypic structure within Southwest Oregon was characterized by a slight excess of heterozygous individuals (*F* - value of -0.026), averaged over all 31 loci (*Table 18, p. 87*). Significant deviations from Hardy-Weinberg expectations were found at 10 loci with highly significant differences at *G6p-2*, *Gdh-1* and *Ugp-2* and significant deviations at *Mpi-1*, *Mpi-2*, *Ugp-1*, *Fum-2* and *Mdh-1*. Positive fixation indices were found for 17 loci while 14 loci showed an excess of heterozygous individuals. Extremely high deficiencies occurred at *G6p-2*, *Gdh-1* and *Mdh-1* with fixation indices of 0.40, 0.34 and 0.16, respectively while a moderate to high excess of heterozygotes was found at *Mpi-1*, *Mpi-2*, *Ugp-1*, *Ugp-2* and *Fum-2*.

Within Central Oregon, overall genotypic structure was characterized by a slight excess of heterozygous individuals (*F* - value of -0.054), averaged over all 31 loci (*Table 19, p. 88*). Deviations from expectations were significant at 7 loci (*Mnr-2*, *Pep-3*, *Mpi-2*, *Ugp-1*, *Ugp-2*, *Fum-2* and *Mdh-3*). Positive *F* - values were found for 8 loci only while 23 loci were characterized by an excess of heterozygous individuals. Deficiencies of heterozygotes were generally less than within Southwest Oregon, with *F* - *values* ranging between 0.005 and 0.165. The same holds for the negative *F* - values, with a range between -0.000 and -0.153.

4.4 Discussion

Estimates of multiplicity and genetic diversity indicate that ponderosa pine in Oregon has substantial genetic variation. Both multiplicity and diversity measures are high compared to values reported for woody plant species in general, as well as for conifer species in particular. The average number of alleles per locus (N_a) is 48% higher, effective number of alleles (N_e) 19% higher, proportion of polymorphic loci (*P*) 29% higher, and the polymorphic index (*PI*) 52% higher than respective estimates reported for woody plant species by HAMRICK et al. (1992)[1]. Compared to results from conifer species only, the estimated N_a is 43% higher, *P* is 24% and *PI* is 30% higher than the mean values reported by HAMRICK et al. (1981)[2] .

Levels of genetic diversity are influenced by life history characteristics and ecological traits such as taxonomic status, geographic range, generation length, mode of reproduction, mating system, fecundity, stage of succession and habitat type (HAMRICK et. al., 1979, 1981, 1992). Moreover, diversity estimates are largely affected by sample size, sampling strategy and investigated gene loci. Heterozygosity estimates often decrease as more loci are added, and are very sensitive to the number of monomorphic loci that are retained in the genetic survey (STRAUSS and CONKLE, 1986). Likewise, estimates of average and effective number of alleles are influenced by sample size, number of monomorphic loci and choice of surveyed enzymes. Hence, differences among studies may reflect true population differences but likewise

[1] *based on 191 investigations, with an average number of 9.2 populations and a mean number of 18.1 loci per investigation*

[2] *based on 20 conifer species, with an average number of 20.1 loci per investigation*

Genotype	11	12	13	14	15	22	23	24	25	33	34	35	44	45	55	Chi-sq	Prob	Chi-sq	Prob	Exact Test	Obs.	Exp.	F	
Frequencies	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N		Chi-sq	LEVENE	LEVENE	Prob P	Hetero-	Hetero-	Index	
LOCUS																pooled	P	pooled	P	pooled	zygotes	zygotes		LOCUS
MNR-1	96	211	3	23	-	123	6	22	-	0	2	-	2	-	-	0.204	0.651	0.225	0.635	0.651	267	272.1	0.018	MNR-1
MNR-2	96	119	90	-	-	57	84	-	-	42	-	-	-	-	-	6.490	0.011*	6.607	0.010 *	0.012 *	293	320.4	0.084	MNR-2
LAP-2	444	12	19	12	-	0	0	0	-	0	0	-	1	-	-	0.001	0.969	0.000	0.987	1.000	43	43.7	0.014	LAP-2
LAP-3	351	81	44	-	-	4	7	-	-	1	-	-	-	-	-	0.048	0.826	0.059	0.808	0.861	132	131.6	-0.004	LAP-3
PEP-1	432	45	11	-	-	0	0	-	-	0	-	-	-	-		1.808	0.179	1.774	0.183	0.392	56	43.4	-0.051	PEP-1
PEP-2	424	42	19	-	-	2	1	-	-	0	-	-	-	-	-	0.246	0.620	0.272	0.602	0.487	62	63.4	0.022	PEP-2
PEP-3	328	76	78	-	-	2	1	-	-	3	-	-	-	-	-	6.776	0.009 **	6.680	0.010 *	0.009 **	155	145	-0.070	PEP-3
PEP-4	450	36	-	-	-	2	-	-	-	-	-	-	-	-	-	1.848	0.174	1.953	0.162	0.188	36	38.4	0.062	PEP-4
MPI-1	319	24	135	-	-	0	1	-	-	9	-	-	-	-	-	3.759	0.053	3.682	0.055	0.068	160	150.3	-0.066	MPI-1
MPI-2	327	52	93	14	-	0	1	0	-	1	0	-	0	-	-	14.275	0.000 ***	14.151	0.000 ***	0.000 ***	160	143.3	-0.118	MPI-2
GOT-1	389	12	75	6	-	0	0	0	-	6	0	-	0	-	-	0.028	0.868	0.036	0.850	0.814	93	95.5	0.025	GOT-1
GOT-2	449	38	1	-	-	0	0	-	-	0	-	-	-	-	-	0.845	0.358	0.823	0.364	1.000	39	37.5	-0.041	GOT-2
GOT-3	400	52	14	10	9	0	0	0	0	2	1	0	0	0	0	0.442	0.506	0.416	0.519	0.787	86	85.2	-0.011	GOT-3
G6P-2	449	15	13	1	-	1	1	0	-	8	0	-	0	-	-	69.315	0.000 ***	70.930	0.000 ***	0.000 ***	30	47.2	0.364	G6P-2
ACP-1	305	122	9	23	7	15	2	1	0	3	0	0	0	0	1	0.016	0.898	0.023	0.881	0.892	164	170.8	0.039	ACP-1
GDH-1	442	33	2	-	-	11	0	-	-	0	-	-	-	-	-	59.058	0.000 ***	60.286	0.000 ***	0.000 ***	35	53.9	0.349	GDH-1
IDH-1	239	207	-	-	-	42	-	-	-	-	-	-	-	-	-	0.089	0.765	0.076	0.782	0.829	207	204.5	-0.014	IDH-1
PGM-1	343	20	1	111	-	0	0	3	-	0	0	-	10	-	-	0.005	0.944	0.009	0.926	0.869	135	135.9	0.005	PGM-1
ACO-1	59	91	7	122	-	46	0	83	-	2	13	-	65	-	-	0.341	0.559	0.368	0.544	0.555	316	331.3	0.045	ACO-1
SKD-1	364	99	12	-	-	11	2	-	-	0	-	-	-	-	-	1.612	0.204	1.678	0.195	0.192	113	119.6	0.055	SKD-1
SKD-2	220	200	-	-	-	68	-	-	-	-	-	-	-	-	-	4.154	0.042 *	4.249	0.039 *	0.045 *	200	220.5	0.092	SKD-2
FDP-2	445	43	-	-	-	0	-	-	-	-	-	-	-	-	-	1.037	0.309	1.011	0.315	0.615	43	41.1	-0.046	FDP-2
UGP-1	294	67	119	-	-	3	2	-	-	3	-	-	-	-	-	12.666	0.000 ***	12.532	0.000 ***	0.000 ***	188	170.1	-0.106	UGP-1
UGP-2	160	185	106	4	-	12	17	0	-	3	1	-	0	-	-	43.002	0.000 ***	42.718	0.000 ***	0.000 ***	313	259.7	-0.207	UGP-2
UGP-3	372	114	-	-	-	2	-	-	-	-	-	-	-	-	-	4.780	0.029 *	4.707	0.030 *	0.031 *	114	103.8	-0.099	UGP-3
FUM-2	271	191	11	-	-	13	1	-	-	1	-	-	-	-	-	9.867	0.002 **	9.740	0.002 **	0.002 **	203	180.2	-0.128	FUM-2
ADH-2	211	207	2	-	-	67	1	-	-	0	-	-	-	-	-	1.936	0.164	2.000	0.157	0.166	210	224.3	0.063	ADH-2
PGI-2	412	34	32	8	-	1	0	0	-	0	0	-	1	-	-	0.473	0.492	0.446	0.504	0.758	74	73.7	-0.005	PGI-2
MDH-1	475	10	2	-	-	0	1	-	-	0	-	-	-	-	-	8.296	0.004 **	9.047	0.003 **	0.090	13	13.8	0.060	MDH-1
MDH-3	387	56	16	16	-	6	3	1	-	1	0	-	2	-	-	7.745	0.005 **	7.918	0.005 **	0.013 *	92	104.3	0.117	MDH-3
MDH-4	438	2	46	-	-	0	0	-	-	2	-	-	-	-	-	0.304	0.581	0.335	0.563	0.639	48	49.4	0.027	MDH-4
MEAN																					131.6	131.4	-0.001	MEAN

***Table 17**: Observed numbers of genotypes, observed and expected numbers of heterozygotes, tests for Hardy-Weinberg expectations, and estimated fixation index at 31 loci: Total area (N = 488)*

Genotype	11	12	13	14	15	22	23	24	25	33	34	35	44	45	55	Chi-sq	Prob	Chi-sq	Prob	Exact Test	Obs.	Exp.	F	
Frequencies	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N		Chi-sq	LEVENE	LEVENE	Prob P	Hetero-	Hetero-	Index	
LOCUS																pooled	P	pooled	P	pooled	zygotes	zygotes		LOCUS
MNR-1	58	103	1	12	-	47	2	10	-	0	1	-	1	-	-	0.037	0.847	0.051	0.821	0.896	129	130.6	0.012	MNR-1
MNR-2	33	62	46	-	-	31	43	-	-	20	-	-	-	-	-	0.049	0.825	0.065	0.799	0.889	151	155.4	0.028	MNR-2
LAP-2	200	10	13	11	-	0	0	0	-	0	0	-	1	-	-	0.122	0.727	0.102	0.749	1.000	34	34.1	0.005	LAP-2
LAP-3	175	33	19	-	-	3	5	-	-	0	-	-	-	-	-	2.638	0.101	2.769	0.096	0.112	57	60.4	0.056	LAP-3
PEP-1	198	27	10	-	-	0	0	-	-	0	-	-	-	-		1.716	0.190	1.666	0.197	0.373	37	34.7	-0.067	PEP-1
PEP-2	202	19	11	-	-	2	1	-	-	0	-	-	-	-	-	2.236	0.135	2.387	0.122	0.140	31	33.8	0.084	PEP-2
PEP-3	171	24	39	-	-	1	0	-	-	0	-	-	-	-	-	3.659	0.056	3.569	0.059	0.093	63	58.2	-0.083	PEP-3
PEP-4	214	19	-	-	-	2	-	-	-	-	-	-	-	-	-	4.058	0.044 *	4.354	0.037 *	0.095	19	21.9	0.131	PEP-4
MPI-1	156	8	69	-	-	0	0	-	-	2	-	-	-	-	-	5.187	0.023 *	5.071	0.024 *	0.022 *	77	68.3	-0.128	MPI-1
MPI-2	165	23	44	3	-	0	0	0	-	0	0	-	0	-	-	7.197	0.007 **	7.082	0.008 **	0.004 **	70	62.1	-0.126	MPI-2
GOT-1	186	4	37	2	-	0	0	0	-	6	0	-	0	-	-	3.085	0.079	3.240	0.072	0.104	43	49.2	0.126	GOT-1
GOT-2	217	17	1	-	-	0	0	-	-	0	-	-	-	-	-	0.373	0.542	0.351	0.553	1.000	18	17.4	-0.038	GOT-2
GOT-3	184	34	8	5	3	0	0	0	0	1	0	0	0	0	0	1.548	0.213	1.487	0.223	0.325	50	47.7	-0.047	GOT-3
G6P-2	202	11	11	1	-	1	1	0	-	8	0	-	0	-	-	39.745	0.000 ***	40.861	0.000 ***	0.000 ***	24	40.0	0.400	G6P-2
ACP-1	135	68	3	9	4	14	0	1	0	0	0	0	0	0	1	0.350	0.554	0.391	0.532	0.598	85	91.5	0.071	ACP-1
GDH-1	194	28	2	-	-	11	0	-	-	0	-	-	-	-	-	29.003	0.000 ***	33.695	0.000 ***	0.000 ***	30	46.5	0.345	GDH-1
IDH-1	152	77	-	-	-	6	-	-	-	-	-	-	-	-	-	1.063	0.302	1.002	0.317	0.397	77	72.1	-0.067	IDH-1
PGM-1	184	5	-	44	-	0	-	0	-	-	-	-	2	-	-	0.415	0.519	0.379	0.538	0.748	49	47.5	-0.031	PGM-1
ACO-1	23	32	0	67	-	23	0	49	-	0	7	-	34	-	-	1.691	0.193	1.608	0.205	0.225	155	156.6	0.010	ACO-1
SKD-1	184	40	6	-	-	3	2	-	-	0	-	-	-	-	-	1.069	0.301	1.152	0.283	0.342	48	50.1	0.043	SKD-1
SKD-2	139	76	-	-	-	20	-	-	-	-	-	-	-	-	-	3.980	0.046 *	4.124	0.042 *	0.053	76	87.4	0.130	SKD-2
FDP-2	230	5	-	-	-	0	-	-	-	-	-	-	-	-	-	0.027	0.869	0.022	0.883	1.000	5	5.0	-0.011	FDP-2
UGP-1	149	32	52	-	-	0	1	-	-	1	-	-	-	-	-	7.151	0.007 **	7.016	0.008 **	0.008 **	85	75.4	-0.128	UGP-1
UGP-2	64	97	57	3	-	6	7	0	-	1	0	-	0	-	-	37.511	0.000 ***	37.127	0.000 ***	0.000 ***	164	129.6	-0.265	UGP-2
UGP-3	176	58	-	-	-	1	-	-	-	-	-	-	-	-	-	2.748	0.097	2.669	0.102	0.141	58	52.3	-0.108	UGP-3
FUM-2	125	96	6	-	-	6	1	-	-	1	-	-	-	-	-	5.586	0.018 *	5.448	0.020 *	0.023 *	103	90.5	-0.139	FUM-2
ADH-2	123	84	2	-	-	25	1	-	-	0	-	-	-	-	-	3.260	0.071	3.385	0.066	0.083	87	98.3	0.115	ADH-2
PGI-2	188	17	24	5	-	1	0	0	-	0	0	-	0	-	-	1.066	0.302	1.014	0.314	0.482	46	44.5	-0.033	PGI-2
MDH-1	230	2	2	-	-	0	1	-	-	0	-	-	-	-	-	24.778	0.000 ***	29.934	0.000 ***	0.032 *	5	5.9	0.159	MDH-1
MDH-3	177	35	12	6	-	2	0	1	-	1	0	-	1	-	-	0.191	0.662	0.223	0.637	0.580	54	56.8	0.049	MDH-3
MDH-4	207	-	26	-	-	-	-	-	-	2	-	-	-	-	-	1.295	0.255	1.409	0.235	0.235	26	28.1	0.074	MDH-4
MEAN																					63.1	61.5	-0.0257	MEAN

***Table 18**: Observed numbers of genotypes, observed and expected numbers of heterozygotes, tests for Hardy-Weinberg expectations, and estimated fixation index at 31 loci: Southwest Oregon (N = 235)*

Genotype	11	12	13	14	15	22	23	24	25	33	34	35	44	45	55	Chi-sq	Prob Chi-sq	Chi-sq LEVENE	Prob LEVENE	Exact Test Prob P	Obs. Hetero-	Exp. Hetero-	F Index	
Frequencies	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N									
LOCUS																pooled	P	pooled	P	pooled	zygotes	zygotes		LOCUS
MNR-1	38	108	2	10	-	76	4	12	-	0	1	-	1	-	-	0.012	0.913	0.020	0.888	0.899	137	137.9	0.005	MNR-1
MNR-2	63	57	44	-	-	26	40	-	-	22	-	-	-	-	-	9.140	0.003 **	9.322	0.002 **	0.003 **	141	162.9	0.133	MNR-2
LAP-2	243	2	6	1	-	0	0	0	-	0	0	-	0	-	-	0.083	0.773	0.074	0.786	1.000	9	8.9	-0.014	LAP-2
LAP-3	175	48	25	-	-	1	2	-	-	1	-	-	-	-	-	1.373	0.241	1.309	0.253	0.349	75	71.1	-0.057	LAP-3
PEP-1	233	18	1	-	-	0	0	-	-	0	-	-	-	-		0.387	0.534	0.366	0.545	1.000	19	18.4	-0.037	PEP-1
PEP-2	221	23	8	-	-	0	0	-	-	0	-	-	-	-	-	1.082	0.298	1.046	0.307	0.608	31	29.5	-0.052	PEP-2
PEP-3	156	52	39	-	-	1	1	-	-	3	-	-	-	-	-	4.051	0.044 *	3.943	0.047 *	0.050 *	92	85.9	-0.073	PEP-3
PEP-4	235	17	-	-	-	0	-	-	-	-	-	-	-	-	-	0.307	0.579	0.288	0.591	1.000	17	16.5	-0.035	PEP-4
MPI-1	162	16	66	-	-	0	1	-	-	7	-	-	-	-	-	0.378	0.539	0.341	0.559	0.688	83	81.8	-0.016	MPI-1
MPI-2	162	29	48	11	-	0	1	0	-	1	0	-	0	-	-	7.293	0.007 **	7.162	0.007 **	0.005 **	89	80.2	-0.112	MPI-2
GOT-1	202	8	38	4	-	0	0	0	-	0	0	-	0	-	-	3.057	0.080	2.990	0.084	0.146	50	46.1	-0.087	GOT-1
GOT-2	232	20	-	-	-	0	-	-	-	0	-	-	-	-	-	0.430	0.512	0.408	0.523	1.000	20	19.2	-0.041	GOT-2
GOT-3	215	18	6	5	6	0	0	0	0	1	1	0	0	0	0	0.188	0.665	0.220	0.639	0.648	36	37.1	0.027	GOT-3
G6P-2	246	4	2	-	-	0	0	-	-	0	-	-	0	-	-	0.037	0.848	0.030	0.862	1.000	6	5.9	-0.009	G6P-2
ACP-1	170	54	6	14	3	1	2	0	0	2	0	0	0	0	0	1.1225	0.268	1.163	0.281	0.377	79	75.9	-0.042	ACP-1
GDH-1	247	5	-	-	-	0	-	-	-	-	-	-	-	-	-	0.025	0.874	0.020	0.887	1.000	5	4.9	-0.010	GDH-1
IDH-1	86	130	-	-	-	36	-	-	-	-	-	-	-	-	-	1.381	0.240	1.308	0.253	0.294	130	121.3	-0.074	IDH-1
PGM-1	158	15	1	67	-	0	0	3	-	0	0	-	8	-	-	0.001	0.981	0.003	0.956	1.000	86	86.6	0.005	PGM-1
ACO-1	36	59	7	55	-	23	0	34	-	2	6	-	30	-	-	0.065	0.799	0.050	0.824	0.894	161	172.1	0.063	ACO-1
SKD-1	179	59	6	-	-	8	0	-	-	0	-	-	-	-	-	0.485	0.486	0.534	0.465	0.483	65	69.0	0.056	SKD-1
SKD-2	80	124	-	-	-	48	-	-	-	-	-	-	-	-	-	0.000	0.997	0.001	0.978	1.000	124	124.2	-0.000	SKD-2
FDP-2	214	38	-	-	-	0	-	-	-	-	-	-	-	-	-	1.676	0.195	1.629	0.202	0.375	38	35.2	-0.082	FDP-2
UGP-1	144	35	67	-	-	3	1	-	-	2	-	-	-	-	-	6.154	0.013 *	6.019	0.014 *	0.018 *	103	94.4	-0.093	UGP-1
UGP-2	95	88	49	1	-	6	10	0	-	2	1	-	0	-	-	10.572	0.001 **	10.376	0.001 **	0.001 **	149	129.4	-0.153	UGP-2
UGP-3	195	56	-	-	-	1	-	-	-	-	-	-	-	-	-	2.090	0.148	2.024	0.155	0.218	56	51.4	-0.091	UGP-3
FUM-2	145	95	5	-	-	7	0	-	-	0	-	-	-	-	-	4.498	0.034 *	4.380	0.036 *	0.046 *	100	89.4	-0.120	FUM-2
ADH-2	87	123	-	-	-	42	-	-	-	0	-	-	-	-	-	0.018	0.895	0.010	0.920	1.000	123	122.2	-0.008	ADH-2
PGI-2	223	17	8	3	-	0	0	0	-	0	0	-	1	-	-	0.015	0.904	0.024	0.877	0.600	28	28.8	0.025	PGI-2
MDH-1	244	8	-	-	-	0	-	-	-	0	-	-	-	-	-	0.066	0.798	0.057	0.811	1.000	8	7.9	-0.016	MDH-1
MDH-3	210	21	4	10	-	3	3	0	-	0	0	-	1	-	-	10.986	0.001 **	11.364	0.001 **	0.004 **	38	45.6	0.165	MDH-3
MDH-4	230	2	20	-	-	0	0	-	-	0	-	-	-	-	-	0.525	0.469	0.500	0.479	1.000	22	21.2	-0.042	MDH-4
MEAN																					68.4	64.9	-0.0537	MEAN

***Table 19**: Observed numbers of genotypes, observed and expected numbers of heterozygotes, tests for Hardy-Weinberg expectations, and estimated fixation index at 31 loci: Central Oregon (N = 252)*

they may, at least partially, arise from analyzing different loci, different numbers of loci, dissimilar samples of trees, and from using different laboratory techniques. Comparisons of results therefore have to be made with some caution.

The estimates of genetic diversity for ponderosa pine in Oregon are most likely inflated, relative to similar studies, due to the sampling strategy and the methodology used in the electrophoretic survey. Individual trees were sampled over a wide range of sites and habitat conditions, capturing a maximum of ecological (and most likely genetic) variation. All 31 loci used for the investigation were polymorphic, *i.e.* had more than one allele per locus. All loci, which proved to be monomorphic in the preliminary test runs were excluded from the survey. In addition, many of the enzyme systems that were surveyed are known to be highly variable compared to others (MNR, PEP, GOT, PGM, UGP and MDH). All these factors lead to estimates of genetic diversity which are most likely high compared to other studies. Nevertheless, it may be interesting to compare the levels of genetic diversity found in Oregon with results reported for ponderosa pine from other areas, using similar methodology.

ALLENDORF et al. (1982) reported an average polymorphic index of 0.186 for ponderosa pine from Washington, Idaho and Montana[3]. Excluding 3 loci which were monomorphic, a polymorphic index of 0.208 can be calculated from the published data. Sampling design as well as number of loci were similar to the present survey. Unfortunately, no other diversity measures were reported.

An average number of alleles of 1.99, polymorphic index of 0.123 and proportion of polymorphic loci of 56.5% were found for ponderosa pine in eastern Montana by WOODS et al. (1983)[4]. Samples were taken from six small and isolated stands within a radius of 9 km only. The comparatively low diversity estimates most likely are the result of this restricted sampling.

Results of a range-wide survey of genetic variation in ponderosa pine were reported by NIEBLING and CONKLE (1990). Twenty eight loci were surveyed, 19 were polymorphic and 9 monomorphic. Surveyed gene loci were similar but the sampling design differed from the present investigation. For the Pacific race, represented only by two populations, P was 70 % in both populations, N_a was 2.29 and 2.39, and PI 0.137 and 0.150. The North Plateau race was represented with samples covering a much wider area[5]; consequently, as would be expected, average diversity was higher than in the Pacific race (P = 64%, N_a = 2.77, PI = 0.178). Similar diversity values were found within the Rocky Mountain race with values of P = 75%, N_a = 2.77 and PI = 0.164. These values indicate that ponderosa pine maintains relatively high genetic diversity within small geographic areas.

High diversity within a relatively small area was also reported for ponderosa pine in California[6] (CONKLE and WESTFALL, 1984): P = 89%, N_a = 3.0, and PI = 0.210. Sampling design and number of analyzed loci (30) were very similar to the present study. Consequently, the results might be expected to be more or less comparable.

Ponderosa pine in Oregon seems to maintain a high genetic diversity comparable to values reported for geographically adjoining parts of the range. Despite the fact that the overall estimates are most likely inflated by the sampling design, the results suggest diversity levels in the upper range found for conifers. Moreover, genetic variation seems well distributed over the area; much of total diversity is found within local areas (or seed zones) as the low F_{ST} - values and the estimates of average diversity within seed zones indicate.

Minor and major polymorphisms are found in about equal proportions. Rare alleles with frequencies smaller than 5 % are, however, found at most loci (24). These alleles are often considered to contribute to the current genetic load while frequent variants may be considered the operating genetic potential (BERGMANN et al., 1990) under the prevailing conditions. Since the capability of enduring a genetic load is prerequisite to the preservation of adaptability to changing environments (GREGORIUS, 1986), rare alleles may also be conceived as the cur-

---

[3] *based on a survey of 400 trees and 29 loci of which 3 were monomorphic*
[4] *based on a survey of 300 trees and 23 loci of which 6 monomorphic*
[5] *bulked seed from 12 populations representing an estimated 102 diploid genotypes*
[6] *based on diploide genotypes of 524 trees originating from 12 seed zones and different elevational bands*

rent latent genetic potential (STEBBINS and HARTL, 1988), being of utmost importance for future adaptability. Thus, the high diversity found in Oregon ponderosa pine may be perceived as a high potential of adaptability to future environmental conditions.

Although Southwest Oregon showed a slightly higher proportion of polymorphic loci (87.1 % vs. 74.2 %), a higher allelic diversity at 18 loci (with 16 expected by chance), and had 5 rare alleles which were not observed in Central Oregon (compared to 2 region specific alleles in Central Oregon), no significant differences in genetic diversity seem to exist between the two regions.

The ocurrence of low-frequency, area-specific alleles at 7 loci may be a result of several possible causes: insufficient sample size, random processes or past migration history. Since the existence of different races or varieties in the two regions has been suggested by several authors (*see section 2.2*), it would be tempting to view the occurrence of region-specific alleles as supporting evidence for a different evolutionary history of ponderosa pine in the two regions (different refugia, postglacial history). Glacial refugia which have long been isolated should have differed mainly in the occurrence of low-frequency alleles due to mutation, genetic drift and weak selection (assuming that these allelic variants are selectively neutral or only slighly deleterious, so that they are not removed by selection), whereas frequent and functionally important alleles are expected to occur in all regions of the species. Therefore, it may reasonably be assumed that populations having different rare variants at several loci originated with a higher probability from different glacial refugia than populations sharing the same rare alleles (WHEELER and GURIES, 1982; KONNERT and BERGMANN, 1995). The detection of low-frequency alleles, however, largely depends on sample size. According to GREGORIUS (1980) and HATTEMER et al. (1980), the minimum sample size of genotypes required to ensure that an allele at a locus is detected with 95 % probability is $N = 218$ for an allele with a frequency of 1 % and $N = 436$ for a variant with a frequency of 0.5 % (assuming Hardy-Weinberg structure and 4 low frequency alleles at a locus). Since frequencies of the 7 region-specific alleles were all below 0.6 %, our sample size was clearly not large enough to allow for the detection of these alleles with sufficient precision. Consequently, the presence or absence of these alleles in each of the two regions is rather a result of sampling error than a reflection of real differences caused by evolutionary history.

Since gene flow is probably minimal across the natural physiogeographic barrier of the Cascade crest, racial differences, if existent, should be reflected in differentiation of allele frequencies between the east and west sides of the Cascade Range. Allelic frequencies are in fact significantly heterogeneous between the two regions when all 31 loci are considered. The two regions differ on average by a proportion of 6.1% unshared alleles. Moreover, much higher differences exist for some of the loci, reaching values as high as 21.2%. Interestingly, loci which revealed a large differentiation between the Pacific and the North Plateau race in the investigation of CONKLE and CRITCHFIELD (1988), were also highly differentiated in the present investigation (*Idh*, *Adh*, *Pgm*). However, in contrast to the results of CONKLE and CRITCHFIELD (1988), estimated genetic distance between the two regions does not support the existence of two races. With an average of 0.007 for NEI's unbiased genetic distance, differentiation between the regions is small, being in the range of distances commonly found between populations (NIEBLING and CONKLE, 1990). Based on electrophoretic data from a number of studies, NEI (1974) presented a generalized scale of genetic distances and reported that species were characterized by distances of 0.1 to 1.0, subspecies and varieties by 0.02 to 0.2 and races by 0.01 to 0.05. While such a scale alone cannot be used to determine taxonomic relationships, it can provide a guideline for evaluating electrophoretic data. According to this scale, the estimated differentiation is too small to support the existence of different races in the two regions. Moreover, differentiation is no higher among the regions, as would be expected in the case of different races, than among populations within regions. A similar conclusion is reached from results of F - statistics. According to F - statistics, overall differentiation among regions is very small; the fraction of total gene diversity which may be attributed to regional differences was only 1%. This is substantially smaller than the fractions attributable to differences among populations within regions, which were estimated as 7% for Central Oregon and as 8% for Southwest Oregon (although within-region estimates of differentiation probably contain much noise generated by the small sample sizes).

Based on the results from univariate measures of differentiation, overall allelic structures seem

rather similar in both regions. With respect to isozyme diversity, the results do not seem to support the existence of different races in the two regions. Conclusions from univariate isozyme analyses seem to contrast with results from monoterpene analyses which found a clear racial differentiation between the east and the west side of the Sierra Nevada and the Cascade Mountains. Contrasting conclusions based on isozyme and monoterpene data have already been reported by NIEBLING and CONKLE (1990) and CONKLE and CRITCHFIELD (1988).

Univariate measures of differentiation, however, have their limitations. Therefore, some caution must be exercised when drawing conclusions. Both *F* - statistics and NEI's genetic distance have clear limitations in portraying population differentiation. WRIGHT's (1965) method of partitioning the genetic variation to different hierarchical levels relies on estimates of genotypic frequencies. The precision of such estimates depends upon the number of individuals sampled, the number of populations and their distribution as well as the loci assayed per population. Variation in intra- and inter-locus estimates are reduced as sample size increases. Estimates of within region F_{ST} - *values* are thus expected to be upwardly biased due to the small sample sizes of the artifically formed populations. Moreover, the univariate approach of measuring population differentiation using WRIGHT's (1965) or NEI's (1973) method which weights all loci equally by taking the mean over all loci and disregards relative subpopulation size has been proven inadequate (GREGORIUS, 1978). In several studies, comparing *F* - statistics with differentiation according to GREGORIUS and ROBERDS (1986), subpopulation differentiation δ was substantially higher than the F_{ST} - values (6.1 versus 2.4 (GREGORIUS and ROBERDS, 1986); 3.1 versus 0.4 (GREGORIUS et al., 1986); 4.9 versus 1.96 (MÜLLER-STARCK and GREGORIUS, 1986)), leading to different conclusions. In all cases, differentiation measures according to GREGORIUS and ROBERDS (1986) were at least 2.5 times higher than the F_{ST} - values, perhaps indicating the greater sensitivity of the former method. It is noteworthy, however, that F_{ST} - values are not directly comparable with δ since F_{ST} depends on the amount of total variation present. If there is a lot of variation within populations relative to variation between populations, F_{ST} will be small, but subpopulation differentiation may still be high.

Likewise, NEI's genetic distance seems defective in theory as well as in practice, at least as far as its application to forest tree species is concerned (FALKENHAGEN, 1985). Data published by BERGMANN (1973) and LUNDKVIST and RUDIN (1977) clearly demonstrate, that NEI's distance is not a metric and cannot be interpreted in terms of provenance or racial variation. Undesirable properties as well as difficulties in interpretation of NEI's distance measure have also been pointed out by GREGORIUS (1974).

Subpopulation differentiation δ according to GREGORIUS and ROBERDS (1986), on the other hand, has clearly defined properties and the results may be interpreted directly as proportions of genetic disparity *i.e.* as the mean proportion of the effective numbers of genetic elements by which subpopulations differ from their complements. Estimated overall differentiation among the regions was in fact substantially higher based on δ than on F_{ST} (6.1 % vs. 1 %). Subpopulation differentiation was still small, however, relative to differentiation among populations within regions (δ = 5.5% and δ = 7.7%). Even if the within-region estimates probably are upwardly biased due to sampling design, differentiation among regions does not seem to be much higher than differentiation within regions. Although there are no references to validate the 6.1% in terms of racial differences, results of overall differentiation do not seem to support the existence of different races in the two regions since the ratio of among to within region differentiation argues against it.

It must be emphasized, however, that all presented measures of differentiation are composite indices, weighting all loci equally and averaging the differences over all loci. They equalize differentiation over all loci, which may not be adequate to describe the real situation. Although the different measures of differentiation vary somewhat in their sensitivity, they all are not apt to reflect differentiation completely, especially regarding differences among the various gene loci. Patterns of differentiation may differ at individual loci which cannot be reflected in a composite index. With respect to racial differentiation, it is highly unlikely that differences would occur equally at all loci. Whatever the forces which lead to racial differentiation (different selection pressures, genetic drift, bottlenecks, introgression), their effects probably affect individual gene loci differently, creating different degrees of differentiation.

Single locus F_{ST} - values as well as the subpopulation differentiation measures δ were in fact highly variable between loci. Several loci exhibited high differentiation among the regions while others lacked differentiation. Thus, despite the observed small overall differentiation between regions, major differences were found at the single locus level.

Differentiation in allele frequencies at some loci were interpreted as evidence for the existence of races in ponderosa pine by CONKLE and CRITCHFIELD (1988) and NIEBLING and CONKLE (1990). While the Pacific and North Plateau races were closely aligned, exhibiting only moderate differences in allele frequencies, Rocky Mountain ponderosa pine (*var. scopulorum*) was highly differentiated at several loci from Pacific ponderosa pine (*var. ponderosa*). A phylogenetic tree based on allozyme frequencies separated Rocky Mountain ponderosa pine by a long genetic path (with a distance of 0.14 from the root) from Pacific ponderosa pine, suggesting a long divergence and distinct evolution from the root (NIEBLING and CONKLE, 1990). Distinct morphological differences, divergent monoterpene types in the south and a west-east differentiation in allozyme variation at several loci suggest that the two varieties (*ponderosa* and *scopulorum*) may have evolved from two distant progenitor lines, one perhaps from a west coast, the other from a more central progenitor in Mexico. Consequently, the genetic distinctness of Rocky Mountain ponderosa pine is most likely the result of a long and different evolutionary history which was conserved by restricted gene flow across the Continental Divide. On the basis of a neutral gene model (NEI, 1975), a rough approximation of the time of divergence between the two varieties *var. scopulorum* and *var. ponderosa* can be estimated, using the formula

$$t = 5x10^6 x D$$

where t is the divergence time in years and D is the average genetic distance between taxa.

Based on the genetic distances reported by NIEBLING and CONKLE (1990), estimated time of divergence between the varieties is between 300'000 and 400'000 years. This approximation suggests that the two varieties have been in existence for a long time, having diverged already during the late Pleistocene.

In contrast, separation of the two races within Pacific ponderosa pine (*var. ponderosa*) seems to be quite recent. North Plateau race exhibits a shorter distance from the root than the Pacific race, suggesting that it may have been the progenitor of the Pacific race (NIEBLING and CONKLE (1990). Thus, inferred from allozyme differentiation, the two lines seem to have a long common evolutionary history. Since the species has a relatively short history of immigration (3'000 to 6'000 years), it may be speculated that division of Pacific ponderosa pine into two races began only recently and that it is still in progress. Differentiation is most likely the result of the extreme geographic barriers separating the two areas of distribution. Restricted gene flow across these physiographic barriers combined with different selection pressures due to climatic differences are most likely the driving forces in this differentiation process. Due to extended population size in both areas, random processes are not likely to play a major role. Under these assumptions, any observed differentiation is expected to rather reflect climatic differences between the regions, *i.e.,* current adaptation to different environments rather than historic events or random processes. Regarding morphological traits, the Pacific and North Plateau races are closely aligned. Needle hypoderms are thicker in the North Plateau race and trees are more cold-hardy than trees of the Pacific race. These traits combined with earlier initiation of growth and slower growth rate of North Plateau seedlings, strongly support the assumption that adaptation to different environmental conditions is a major force of differentiation. Nevertheless, other forces may have contributed partially to observed differentiation as well. Selection, for example may only operate with genetic variants that are present in the species or populations; thus, differences in genetic makeup due to a different evolutionary history may affect adaptation to present environments, producing different patterns of adaptation for different species under equal environmental conditions (REHFELDT, 1979, 1980, 1983b).

Approximate time of evolutionary divergence, based on the average genetic distance of 0.007 between the areas, is estimated as 35,000 years which underlines the very recent divergence between the Pacific and North Plateau races. Since only polymorphic loci were assayed and used for calculating average genetic distance D (while D is supposed to be an average over all

loci, including monomorphic loci), time of evolutionary divergence is most likely overestimated. Moreover, average genetic distance *D* is only an estimate with a certain error attached to it. This may be one of the reasons why the estimated time of divergence does not fit with the elapsed time of 3'000 to 6'000 years since immigration into the area. Two alternative explanations for this discrepancy seem plausible. Divergence may have occurred before ponderosa pine colonized the area, arguing for two different refugia in *var. ponderosa* and supporting the occurrence of two races. Secondly, since NEI's evolutionary interpretation of his genetic distance assumes neutrality, selection and migration are considered as negligible. The observed discrepancy may then be interpreted as an argument against neutrality of the gene markers, not only suggesting adaptivity of certain genes but also supporting the assumption that selection plays a major role in shaping today's observed genetic structures.

Deciding on the taxonomic status of species, varieties or races involves some methodological problems which may give rise to controversial conclusions. Races, according to ZOBEL and TALBERT (1984), are groups of populations that generally interbreed with one another and that intergrade more or less continuously. From this definition, several fundamental problems separating races become apparent. To divide continuously intergrading populations into different taxonomic units, a subjective decision is needed regarding the amount of differentiation necessary for taxonomic separation. Moreover, a number of different characteristics may be used to estimate differentiation. Although NEI (1974) has presented a generalized scale of genetic distances for allozyme differentiation which may help to answer taxonomic questions, the results may be controversial compared to those based on other characteristics such as morphological traits, other biochemical markers or crossability results.

The case of Washoe pine may serve as an example to illustrate the problems involved in taxonomy of species, subspecies, varieties and races. Washoe pine (*Pinus washoensis Mason and Stockwell),* an endemic pine species in subsection ponderosa, is considered to be a separate species, but a close relative of ponderosa pine. Washoe pine's monoterpene and morphological traits are within the ranges of variation for ponderosa pine's North Plateau and Rocky Mountain races, but Washoe pine differs notably from the Pacific race (CRITCHFIELD, 1984). Controlled test crosses established a strong and direct evolutionary relationship between Washoe pine and the Rocky Mountain races of ponderosa pine since interspecies crosses produced more sound seed per cone than crosses within Washoe pine itself. With respect to allozyme diversity, however, Washoe pine is highly differentiated from Rocky Mountain ponderosa pine (with an average genetic distance of 0.066) but genetically close to the Pacific race (with an average genetic distance of 0.013). As inferred from isozymes, Washoe pine is phylogenetically placed in a lineage closely related to variety *ponderosa*, but is clearly distinct from variety *scopulorum* (NIEBLING and CONKLE (1990).

As a further illustration of the difficulties in determining taxonomic relationships from electrophoretic data, some comparisons of genetic distances between species, varieties and races may be looked at. The relatively large genetic distances between the Rocky Mountain and Pacific races of ponderosa pine, reported by NIEBLING and CONKLE, (1990), are within the range of distances reported for varieties by NEI (1974). However, the estimated genetic distances from the Rocky Mountain race, ranging from 0.06 (Pacific race) to 0.082 (North Plateau race), exceed some of the distances which were reported for interspecies comparisons in other conifers: for example, genetic distance between *Pinus clausa Chapm.* and *Pinus virginiata Mill.* has been estimated as only 0.014 (WHEELER et al., 1983), and between *Abies balsamea var. balsamea (L.) Mill.* and *Abies fraseri (Push) Poir.* as only 0.060 (JACOBS et al., 1984).

Hence, decisions regarding the taxonomic status of a taxa involve several difficulties and some complex and controversial aspects which are rather more fundamental than specific for gene markers. Moreover, due to the limitations of univariate measures of differentiation, the question regarding the existence of different races in the two regions, as a consequence of a different evolutionary history, cannot be answered. The small average differentiation between the two regions argues against the existence of races. In contrast, allele frequency differences at several loci which exist between the two regions argue in favor of a possible racial differentiation. Allele frequency differences may, however, not only be the result of a separate evolution in the past; they may also be a result of natural selection in different environments or a consequence of random genetic process.

Since climate not only differs substantially between the two regions but varies considerably within each region (*see section 3.5.2*), genetic differentiation among as well as within regions is expected to reflect this environmental variation if selection were a major force of differentiation. Despite the fact that the highest proportion of total variation in temperature (climate *Factor 1*) is found among the regions (64%), 36% of total variation is due to differences among locations within regions (20.4% of it within Central Oregon, 15.2% within Southwest Oregon) (*see Table 7, p. 45*). The largest proportion of variation in water balance (climate *Factor 2*) is found within Central Oregon (68.9%), while 31% is found within Southwest Oregon. Only 0.1% of total variation is due to differences between the two areas. For both climate factors combined, climate is most variable within Central Oregon, accounting for an average proportion of 44.6% of total variation. Variation within Southwest Oregon accounts for 23.1% of total variation, which is only half the amount found within Central Oregon. The difference in climate conditions between the two areas accounts for 32.3% of total variation in temperature and water balance.

Although estimates of within-region differentiation are expected to be associated with large errors due to the small and unequal sample sizes, the amounts of differentiation within regions, calculated as subpopulation differentiation δ among all 37 pairs of seed zones and averaged over all 31 loci, parallel the existing differences in climate conditions, suggesting that genetic differentiation and environmental differentiation may be associated. Highest average within region differentiation was found in the climatically more variable area of Central Oregon (on average 7.7% of the alleles differed among the seed zones), while genetic differentiation within Southwest Oregon was lowest (5.5%), probably reflecting the lower environmental variation in this area. Finally, average genetic differentiation among the regions was intermediate (6.1%), paralleling the intermediate variation of climate conditions found among the two regions. It must be emphasized, that genetic differentiation within Southwest Oregon is actually smaller than the average value of 5.5% indicates since seed zone 90 has an extremely high genetic differentiation from all the other zones. The outstanding genetic structure of this zone is reflected in an average genetic distance of 0.084 to all the other zones within the area. This distance equals the distance found between the varieties *var. scopulorum* and *var. ponderosa*. Except sampling error (zone 90 contained only 7 individuals), we know of no reason for this outstanding genetic structure of this zone. Some Jeffrey pines (*Pinus jeffreyi Grev. & Balf.*) may have been erroneously sampled as ponderosa pines, since both species are common in this area and are difficult to distinguish (in fact, several Jeffrey pines were sampled by mistake; possibly not all errors were detected).

The average unbiased genetic distance by NEI among all pairs of seed zones within Central Oregon equaled the average distance found between the two regions while Southwest Oregon had less differentiation among the zones. When seed zone 90 was excluded, average genetic distance within Southwest Oregon was only half the distance found within Central Oregon and half of what was found among the regions. These results are in line with the results obtained from analyses using subpopulation differentiation δ. Results from F - statistics, on the other hand, estimating amounts of differentiation among and within regions, do not correspond to environmental variation. As discussed before, results from F -statistics are most likely the least liable to describe differentiation under the given circumstances. Thus, results from F - statistics will not be discussed any further.

Patterns of differentiation emerging from cluster analyses either on NEI's unbiased genetic distance or on subpopulation differentiation δ among all pairs of seed zones, seem to confirm the existence of associations between genetic structures and habitat conditions. While results differed according to the fusion strategy used to form the clusters (*for fusion strategies see 3.6.4.2*), dendrograms from complete linkage cluster analyses revealed patterns of differentiation which support such a relationship. UPGMA clustering, on the other hand, produced less distinct patterns of differentiation, which were not readily apparent nor easily interpretable. Although complete linkage is a space dilating strategy, changing the structure of the multidimensional data cloud, it minimizes dissimilarities between the entities of each group and thus produces distinct and homogeneous clusters. Hence, complex patterns of genetic differentiation become more easily recognizable than with other clustering allgorithms. Consequently, dendrograms resulting from complete linkage clustering most probably illustrate the complex patterns of genetic differentiation better than cluster diagrams using the UPGMA technique. Results of

average linkage cluster analyses were presented mainly for comparison, since UPGMA clustering has been widely used to illustrate patterns of genetic differentiation. Discussion, however, will focus on the results of complete linkage cluster analyses only.

Considering the small sample sizes which probably lead to large errors associated with parameter estimates, cluster diagrams of genetic distance among seed zones are expected to show only major trends. The dendrogram resulting from complete linkage clustering of seed zones based on NEI's unbiased genetic distances among zones showed a more or less distinct separation of the two regions (*Figure 14, p. 75*). Seed zone 661 clustered with the southwestern zones. This result may reflect a Pacific Coast climatic influence at this inland sites since the Columbia river gorge forms the only corridor through the Mountain range (SORENSEN, 1994). Except sampling error, we have no other plausible explanation for the other 3 zones from Central Oregon which clustered with southwestern zones. The more homogeneous climate conditions found within Southwest Oregon seem to be reflected in a smaller genetic differentiation in this area. Likewise, the more variable habitat conditions prevailing in Central Oregon are matched by a higher genetic differentiation. Furthermore, the pattern of differentiation within Central Oregon showed a trend of similar genetic structures in the northern and southern parts while the central part was distinct. This pattern seems to match variations in temperature, with harsher conditions and a higher degree of variation in the center of the area compared to the northern and southern parts (*see section 3.5.2, Figure 7*) .

Complete linkage cluster analysis of subpopulation differentiation δ between all pairs of seed zones revealed similar but even more distinct patterns of differentiation (*Figure 16, p. 77*). Highly distinct genetic structures were manifest for both regions. Seed zones of Southwest Oregon were clearly separated from zones in Central Oregon. Again, little differentiation was observed within Southwest Oregon while four genetically distinct groups were clearly discernible in Central Oregon. The distinctness of allelic structures of the five main clusters could be confirmed by descriptive discriminant analysis. It is interesting to note in this context that cluster formation based on simple univariate measures of differentiation proved to be rather effective in uncovering patterns of differentiation, in spite of the limitations of such composite indices. Roughly 40% of the total variation of all 71 original allozyme variables were associated with group differences, indicating that major patterns of differentiation were most likely revealed by cluster analysis using simple univariate measures of differentiation. This conclusion is also sustained by results from predictive discriminant analysis. The distinct genetic structure of the clusters was confirmed by the relatively high classification accuracy of individual trees into their initial clusters, using discriminant functions. 48% of all individuals were correctly assigned to their initial clusters, based on the most discriminating alleles (22). On average, classification was 28% better than would be expected by chance alone. Very distinct clusters, such as the group encompassing trees of Southwest Oregon, had classification rates as high as 40% above their prior probability.

If natural selection is playing a major role in shaping genetic variation, associations between genetic and environmental variation are expected. Our results seem to indicate such associations. However, results of within-region differentiation are only crude estimates which most likely are associated with large sampling errors. The larger within-region differentiation observed in Central Oregon may primarily be a reflection of a larger sampling error due to smaller sample sizes in this region. Moreover, the observed difference between the two regions, if real, is still very small (5.5 % vs. 7.7 %). Furthermore, other plausible explanations are conceivable for associations between genetic and environmental variation. For example, greater climate variation in Central Oregon may mean more extreme environments resulting in patchy distributions of ponderosa pine, either now or in the past, resulting in more genetic drift and less gene flow between populations. Hence, our results are non conclusive regarding the possible underlying forces which are responsible for the observed patterns of differentiation. According to our results, natural selection is a plausible but not the only possible explanation for the observed patterns of differentiation.

Analysis of genotypic structure is strongly limited due to the sampling design. Since no population samples were available, genotypic structures are expected to reflect primarily sampling design rather than real population genetic effects. While overall genotypic structure was in equilibrium, deviations from Hardy-Weinberg expectations were manifest at several loci. Be-

cause significant deviations mainly occurred at loci with abundant differentiation within and among the areas, the deviations might, at least partly, be a result of the Wahlund effect. Mating normally occurs within sub-populations. If sampling is carried out on a grid over the area, the individuals come from many different sub-population (demes). For alleles with regional differentiation, showing higher frequencies in some demes than in all other demes, the alleles cannot be combined randomly into individual genotypes since mating does not occur at random over the total area. Hardy-Weinberg equilibrium, however, may only be attained in the case of random mating among all individuals. As a consequence, less heterozygote individuals than expected will be observed. The Wahlund effect creates heterozygote deficiencies which vary significantly among the loci, because allele frequency differences among demes, causing the effect, differ for each locus. In contrast, inbreeding which may also cause deficiencies in heterozygous individuals, affects all loci equally; *F* - values therefore should not vary much among the loci.

For total area, a deficiency of heterozygous individuals was manifest at 17, an excess at 16 loci. *F* - values varied substantially among the loci. From the 14 loci which showed significant allele frequency differences between the two region, 10 showed a positive, 4 a negative fixation index, suggesting that the Wahlund effect may indeed partly be responsible for observed deficiencies in heterozygous individuals. Within regions, sampling error probably is the most plausible explanation for the observed deviations from Hardy-Weinberg expectations. Genotypic structures, therefore, will not be discussed in more detail.

4.5 Summary

Ponderosa pine in Oregon maintains a high genetic diversity which is comparable to values reported for geographically adjoining parts of the range. Despite the fact, that the overall estimates are most likely inflated by the sampling design, the results suggest diversity levels in the upper range found for conifers. Moreover, genetic variation seems well distributed over the area; much of total diversity is found within local areas (or seed zones) as low F_{ST} - values and the estimates of average diversity within seed zones are indicating. Minor and major polymorphisms are found in about equal proportions. Rare alleles with frequencies smaller than 5 % are, however, found at most loci (24). The high genetic diversity may be perceived as a high potential of adaptability to the extreme and variable habitat conditions which prevail in the area. Although Southwest Oregon had a slightly higher mean number of alleles per locus a higher proportion of polymorphic loci and showed 5 rare alleles not found in Central Oregon, differences in genetic diversity between the two regions were small and non significant. Moreover, observed differences most likely are rather reflecting the unequal sample sizes (leading to large sampling error) than real differences.

Allelic frequencies are significantly heterogeneous between the two regions when all 31 loci are considered. The two regions differ on average by a proportion of 6.1% unshared alleles. With an average of 0.007 for NEI's unbiased genetic distance, differentiation between the regions, however, is rather small, being in the range of distances commonly found between populations in other studies of this species and in other conifers. Differentiation is no higher among the two regions than among populations within regions (within-region estimates are however upwardly biased due to small sample sizes). The small average differentiation between the two regions argues against the existence of different races (Pacific race, North Plateau race) in the areas east and west of the Cascade Range. Methodological limitations of composite measures of differentiation as well as fundamental difficulties involved in decisions on the taxonomic status of taxa do not allow a final conclusion, however.

In contrast, allele frequency differences at several loci which exist between the two regions would argue in favor of a possible racial differentiation. Allele frequency differences may, however, not only be the result of a separate evolution in the past; they may also be a result of natural selection in different environments or a consequence of random genetic processes.

Patterns of differentiation, emerging from cluster analysis on genetic distance by NEI or GREGORIUS, suggest that genetic and environmental variation parallel each other. Such associations between genetic and environmental variation are expected if natural selection plays a major role in shaping genetic variation. Results from univariate genetic analyses are non con-

clusive, however. Natural selection is a plausible but not the only possible explanation for the observed patterns of differentiation.

To gain further insight into the processes leading to observed differentiation, associations of allozyme variation with present climatic conditions will have to be analyzed in order to estimate the degree of adaptation to the present environment and to separate this pattern from historic events and random processes. The question regarding the existence of different races in the areas east and west of the Cascade Range will therefore have to be addressed again in later chapters and under different perspectives, using different analytical tools.

5. Patterns of genetic variation - spatial structures and environmental associations of single alleles

5.1 Spatial patterns of single locus allele frequencies, inferred from spatial autocorrelation analysis

The study of spatial structures of protein polymorphisms is a well established field of gene-ecology (HARTL and CLARK, 1989) and several statistical approaches are available to describe and test for patterns of geographic variation of allele frequencies. One of the most powerful among such tools is spatial autocorrelation analysis. Differentiation of allele frequencies over a geographic space subsumes two different concepts which need to be distinguished - statistical heterogeneity and pattern (SOKAL and ODEN, 1978b). Statistical heterogeneity of values does not necessarily imply geographic patterns *i.e.* a departure from random spatial arrangement of these values. Statistically heterogeneous values, such as allele frequencies, can be randomly positioned in a geographic space. Thus heterogeneity and spatial patterning are potentially independent and geographic patterns must be examined separately. Consequently, heterogeneities in allele frequencies found between the two areas as well as among the seed zones within each area do not necessarily imply spatial structures of allele frequencies which depart from random spatial arrangements. Spatial autocorrelation analysis, however, provides a description of allele frequency variation in space. The analysis is independent of preliminary assumptions about the underlying population structure. Despite the fact that a description of patterns does not necessarily imply the understanding of the causes of such variation, spatial structures of allele frequencies may yield useful information about the processes which may have been involved in generating the patterns.

Spatial autocorrelation analysis was performed on the basis of single alleles; 96 alleles were analyzed separately. Hence, spatial autocorrelation is not a multivariate but a multivariable approach which may be regarded as a repeated application of univariate methods (SOKAL, 1979b). Autocorrelation of allele frequencies was analyzed using *Moran's I* coefficients (*section 3.6.1.2*). For all pairs of individuals, geographic distance in kilometers was calculated and then classified into five distance classes. Deviation of *I* values from random expectations *E(I)* were then tested for significance at $p < 0.01$ and $p < 0.05$. High positive values of *I* indicate that allele frequencies of individuals are similar within the distance class considered, *i.e.*, that they deviate significantly from the mean in the same direction, while negative spatial autocorrelation coefficients signify dissimilarities in allele frequencies among individuals within the respective distance class.

Results of spatial autocorrelation analysis over total sampling area are summarized in the correlograms shown in *Figure 22, p. 100, 101 and 102*. Correlograms are only shown for alleles which displayed a significant spatial structure. For *Gdh* and *Adh*, only one allele is illustrated since the alternative allele carries the same spatial information. Of the 96 alleles tested, 26 showed a significant spatial structure in their allele frequencies, *i.e.* frequencies were significantly (p at least < 0.05, mostly $p < 0.01$) autocorrelated over most distance classes. Spatial patterns which deviated from random arrangement were manifest for one to three alleles at each of 15 loci *i.e.* at *Mnr-1*, *Mnr-2*, *Lap-2*, *Pep-1*, *Pep-3*, *Got-1*, *G6p-2*, *Acp-1*, *Gdh-1*, *Idh-1*, *Pgm-1*, *Aco-1*, *Skd-2*, *Fdp-2* and *Adh-2*. Of the 26 allele frequencies, 17 exhibited a monotonically decreasing trend over all five distance classes (e.g. starting from high positive values and ending at high negative values). Such a pattern is known as a "cline". Clinal structures are likely the result of either isotropic population movements or some form of differential selection. Six allele frequencies were characterized by a positive *I* value in the first distance class, a near-zero non-significant value in the second, negatively significant *I* values in the third and fourth class and again a near-zero non-significant value in the last distance class. Such a spatial structure is called a "depression", and it is likely to represent the result of a circular cline or a double cline that runs in opposite directions (for example most western and eastern demes are similar, but differ from the central ones). Finally, 3 alleles were characterized by a "crazy quilt" pattern *i.e.* high values of *I* were surrounded by low values without any clear ordering.

Patches are homogeneous areas in the surface of variables such as allele frequencies. Patchy

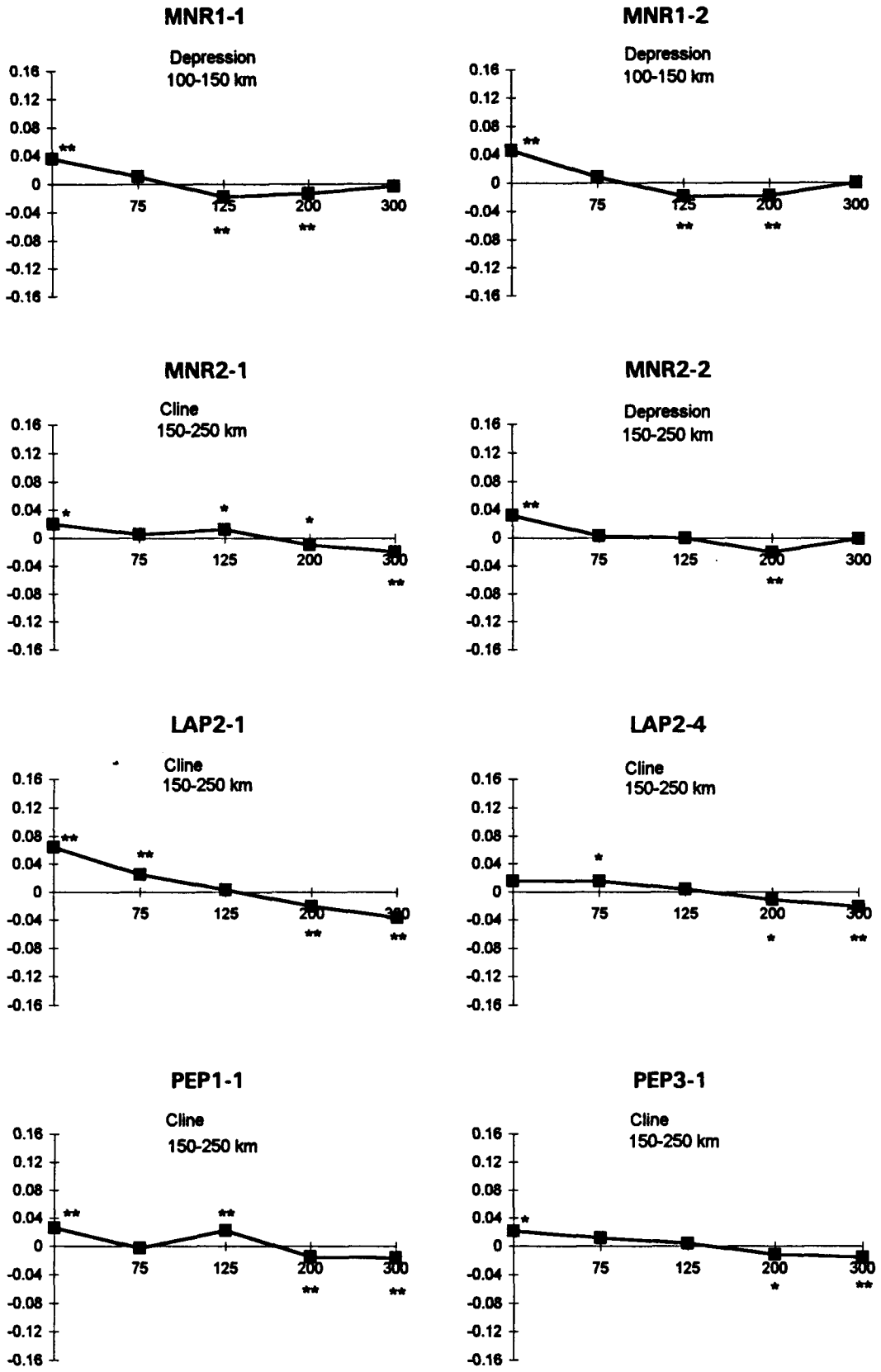


Figure 22: Correlograms of alleles with significant spatial structures: Total sampling area. X - axis: Mean values of the 5 distance classes in km; Y - axis: Moran's I coefficients (significance: p=0.01 (**) and p=0.05 (*)). Patterns and patch sizes are indicated.

Figure 22 (continued)

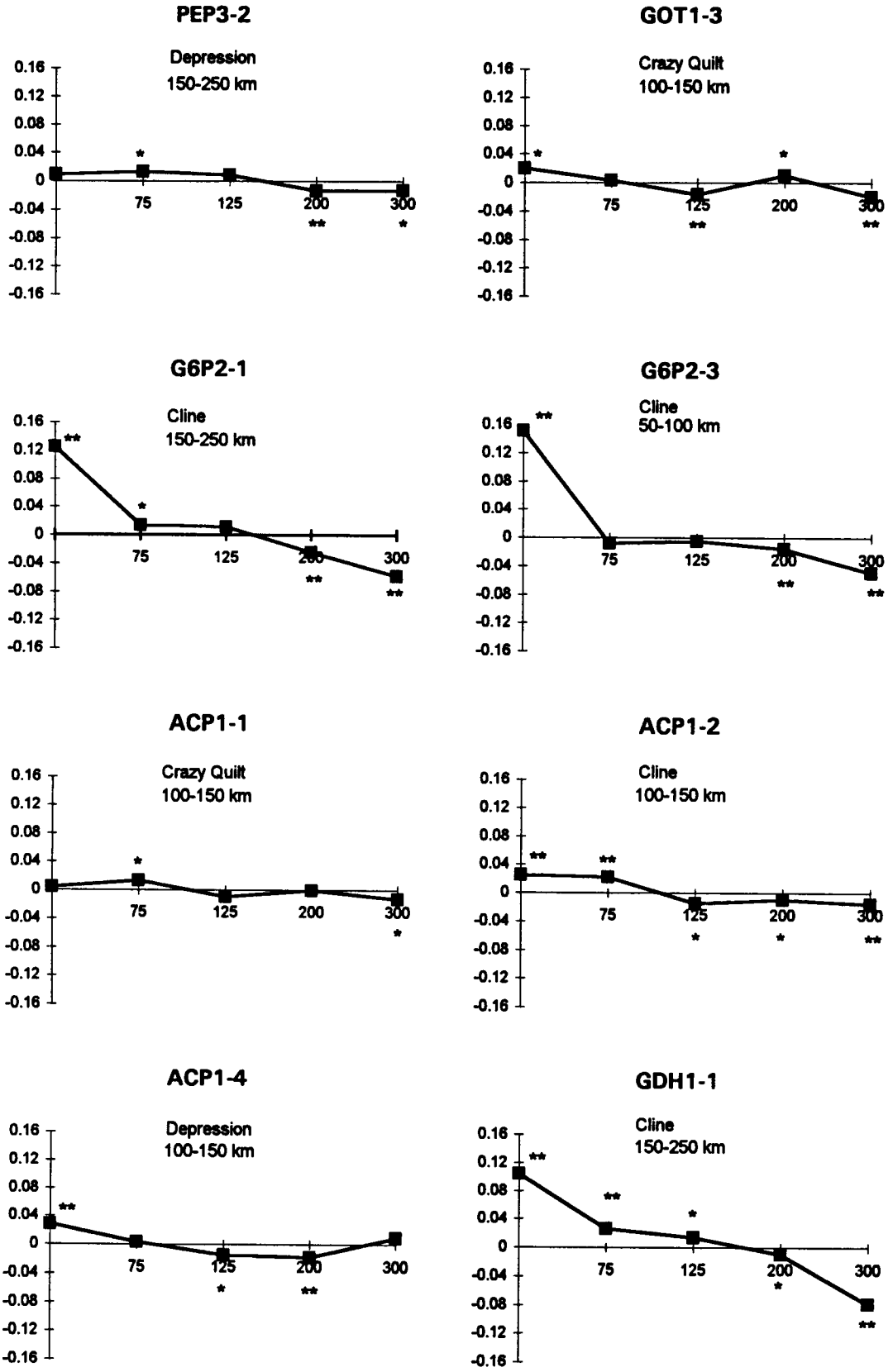
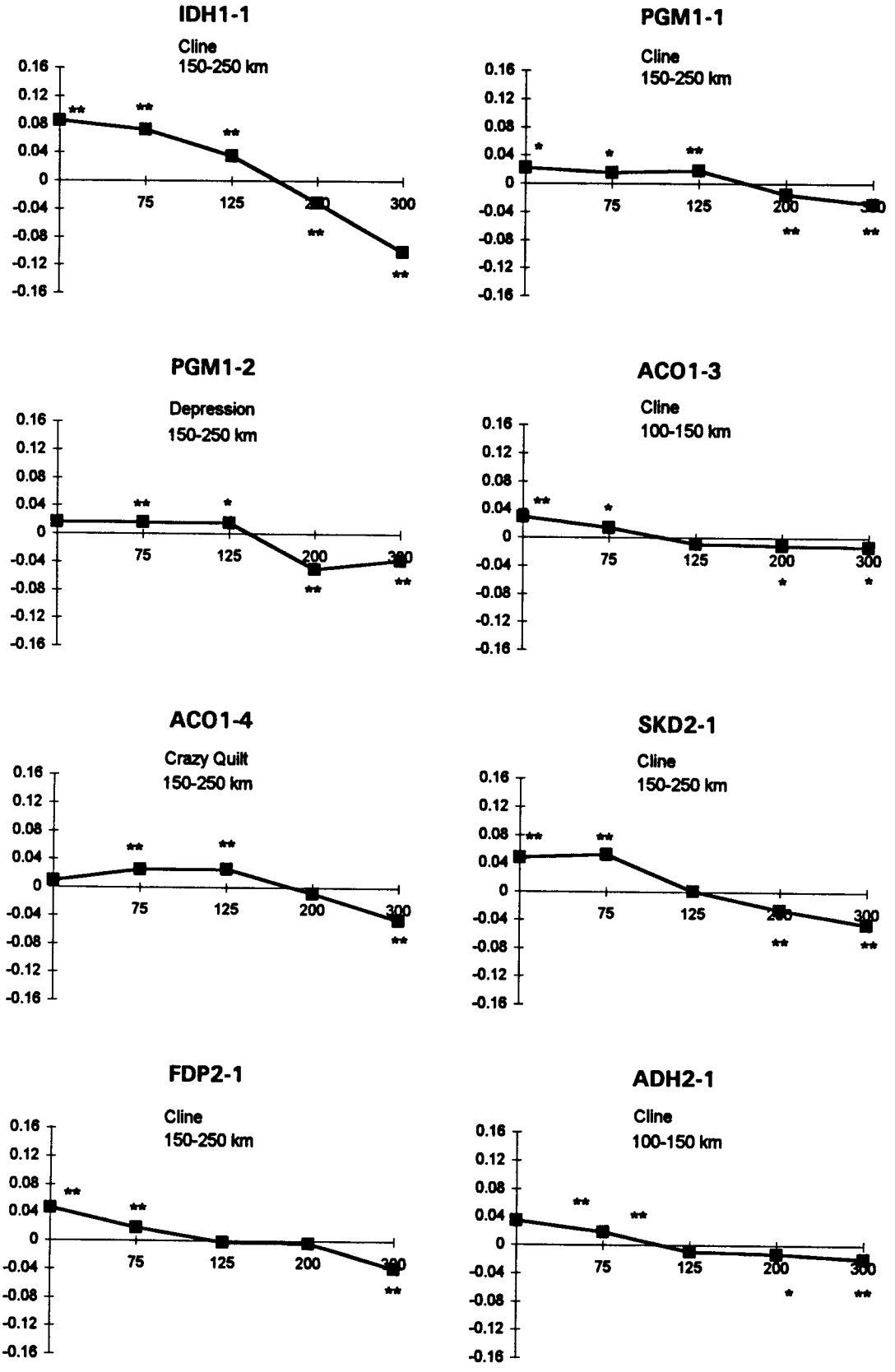


Figure 22 (continued)



structures are detected by spatial autocorrelation only if the spacing of the sample locations is less than the average diameter of the patches and if the chosen distance classes are smaller than patch size. SOKAL (1979b) has found that spatial correlograms indicate patch size (*i.e.* the diameter of the homogeneous area) by the distance class showing the first negative or zero autocorrelation.

Alleles with spatial structures showed substantial differences in the amount of autocorrelation, especially in the first and last distance classes. High values of *I*, indicating similar allele frequencies within patches smaller than 50 km and thus suggesting a rather small scale structure (for example due to adaptation to specific habitat conditions), were found for several alleles such as *G6p2-1*, *G6p2-3*, *Gdh1-1*, *Idh1-1*, *Skd2-1*, *Fdp2-1* and *Adh2-1*. Substantial dissimilarities within patches greater than 250 km, *i.e.* extended variation on an area-wide scale, were manifest for *Idh1-1*, *G6p2-1*, *G6p2-3*, *Gdh1-1*, *Skd2-1* and *Aco1-4*.

With only few exceptions, patch sizes of the significantly autocorrelated allele frequencies were in the range between 150 and 250 km (*Figure 22, p. 100, 101 and 102*), suggesting that most variation in allele frequencies is associated with the two regions. The sampled area in Southwest Oregon extends over 130 km in the west-east direction while the Central Oregon sample covers about 250 km from west to east. Thus, patch size seems to parallel the two areas fairly well, exhibiting more homogeneous allele frequencies within the two areas (*i.e.* patches) relative to the overall variation (*i.e.* the sum of the two areas). This conclusion is further substantiated by the fact that all loci, exhibiting spatial structures for one or more alleles, showed significant differences in allele frequencies between the two regions (with the exceptions of *Got-1*, *Pgi-2* and *Mdh-1*).

Allele frequency patterns over the total area differed substantially among the alleles which showed spatial structure. Six examples of such variation patterns are illustrated in *Figure 23, p. 104, 105 and 106*, where patterns of variation are depicted as allele frequency surfaces, which were smoothed using spline interpolation among the individuals.

Spatial structures of allele frequencies were also analyzed for both regions separately. With the exception of smaller intervals chosen for the distance classes, the same methodology was used for the analyses (*see section 3.6.1.2*).

Correlograms, summarizing results of autocorrelation analysis restricted to Southwest Oregon samples only, are shown in *Figure 24, p. 108 and 109*. Within Southwest Oregon, significant spatial patterns of allele frequencies over most distance classes were manifest for 16 alleles at each of 11 loci *i.e.* at *Mnr-1*, *Lap-2*, *Pep-1*, *Pep-4*, *Got-1*, *Got-3*, *G6p-2*, *Acp-1*, *Gdh-1*, *Skd-1* and *Fum-2*. Of these 16 alleles deviating from random spatial arrangement, 8 exhibited a depression structure, 5 showed a clinal pattern, while 3 were characterized as a crazy quilt pattern (*Figure 24, p. 108 and 109*).

Alleles with spatial structures showed substantial differences in the amount of autocorrelation in the first and last distance classes. High autocorrelation coefficients within patches smaller than 25 km, indicating a small scale patchy structure, were found for several alleles, especially for *Got1-1*, *Got1-3*, *Skd1-3*, *Gdh1-1*, *Acp1-4*, *G6p2-1* and *G6p2-3*. The distance class greater than 150 km was characterized by marked negative autocorrelations for most of the 16 alleles. Especially high negative values, indicating a substantial dissimilarity in allelic structure among individuals separated by more than 150 km, were observed for *G6p2-1* (-0.70), *G6p2-3* (-0.53), *Acp1-4* (-0.30), *Skd1-3* (-0.25) and *Mnr1-2* (-0.21). However, as may be seen from the correlograms, autocorrelation values for most of the alleles changed rather drastically between 125 and 175 km. Hence, allelic structures, being rather homogeneous in an area up to 125 to 150 km in diameter, shifted quite abruptly when distances among trees exceeded 150 km. Since sampling covers a total distance of about 130 km between the Pacific coast and the Cascade Range, the fifth distance class focuses on the diagonal extension of the area, *i.e.*, grouping the trees that are located at the southwestern-northeastern or at the northwestern-southeastern borders of the area. As may be seen from the variation surfaces (*Figure 25, p. 110, 111 and 112*), the abrupt changes in allelic structures in the last distance class indicate in part the existence of a small zone with distinct genetic makeup in the southwestern corner of the area (see

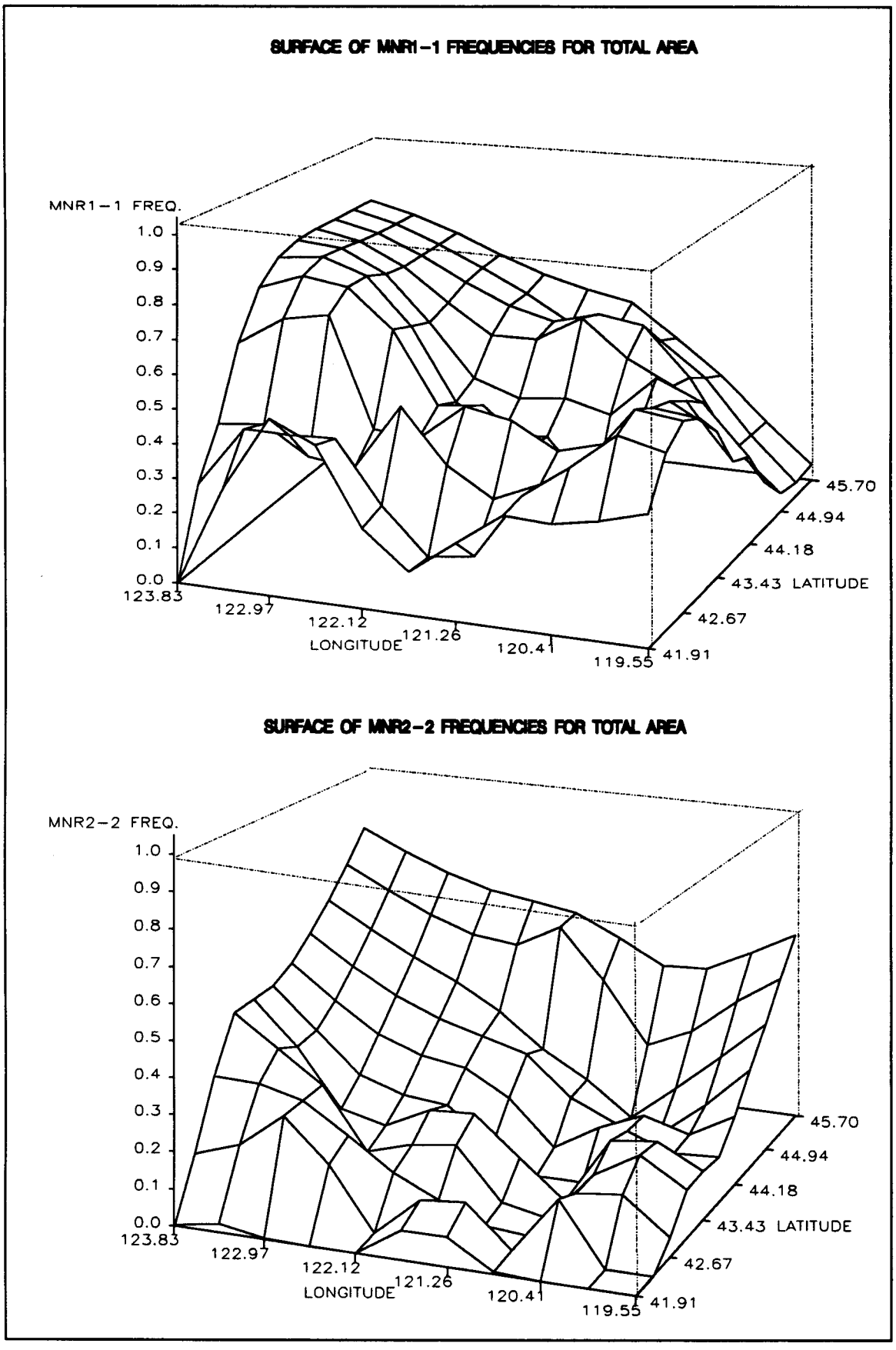


***Figure 23**: Examples of variation patterns, illustrated as smoothed surfaces of allele frequencies over total sampling area*

Figure 23 (continued)

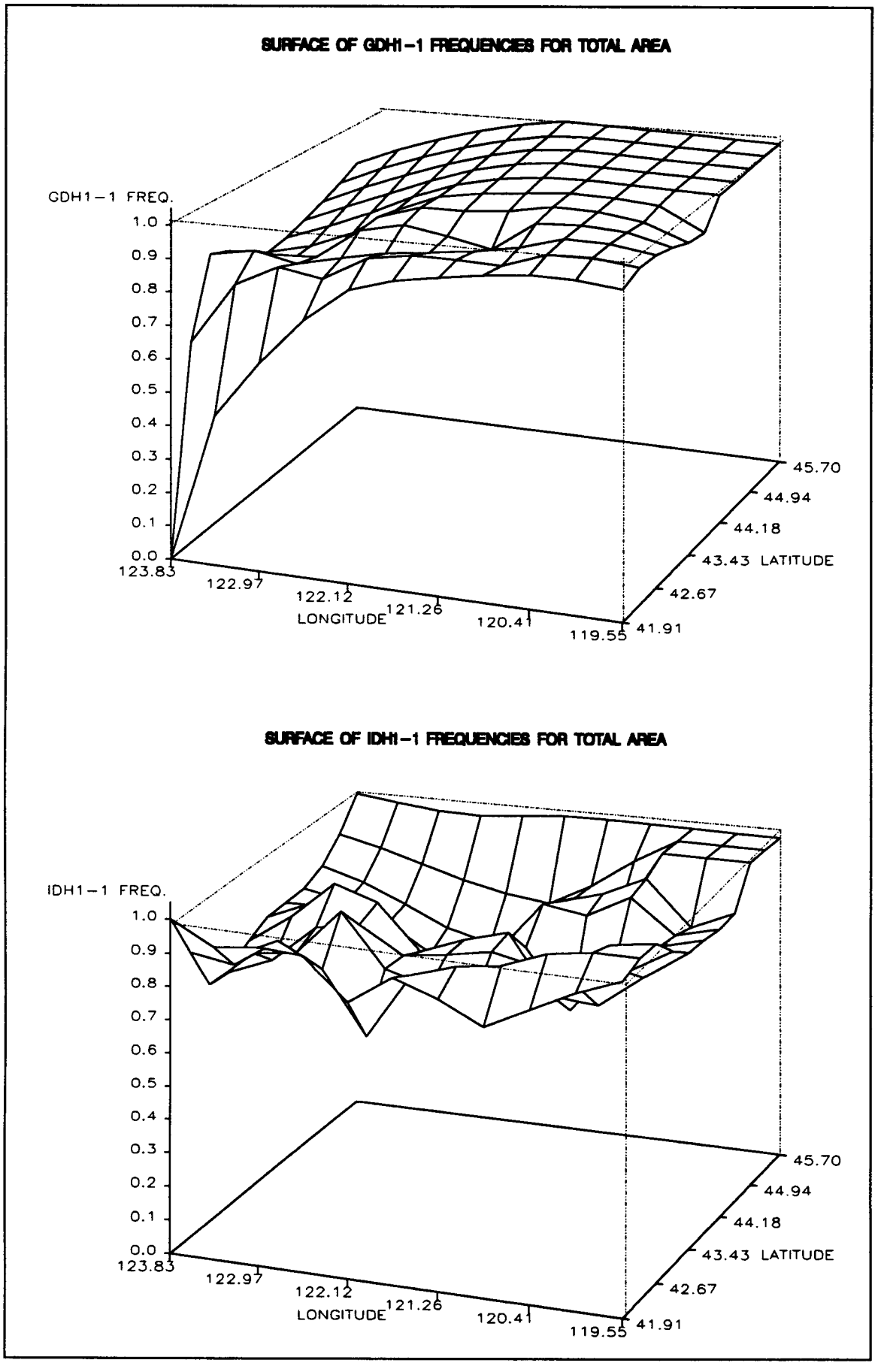
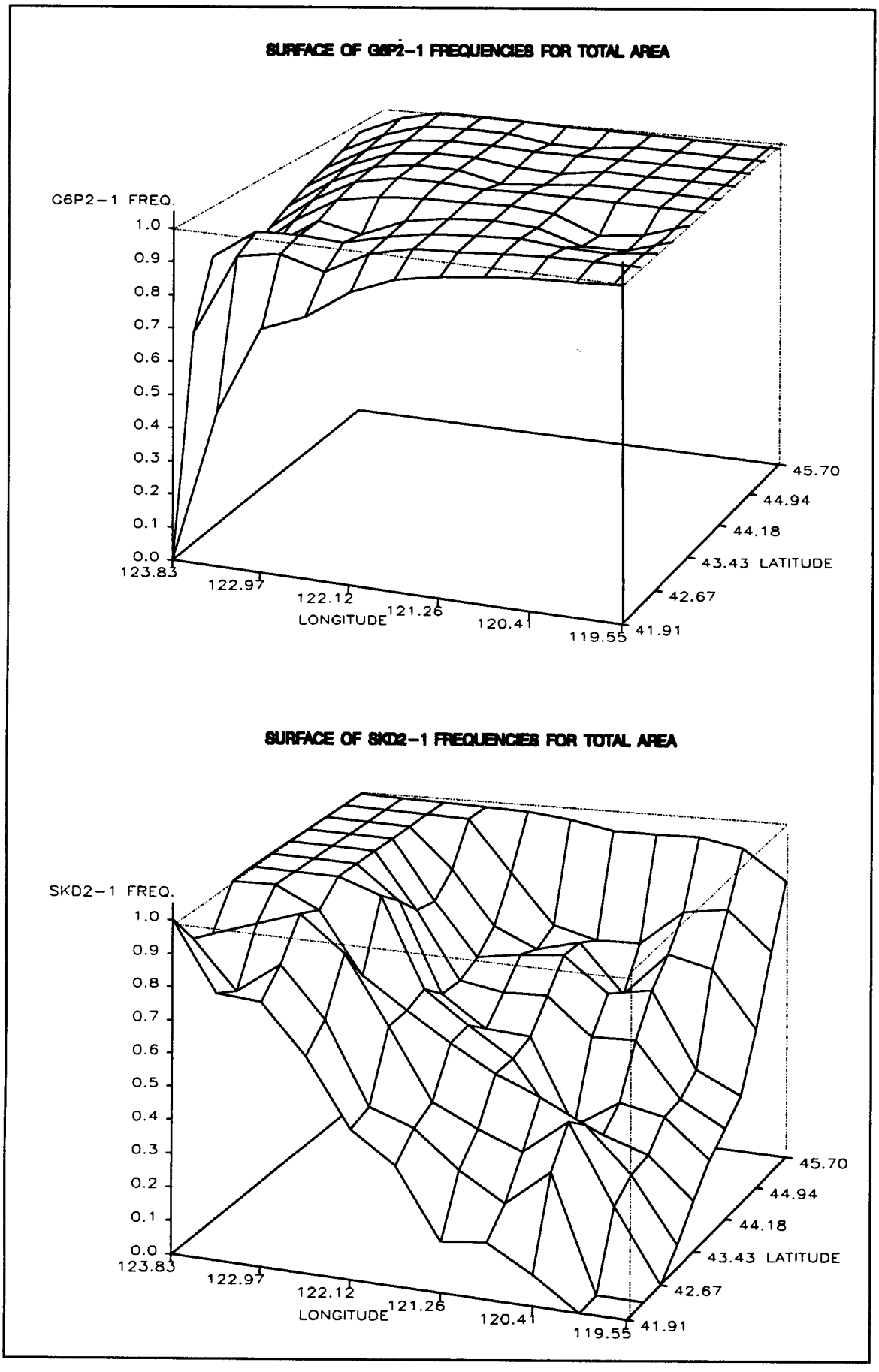


Figure 23 (continued)



surfaces for *Mnr1-1*, *Got3-2*, *Acp1-4*, *G6p2-1*). Seed zone 90, representing this part of the sampling area, was shown in previous analyses to be quite distinct in allozyme frequencies (*section 4.2*).

Patch sizes of significantly autocorrelated allele frequencies confirmed a small-scale patchy structure for most of the alleles (*Figure 24, p. 108 and 109*). A patch size between 25 to 50 km was found for 9 alleles, 6 alleles showed a more or less homogeneous allelic structure within an area of 50 to 100 km in size, and only one allele was characterized by a patch size about equal to the size of the sampling area (100-150 km)(*Got1-3*). Thus, spatial structure of allele frequencies within Southwest Oregon, as revealed by spatial autocorrelation analysis, may be described as a rather small scale allelic variation for 16 alleles while all others exhibit no deviation from spatial random arrangement.

Figure 26, p. 113 shows the correlograms which summarize results of spatial autocorrelation analysis for Central Oregon. Significant spatial structures were scarce. Compared to Southwest Oregon, allele frequency variation in Central Oregon showed much less association with geographic location. Only 6 out of 96 alleles tested showed a spatial pattern which significantly deviated from random arrangement. Spatial patterns were found for one to two alleles at locus *Mnr-2*, *Got-3*, *Aco-1*, *Ugp-2* and *Ugp-3*. Of these 6 alleles, 4 exhibited a depression type of structure while 2 were characterized as a crazy quilt pattern. With one exception, patch size was in the range of 50 to 100 km, suggesting that allelic variation of these alleles is organized in larger homogeneous units than within Southwest Oregon. In contrast to Southwest Oregon, no drastic changes in autocorrelation was observed between the fourth and fifth distance classes. For most alleles, autocorrelation coefficients did not deviate from random expectation in the last distance class.

Spatial structure of allele frequencies within Central Oregon, as revealed by spatial autocorrelation analysis, may be described as a random spatial arrangement for most of the alleles. Although allele frequency differentiation within Central Oregon did not seem to differ much from Southwest Oregon (*see section 4.2*), spatial structures are much scarcer within the Central Oregon sample. As has been stated before, spatial patterns and heterogeneity of variables are potentially independent of each other. Results, therefore, are not contradictory as will be discussed in one of the following sections.

5.2 Associations between genetic, geographic and environmental variation, inferred by Mantel test statistics

Associations between genetic, geographic and environmental patterns were assessed using a statistical technique developed by MANTEL (1967). The value of this technique has been emphasized by ROYALTEY et al. (1975) and SOKAL (1979a) who have adapted the Mantel test for use in analyzing variation in allele frequencies. The Mantel statistic tests the independence of two matrices, one containing values of any specified measure of differences in allele frequencies, and a second matrix containing values of any specified measure of spatial distance or relationship. Hence, the Mantel statistic may be applied to test for randomness of geographic patterns. It may be used to examine whether differences in allele frequencies between pairs of sample localities are statistically associated in a linear manner with the geographic distances between the localities or with differences in environmental conditions that exist between these localities. Two matrices, one of genetic distance and the other of either geographic or environmental distance, are constructed and the sum of the products of their corresponding elements compared with that expected based on a null hypothesis of random permutation. The observed association (Mantel's t - value) between the two distance matrices is compared with a randomized distribution of t (calculated by randomly permutating one of the matrices and calculating t) and the probability of obtaining a value greater than t is calculated. For each Mantel test, 1000 random permutations were carried out to generate the random distribution of t. According to CHEVERUD et al. (1989), this is sufficient to obtain a good empirical distribution of the matrix correlation coefficients. The Mantel test statistic and the distance measures used for the comparisons were described in detail in *section 3.6.1.1*.

Allele frequency differences between localities were tested for associations with geographic dis-

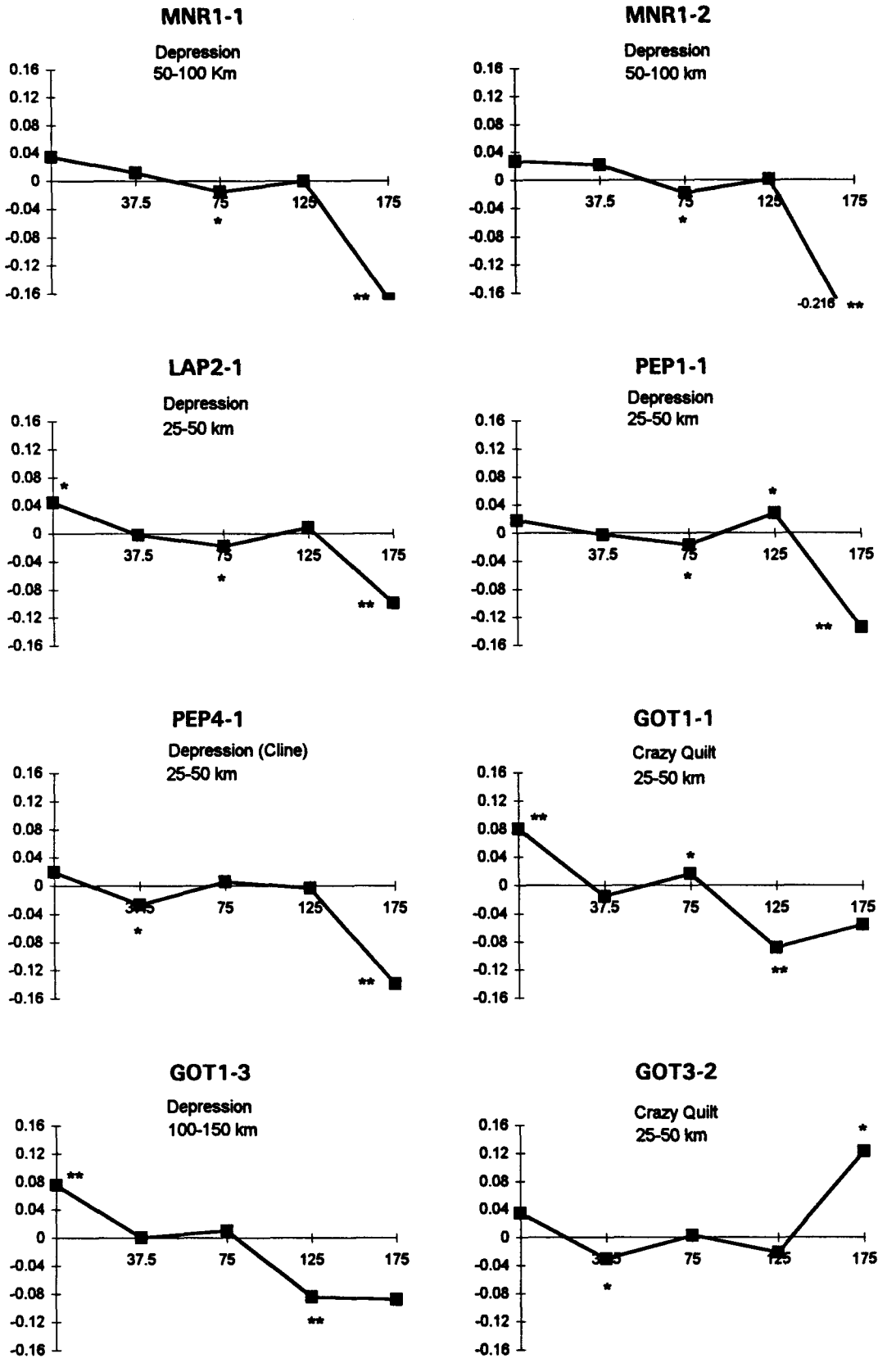
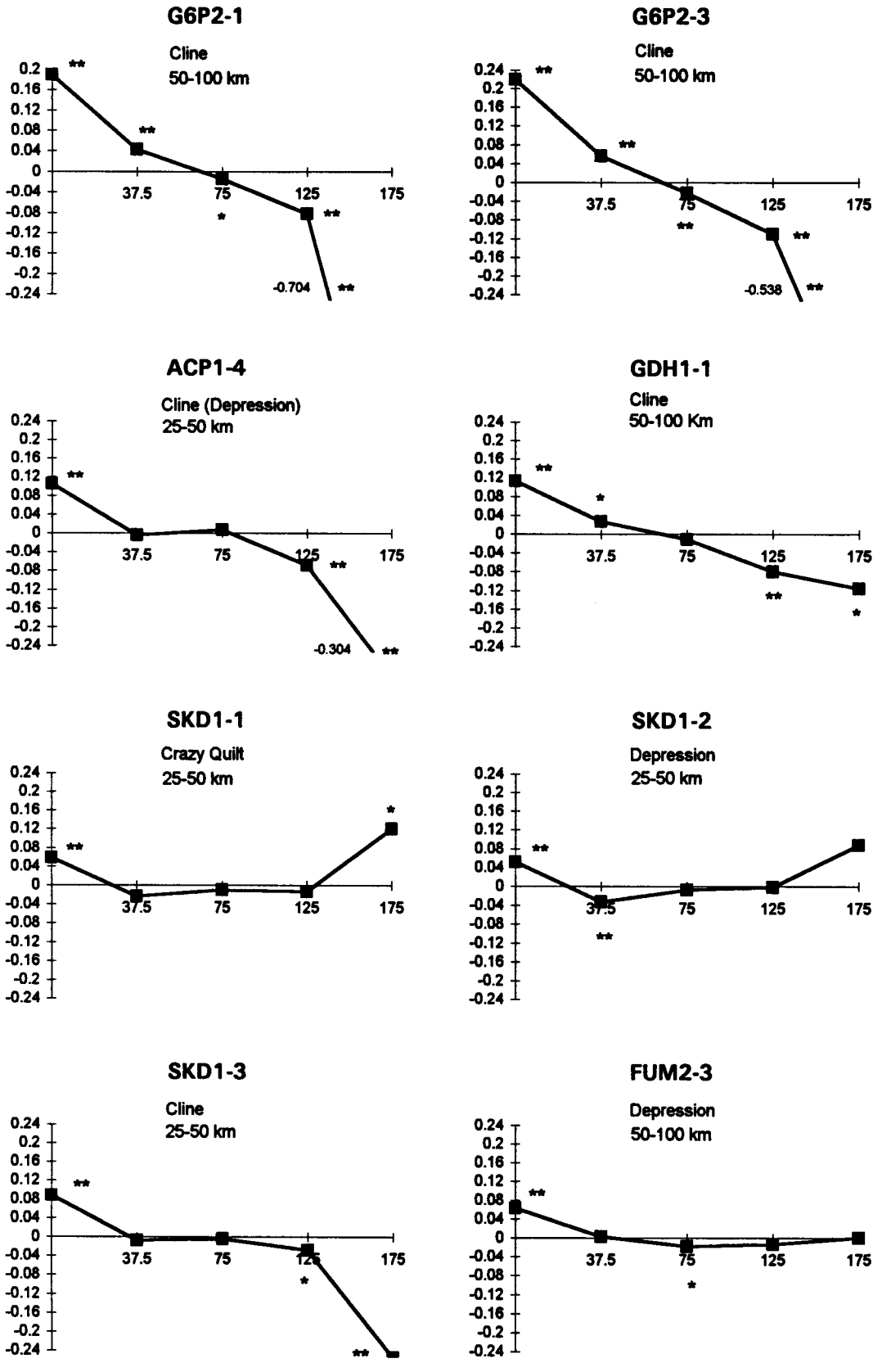


Figure 24: Correlograms of alleles with significant spatial structures: Southwest Oregon. X - axis: Mean values of the 5 distance classes in km; Y - axis: Moran's I coefficients (significance: p=0.01 (**) and p=0.05 (*)). Patterns and patch sizes are indicated

Figure 24 (continued)



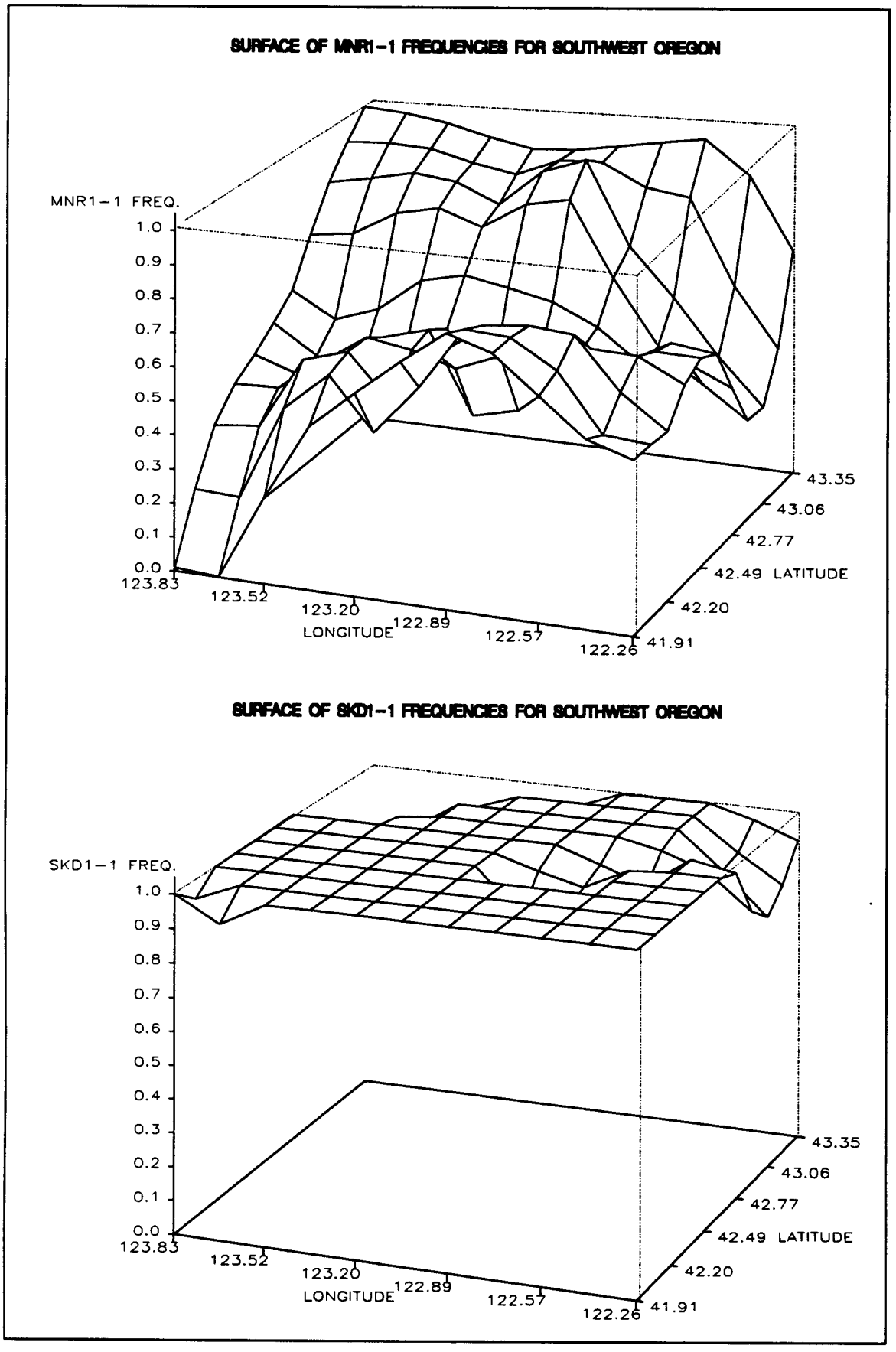


Figure 25: *Examples of variation patterns, illustrated as smoothed surfaces of allele frequencies for Southwest Oregon*

Figure 25 (continued)

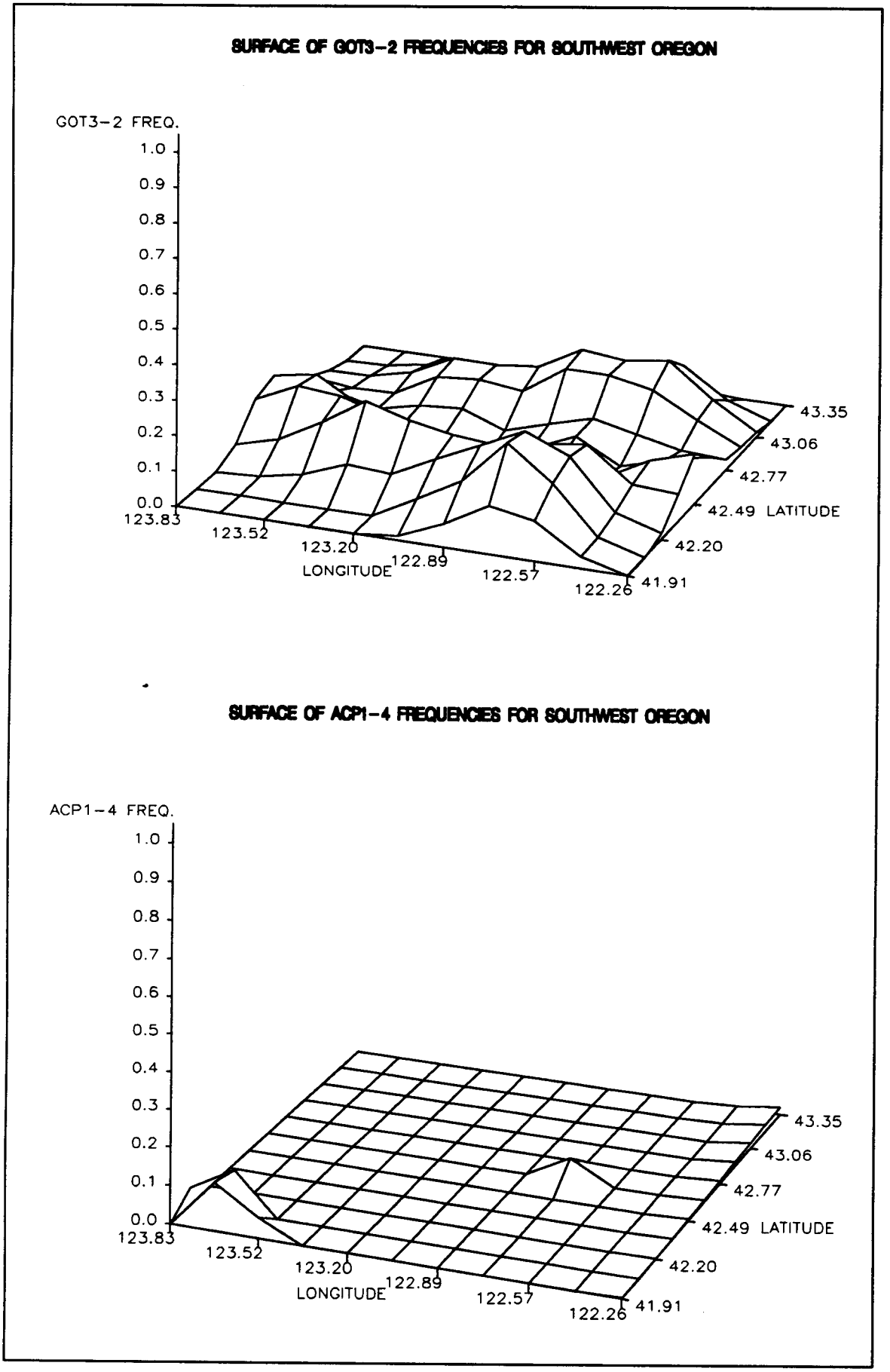
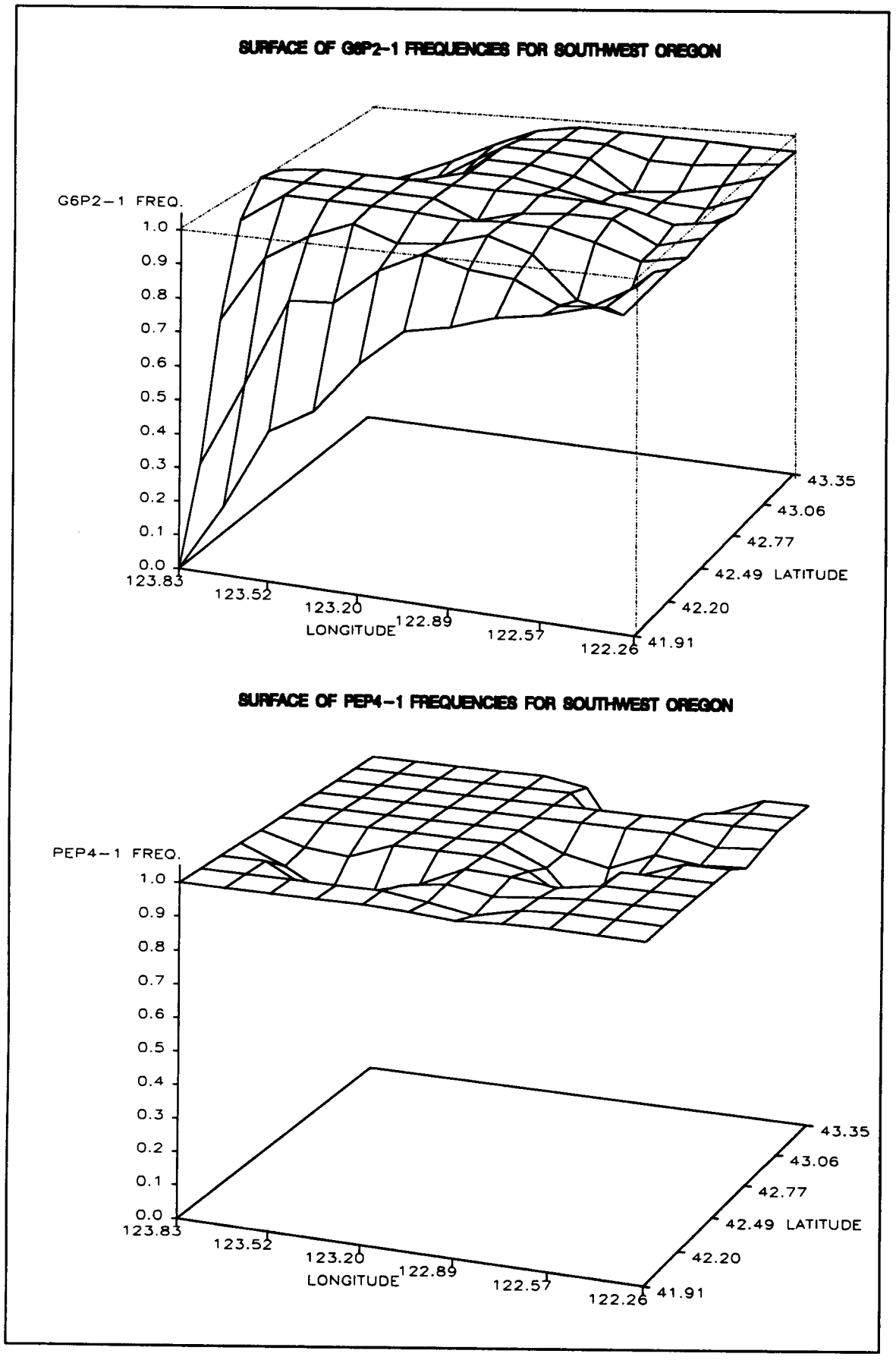


Figure 25 (continued)



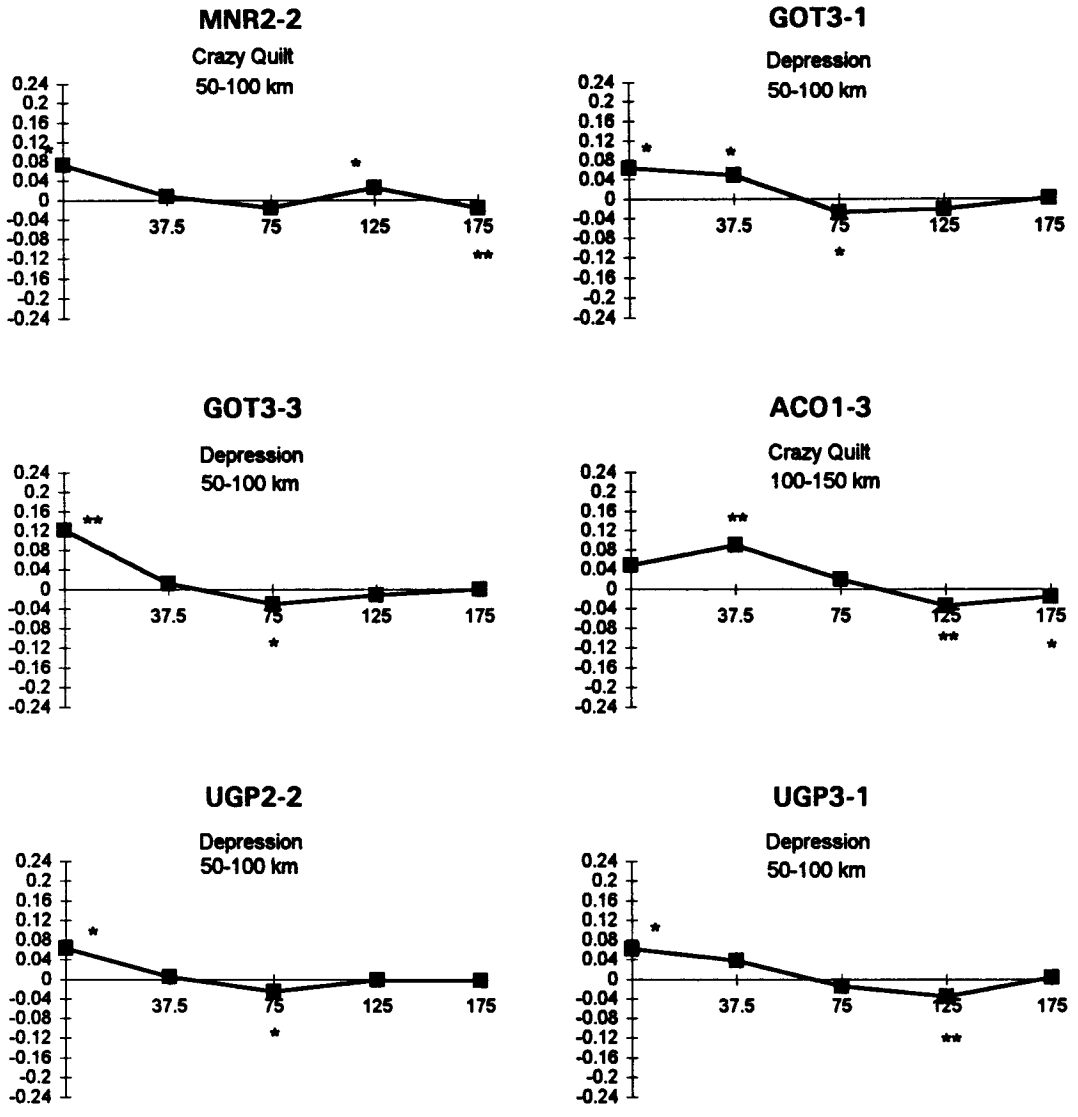


Figure 26: Correlograms of alleles with significant spatial structures: Central Oregon. X - axis: Mean values of the 5 distance classes in km; Y - axis: Moran's I coefficients (significance: p=0.01 (**) and p=0.05 (*)). Patterns and patch sizes are indicated

stance, with the reciprocal of geographic distance and with climatic distance. Furthermore, associations of genetic distances with the two regions were assessed using a dummy coding for the regions. In effect, tests involving reciprocals of geographic distance consider all longer distances to be about equal, while emphasizing differences between shorter distances. This test increases the statistical power to reveal local geographic patterning, whereas tests involving actual distances focus more on regional geographic patterns.

The Mantel-test results, testing associations among genetic, geographic and climatic variables over the total sampling area, are given in *Table 20, p. 114.* The *t*- values of the Mantel test statistics, as well as the significance levels of the observed *t* - values are given in the upper triangular matrix of the table while the lower matrix contains the matrix correlations *r* (Pearson's coefficients of correlation) between the elements of the respective matrices. Since multiple comparisons were made, a conservative level of significance should be selected ($p < 0.01$) in order to avoid the possibility of obtaining a few significant tests just as a consequence of the high number of tests carried out.

Distance matrices	Geographic distance [km]	Climatic distances:			Genetic distances:				Distance matrices
Abreviation	GEO	PCA-Factor 1 CLIM-F1	PCA-Factor 2 CLIM-F2	PCA-Factor 1+2 CLIM-F1/2	GREGORIUS 31 loci GREGORIUS-31	GREGORIUS 15 loci GREGORIUS-15	NEI 31 loci NEI-31	NEI 15 loci NEI-15	Abreviation
GEO		25.679 (<0.001)	14.459 (<0.001)	26.552 (<0.001)	1.006 (0.1860)	1.848 (0.037)	1.371 (0.100)	1.564 (0.068)	GEO
CLIM-F1	0.356		3.776 (<0.001)	-	4.398 (<0.001)	6.782 (<0.001)	4.895 (<0.001)	6.845 (<0.001)	CLIM-F1
CLIM-F2	0.201	0.039		-	-0.308 (0.636)	0.661 (0.279)	-0.049 (0.509)	0.531 (0.299)	CLIM-F2
CLIM-F1/2	0.386	-	-		2.679 (0.004)	4.930 (<0.001)	3.205 (<0.001)	4.867 (<0.001)	CLIM-F1/2
GREGORIUS-31	0.021	0.072	-0.005	0.046		32.878 (<0.001)	34.570 (<0.001)	29.200 (<0.001)	GREGORIUS-31
GREGORIUS-15	0.037	0.099	0.010	0.076	0.764		31.074 (<0.001)	42.701 (<0.001)	GREGORIUS-15
NEI-31	0.029	0.082	-0.001	0.056	0.943	0.759		32.433 (<0.001)	NEI-31
NEI-15	0.032	0.104	0.008	0.078	0.714	0.930	0.812		NEI-15

Distance matrices	Geographic:		Genetic:			
	reciprocal	regional	GREGORIUS 31 loci	GREGORIUS 15 loc	NEI 31 loci	NEI 15 loci
Abreviation	1/GEO+1	REG	GREGORIUS-31	GREGORIUS-15	NEI -31	NEI-15
1/GEO+1		-	-3.825 (<0.001)	-4.078 (<0.001)	-4.196 (<0.001)	-4.221 (<0.001)
REG	-		10.162 (<0.001)	13.053 (<0.001)	10.098 (<0.001)	12.462 (<0.001)
GREGORIUS-31	-0.018	0.052		32.878 (<0.001)	34.570 (<0.001)	29.200 (<0.001)
GREGORIUS-15	-0.021	0.062	0.764		31.074 (<0.001)	42.701 (<0.001)
NEI -31	-0.020	0.052	0.943	0.759		32.433 (<0.001)
NEI-15	-0.022	0.061	0.714	0.930	0.812	

UPPER TRIANGULAR MATRIX:

- Values of Mantel test statistics: t= [Z - E(Z)] / var(Z)]
- in brakets are the significance levels of the t - values. Probabilities were derived by comparing the t-values to a randomized distribution of t, generated by 1000 random permutations of one of the matrices for each comparison

LOWER TRIANGULAR MATRIX:

- Values of matrix correlations r (Pearson's coefficient of correlation between the elements of the two matrices)

Geographic distances:

- Distances in km between sample trees
- Reciprocal of distances [km] between sample trees (1/distance+1), emphasizing differences between shorter distances (local patterning)
- Regional: Dummy coding: 0 for same region, 1 for different region

Climatic distances:

- Distances calculated based on rank differences for PCA Factor 1, Factor 2 and combined according to MANTEL and VALAND, 1970

Genetic distances:

- Distances between individual trees according to:
  - GREGORIUS (GREGORIUS, 1974) (=Prevosti)
  - NEI (NEI, 1972)
- 31 loci: All loci used for calculation of distance matrix
- 15 loci: Only loci with significant spatial autocorrelation (Moran's I) used for calculation of distance matrix [MNR1, MNR2, LAP2, PEP1, PEP3, GOT1, G6P1, ACP1, GDH1, IDH1, PGM1, ACO1, SKD2, FDP2, ADH1]

***Table 20**: Mantel tests of matrix comparisons between genetic, geographic and climatic distances among individual trees (N=488 resp. 411): Total sampling area*

Southwest Oregon						
	Genetic distances:					
	GREGORIUS 31 loci			GREGORIUS 11 loci		
Association with:	Mantel t	Prob p	Pearson r	Mantel t	Prob p	Pearson r
Geographic distance	-0.363	0.369	-0.008	0.314	0.382	0.008
Reciprocal geographic distance	-2.265	0.014	-0.013	-0.419	0.330	-0.002
Climatic distance Factor 1	-0.685	0.764	-0.015	-0.882	0.172	-0.020
Climatic distance Factor 2	-0.341	0.618	-0.007	-1.036	0.147	-0.024
Climatic distance Factor 1/2	-0.750	0.781	-0.016	-1.401	0.081	-0.031
Central Oregon						
	Genetic distances:					
	GREGORIUS 31 loci			GREGORIUS 5 loci		
Association with:	Mantel t	Prob p	Pearson r	Mantel t	Prob p	Pearson r
Geographic distance	0.938	0.185	0.028	0.155	0.450	0.004
Reciprocal geographic distance	-1.288	0.107	-0.012	-0.768	0.219	-0.007
Climatic distance Factor 1	0.962	0.177	0.024	2.216	0.013	0.050
Climatic distance Factor 2	0.036	0.378	0.001	1.272	0.112	0.029
Climatic distance Factor 1/2	0.667	0.273	0.016	2.391	0.018	0.053

Table 21: *Mantel tests of matrix comparisons between genetic distance by GREGORIUS (1974) and geographic and climatic distances between individual trees for the two regions*

Climatic distances (*Factor 1* (temperature), *Factor 2* (water balance) and both factors combined) were significantly associated ($p < 0.001$) with geographic distances between localities. Associations were rather weak, however, suggesting that climate conditions are only weakly patterned in distinct geographic patches of similar climate. Similar climate conditions seem to occur at rather irregular spacings over geographic space. Temperature (*Factor 1*) was more associated with geography than water balance (*Factor 2*). Climate distances involving both components combined showed a correlation of $r = 0.386$ with distance between location, which means that only 15% of the distance in climate conditions between samples was associated with the geographic distance between the samples.

Both measures of genetic distance by GREGORIUS (1974) and by NEI (1972) were tested for associations with geographic distance. Calculation of these distances were either based on all 31 loci or on the 15 loci only which displayed a spatial structure. Correlations between these two genetic distances, one based on all 31 loci, the other based on 15 loci only, were high ($r = 0.764$ or $r = 0.812$) for both distance measures. The 15 loci, constituting 48% of all investigated loci, represented 58% (GREGORIUS distance) and 66% (NEI's distance) of total variance in all 31 loci. Hence, each of these 15 loci, selected by spatial autocorrelation analysis, obviously assembled more variation than the average locus. Correlations between the two distance measures were very high ($r = 0.943$ (31 loci) and $r = 0.930$ (15 loci)).

All tested associations between genetic and geographic distance were not significant ($p > 0.01$). Hence, the distribution of allele frequencies, expressed as a composite index over 31 loci, could not be distinguished from a random pattern. The same was true for the 15 loci which revealed a spatial structure in spatial autocorrelation analysis. Correlation coefficients were slightly higher when only the 15 loci were analyzed, but associations with geographic distance remained very low and not significant. Although comparisons involving the reciprocals of distance resulted in significant ($p < 0.001$) t - values for all genetic distance measures, estimated correlations were very low. Nonetheless, the significant t -values, being about 3 to 4 times higher than t - values resulting from the tests using untransformed distance, suggest that genetic variation is characterized rather by local patterning than by large scale patterns. Genetic distance was only weakly associated with the two regions for all distance measures. Although t - values were significant ($p < 0.001$), indicating a regional effect, correlation coefficients were all below $r = 0.062$.

Genetic distance measures were significantly ($p < 0.001$) associated with climate *Factor 1* (temperature), but showed no associations with climate *Factor 2* (water balance). Although correlation coefficients were all below $r = 0.10$, indicating only weak relationships between the composite measures of genetic variation and climate conditions, the test statistics clearly suggest that genetic distance is more closely associated with distance in climate *Factor 1* than with *Factor 2*. The autocorrelated 15 loci showed a stronger association with climate than did the composite measures based on all 31 loci, indicating that the observed patterns of these loci may be related to variation in temperature regime. Associations remained low, however, with a shared variance between genetic and climatic distance of 1% at best.

Results of Mantel test statistics for the two regions separately, utilizing just the GREGORIUS distance measure for genetic distance, are presented in *Table 21, p. 115*. All comparisons resulted in insignificant associations and very low coefficients of correlation, indicating that the distribution of genetic variation within both regions cannot be distinguished from a random pattern. Within Southwest Oregon, genetic distance based on 31 loci was nearly significantly ($p = 0.014$) associated with the reciprocal of geographic distance, suggesting that patterning is rather local. Within Central Oregon, a nearly significant association ($p = 0.013$) between climate *Factor 1* and genetic distance was observed. But in neither case was Pearson's r very large.

5.3 Associations between single locus genotypes and climate, inferred by multinomial response models

Multinomial response models were used to study associations between genotypes and climate variables. Allelic genotypes of the 411 individuals, where climate data was available for the source locations, were used as response variables and were tested in different models. The response variables were either 1 for homozygous individuals with two copies of the allele, 0.5 for heterozygotes carrying one copy of the allele, or zero when no copy of the allele in question was present. Climate variables *Factor 1* and *Factor 2* were used as predictor variables. Multinomial response models and the specified models used in the analysis were described in detail in *section 3.6.2*.

Results of multinomial response models for the total sampling area are furnished in *Table 22, p. 118*. Significance levels of the three effects (climate *Factor 1*, *Factor 2*, interaction *F1*F2*) in the full model are given. In addition, for each model with a signficant effect, a reduced model with the respective significance level of the predictor variable is given. Since likelihood ratio tests are not appropriate for continuous predictor variables, the square of the respective predictor variable was added to the reduced model; linearity (and sufficient model fit) was assumed when this second order term was non significant.

Significant associations between genotypes and climate were found for 33 alleles at 18 loci at a significance level below $p < 0.05$ (*Table 22, p. 118*, reduced models). Considering the high number of tests which were performed, a more conservative level of significance should be applied, however, in order to avoid the probability of obtaining significant associations only as a consequence of the high number of tests carried out. Nineteen alleles at 12 loci *i.e.* at *Mnr-1*, *Mnr-2*, *Lap-2*, *Pep-3*, *G6p-2*, *Acp-1*, *Gdh-1*, *Idh-1*, *Pgm-1*, *Aco-1*, *Skd-2* and *Fdp-2* exhibited association with one of the predictor variables at $p < 0.01$. All significant models showed associations with climate *Factor 1* (temperature regime) while associations with climate *Factor 2* (water balance) were scarce (*Mpi1-1*, *Acp1-3*, *Aco1-3* and *Fum2-3*) and significant only at $p < 0.05$ but not at $p < 0.01$. Based on significance of the second order terms in the reduced models, linearity was not violated in any of the significant models.

Some of the significant associations between genotypes and climate may have been caused by indirect effects. Such indirect effects may be interpreted as adaptation to climate conditions although the associations are in fact due to other effects. Since climate conditions differ substantially between the two regions, all other effects which are related to the two regions (for example a different evolutionary history in the two areas) are likely to cause such indirect associations, mimicking climate associations. To preclude false interpretations, nested models with region and climate *Factor 1* nested within region and nested-by-value effect models were tested in a second step. Strong regional effects accompanied by insignificant effects of climate

Factor 1 within region are indications for such indirect effects. Climate effects may, however, not be separated completely from other effects. Strong regional effects may at least partly be caused by climate as well since climate conditions differ substantially between regions. Non-significant effects in the nested models, on the other hand, may partly be caused by the reduction in sample size (Central Oregon=222, Southwest Oregon=189) or by adaptive differences between the two regions, *i.e.*, by a significant association in one region and a non significant association in the other region, which in turn leads to an overall not significant effect. If both within-region effects are not significant, then indirect effects are likely. Indirect effects are also indicated by different signs of the parameters in the probability functions. The signs of the parameters are expected to be the same in both regions if a direct effect of climate is postulated. While significance levels may differ due to different selection pressures under different climatic conditions, directions of genotypic changes should be the same in both regions, if adaptation was in fact responsible for the association.

Results of nested multinomial response models are shown in *Table 23, p. 119 and 120.* Only climate *Factor 1* has been used in the models since climate *Factor 2* has been shown to lack associations with genotypic frequencies in the first step of the analysis. Some genotypes could not be tested due to the imbalanced data structure leading to infinite parameter estimates.

Several of the significant associations revealed in the first analysis (*Table 22, p. 118)* seem to be due to indirect effects. Based on the reasoning mentioned above, such indirect effects may be postulated in the case of *Mnr1-2*, *Acp1-1*, *Acp1-2*, *Adh2-1* and *Adh2-2* which showed different signs of parameters in the two areas. In addition, indirect effects are likely in the case of *Mnr1-1*, *Lap2-1*, *G6p2-3* and *Fdp2-1*, since both within region probabilities were clearly not significant. On the other hand, some of the associations which were not significant at $p < 0.01$ in the first step of the analysis, seem to be associated with climate effects *i.e. Mpi1-1*, *Mpi1-3* and *Ugp1-1* since within region probabilities were significant at $p < 0.05$. Finally, some of the alleles could not be tested due to infinite parameter estimates *i.e. G6p2-1* and *Gdh1-1*.

Hence, based on the combined interpretation of results from reduced and nested models, significant direct associations between genotypic frequencies and temperature may be suggested for 15 alleles *i.e.* for *Mnr2-1*, *Mnr2-2*, *Pep3-1*, *Pep3-2*, *Mpi1-1*, *Mpi1-3*, *Idh1-1*, *Pgm1-1*, *Pgm1-2*, *Pgm1-4*, *Aco1-4*, *Skd2-1*, *Ugp1-1*, *Mdh3-1* and *Mdh3-2* while *Gdh1-1* and *G6p2-1* are likely candidates but indirect effects could not be precluded.

Significant associations between genotypic frequencies and temperature (climate *Factor 1*) were very scarce within Southwest Oregon (*Mnr2-2* at $p < 0.05$ only) while more frequent associations were found within Central Oregon *i.e.* for alleles *Mnr2-1*, *Mnr2-2*, *Pep3-1*, *Mpi1-1*, *Mpi1-3*, *Pgm1-1* and *Pgm1-2* at $p < 0.01$, and *Pep3-2*, *Idh1-1*, *Pgm1-4*, *Skd2-1*, *Ugp1-1* and *Ugp2-4* at $p < 0.05$.

In summary, genotypic frequencies of several alleles showed associations with temperature conditions at source location. In contrast, genetic structure was not related to water balance. Significant associations were scarce within Southwest Oregon while frequent associations were manifest within Central Oregon.

5.4 Discussion

Results of spatial autocorrelation analyses produced clear evidence that allele frequencies do not vary randomly over geographic space. Similar genotypes for several alleles clustered in geographic space. Such patterns of variation in allele frequencies may have been induced by different evolutionary forces. SOKAL (1978) has considered four models for the origin of patterns of differentiation. In the first model, a character is differentiated in response to an environmental gradient, producing a cline. In the second, environmental patches are heterogeneous among themselves but homogeneous within with respect to certain ecological variables inducing selection. These patches may or may not exhibit spatial ordering. A third model is differentiation due to the classical isolation by distance model. A fourth model ascribes the differentiation to historical events (founder effect resulting in patchy distributions; immigration into the

LOCUS Allele	Models with climate Factor 1/Factor2 Significance levels of associations in model:						LOCUS Allele	Models with climate Factor 1/Factor2 Significance levels of associations in model:					
	Full model with:			Reduced model with:				Full model with:			Reduced model with:		
	F 1	F2	F1*F2	F1	F2	F*F		F 1	F2	F1*F2	F1	F2	F*F
	Prob	Prob	Prob	Prob	Prob	Prob		Prob	Prob	Prob	Prob	Prob	Prob
	p	p	p	p	p	p		p	p	p	p	p	p
MNR1-1	0.0025	0.9446	0.7085	0.0027	-	0.9659	ACP1-3	0.3711	0.0412	0.0779		0.0411	0.4599
MNR1-2	0.0085	0.5432	0.7931	0.0056	-	0.5544	ACP1-4	0.2873	0.2249	0.5614			
MNR1-3	0.4083	0.1403	0.9656				ACP1-5	0.5818	0.9795	0.9636			
MNR1-4	0.8677	0.3398	0.7443				GDH1-1	0.0001	0.9292	0.6991	0.0001	-	0.4816
MNR2-1	0.0001	0.1631	0.0801	0.0003	-	0.2294	GDH1-2	0.0004	0.9387	0.7827	0.0003	-	0.3895
MNR2-2	0.0000	0.7499	0.4167	0.0000	-	0.9272	GDH1-3	0.4844	0.5200	0.8232			
MNR2-3	0.3283	0.1556	0.5326				IDH1-1	0.0000	0.0748	0.7711	0.0000	-	0.5529
LAP2-1	0.0010	0.3087	0.5214	0.0013	-	-	PGM1-1	0.0009	0.5875	0.9209	0.0017	-	0.7497
LAP2-2	0.0173	0.1678	0.1340	0.0103	-	0.8318	PGM1-2	0.0014	0.6789	0.6614	0.0010	-	0.3120
LAP2-3	0.4367	0.2251	0.6445				PGM1-3	0.4978	0.5762	0.7505			
LAP2-4	0.0198	0.9759	0.4536	0.0195	-	-	PGM1-4	0.0289	0.8744	0.8744	0.0094	-	0.8695
LAP3-1	0.2840	0.7939	0.4794				ACO1-1	0.2913	0.5298	0.3786			
LAP3-2	0.4382	0.9672	0.5117				ACO1-2	0.3242	0.1676	0.8631			
LAP3-3	0.2100	0.3069	0.2069				ACO1-3	0.0859	0.0509	0.0308	-	0.0186	-
PEP1-1	0.0895	0.1302	0.5310	0.0261	-	0.1607	ACO1-4	0.0007	0.6636	0.6495	0.0004	-	0.3106
PEP1-2	0.2194	0.3505	0.5910				SKD1-1	0.4426	0.5116	0.9355			
PEP1-3	0.2113	0.1553	0.9963				SKD1-2	0.4595	0.5638	0.9645			
PEP2-1	0.1847	0.3912	0.4253				SKD1-3	0.9875	0.4938	0.3600			
PEP2-2	0.0965	0.1808	0.3811				SKD2-1	0.0000	0.0808	0.7026	0.0000	-	0.0994
PEP2-3	0.6959	0.9539	0.5379				FDP2-1	0.0007	0.6769	0.6075	0.0006	-	0.0637
PEP3-1	0.0008	0.0910	0.6336	0.0002	-	0.6915	UGP1-1	0.0596	0.3545	0.0257	0.0265	-	0.5387
PEP3-2	0.0015	0.8793	0.7398	0.0010	-	0.4888	UGP1-2	0.1167	0.6803	0.4371	0.0294	-	0.4293
PEP3-3	0.1494	0.1299	0.7939				UGP1-3	0.9626	0.6187	0.2983			
PEP4-1	0.2753	0.3763	0.6239				UGP2-1	0.1199	0.1764	0.8730			
MPI1-1	0.5058	0.0608	0.1686	-	0.0500	0.6265	UGP2-2	0.3016	0.7724	0.1962			
MPI1-2	0.3077	0.4839	0.1514				UGP2-3	0.8139	0.6615	0.1130			
MPI1-3	0.4032	0.1375	0.4945				UGP2-4	0.0986	0.1316	0.0867			
MPI2-1	0.2931	0.3487	0.9288				UGP3-1	0.7406	0.5873	0.4985			
MPI2-2	0.3027	0.3791	0.5688				FUM2-1	0.4984	0.9378	0.1282			
MPI2-3	0.6441	0.3634	0.7636				FUM2-2	0.7940	0.9662	0.0460			
GOT1-1	0.7152	0.6782	0.4206				FUM2-3	0.0920	0.0366	0.2597	-	0.0486	-
GOT1-2	0.4021	0.1270	0.7685				ADH2-1	0.0058	0.8932	0.0835	0.0171	-	0.3935
GOT1-3	0.7500	0.9588	0.4635				ADH2-2	0.0048	0.9303	0.0729	0.0149	-	0.4334
GOT1-4	0.7395	0.5645	0.2828				ADH2-3	0.9701	0.2812	0.6441			
GOT2-1	0.7403	0.4408	0.5205				PGI2-1	0.1177	0.3386	0.9704			
GOT2-2	0.7034	0.4640	0.5323				PGI2-2	0.9173	0.9835	0.5866			
GOT2-3	0.8047	0.7864	0.9338				PGI2-3	0.0842	0.4197	0.1140			
GOT3-1	0.3997	0.5470	0.8226				PGI2-4	0.3270	0.1687	0.5460			
GOT3-2	0.0734	0.9732	0.3120				MDH1-1	0.1482	0.5205	0.4994			
GOT3-3	0.8946	0.2863	0.5753				MDH1-2	0.3011	0.9318	0.6980			
GOT3-4	0.2029	0.2537	0.0216				MDH1-3	0.8512	0.1964	0.7188			
GOT3-5	0.3302	0.5844	0.8683				MDH3-1	0.0208	0.0579	0.9276	0.0148		0.5339
G6P2-1	0.0182	0.3489	0.4038	0.0060	-	0.1694	MDH3-2	0.0299	0.6046	0.7232	0.0115		0.8683
G6P2-2	0.0798	0.8943	0.7853				MDH3-3	0.3808	0.8383	0.8776			
G6P2-3	0.0372	-	0.1774	0.0020	-	-	MDH3-4	0.8549	0.3705	0.9282			
ACP1-1	0.0140	0.7555	0.5718	0.0161	-	0.4633	MDH4-1	0.1961	0.1697	0.8252			
ACP1-2	0.0009	0.9785	0.3155	0.0017	-	0.9028							

Table 22: Associations of genotypes with climate conditions. Multinomial response models with climate Factor 1 and Factor 2 as predictor variables: Total sampling area

LOCUS Allele	Models with Factor 1 within regions (R=S: Southwest Oregon, R=C: Central Oregon)									
	Significance levels of associations in model:				Estimated parameters					
	F 1 (Region	Region	F 1 (R=S)	F 1 (R=C)	F 1 (R=S)		F 1 (R=C)		Region	
	Prob	Prob	Prob	Prob	2	1	2	1	2	1
	p	p	p	p	copies	copy	copies	copy	copies	copy
MNR1-1	0.6674	0.2692	0.8575	0.3559	0.022	-0.130	0.279	0.338	0.400	0.235
MNR1-2	0.4021	0.2174	0.7542	0.1768	0.276	0.154	-0.537	-0.441	-0.420	-0.102
MNR1-3	0.9394	0.5591	0.8340	0.7756	0.210	-	-0.181	-	-0.380	-
MNR1-4	0.1921	0.9373	0.1805	0.2628	-7.101	-0.410	2.208	0.145	0.215	0.088
MNR2-1	0.0148	0.4425	0.2318	0.0089	-0.540	-0.423	-0.978	-0.233	0.223	0.228
MNR2-2	0.0003	0.1069	0.0163	0.0016	0.919	0.716	0.625	0.920	0.336	0.381
MNR2-3	0.6079	0.7771	0.4636	0.5570	-0.566	-0.205	-0.050	0.238	-0.191	-0.089
LAP2-1	0.7802	-	0.4256	0.9759	-4.579	-4.263	-1.301	-1.184	-2.985	-2.438
LAP2-2	0.3312	0.7446	0.1597	0.6297	0.994	-	0.486	-	0.201	-
LAP2-3	0.6388	0.1283	0.3471	0.9134	-0.497	-	-0.073	-	0.624	-
LAP2-4	-	-	-	-						
LAP3-1	0.1573	0.3931	0.1983	0.1841	-1.331	-1.066	1.857	1.718	-0.819	-1.013
LAP3-2	0.1619	0.8114	0.3162	0.1198	3.211	-0.027	-3.575	-0.235	1.007	0.109
LAP3-3	-	-	-	-						
PEP1-1	0.5283	0.0934	0.4575	0.3947	0.255	-	-0.323	-	0.408	-
PEP1-2	0.5633	0.4545	0.7078	0.3155	-0.147	-	0.387	-	0.197	-
PEP1-3	0.6611	0.1138	0.4538	0.6055	-0.460	-	-0.921	-	1.735	-
PEP2-1	-	-	-	-						
PEP2-2	-	-	-	-						
PEP2-3	0.3077	0.1253	0.9066	0.1258	0.064	-	-1.073	-	0.781	-
PEP3-1	0.0123	0.3443	0.2231	0.0008	2.567	-3.014	0.320	-0.489	2.135	2.350
PEP3-2	0.0633	0.6521	0.7039	0.0165	2.675	0.053	1.167	-0.831	-0.148	-0.148
PEP3-3	-	-	-	-						
PEP4-1	-	-	-	-						
MPI1-1	0.0086	0.3438	0.8267	0.0013	-1.112	-1.015	-2.149	-1.994	-1.703	-1.112
MPI1-2	0.9360	0.3009	0.8467	0.8185	0.124	-	0.091	-	-0.402	-
MPI1-3	0.0115	0.5204	0.8468	0.0018	1.103	0.078	2.457	0.130	1.506	0.066
MPI2-1	-	-	-	-						
MPI2-2	0.4597	0.5988	0.7560	0.2273	-0.127	-	-0.427	-	-0.145	-
MPI2-3	-	-	-	-						
GOT1-1	-	-	-	-						
GOT1-2	0.7605	0.8386	0.5370	0.6834	-0.528	-	-0.263	-	-0.104	-
GOT1-3	-	-	-	-						
GOT1-4	0.6108	0.5898	0.9126	0.3237	0.138	-	0.768	-	-0.404	-
GOT2-1	0.2895	0.1583	0.1245	0.7302	0.669	-	0.143	-	-0.400	-
GOT2-2	0.4019	0.2449	0.1918	0.7302	-0.582	-	-0.143	-	0.336	-
GOT2-3	-	-	-	-						
GOT3-1	0.6634	0.5849	0.4612	0.6545	1.727	1.990	0.157	-0.146	-0.311	-0.074
GOT3-2	0.8765	0.1520	0.8658	0.6277	0.064	-	-0.216	-	0.427	-
GOT3-3	0.6660	0.6200	0.3231	0.9427	-1.753	0.924	-0.593	-0.041	0.824	-0.361
GOT3-4	0.8909	0.9608	0.8542	0.6569	0.165	-	-0.342	-	-0.031	-
GOT3-5	0.8100	0.6763	0.7950	0.5520	0.338	-	-0.466	-	0.348	-
G6P2-1	-	-	-	-						
G6P2-2	0.8058	0.1545	0.6815	0.6077	-0.228	-	0.430	-	-0.658	-
G6P2-3	0.5210	-	0.2720	0.7330	5.384	0.847	1.310	0.740	-	-
ACP1-1	0.3904	0.1640	0.5253	0.2430	-0.581	-0.480	1.416	1.540	-1.253	-1.173
ACP1-2	0.8785	0.1094	0.6536	0.8408	0.491	-0.046	-0.901	0.074	0.301	0.491
ACP1-3	0.8088	0.8248	0.9222	0.4873	0.758	-0.151	-1.505	0.023	0.095	-0.424
ACP1-4	0.3030	0.0903	0.1353	0.6912	3.060	-	0.181	-	-2.544	-
ACP1-5	-	-	-	-						
GDH1-1	-	-	-	-						
GDH1-2	0.2926	-	0.0995	0.8466	0.500	1.030	1.342	-0.500	3.811	0.642
GDH1-3	-	-	-	-						

Table 23: Associations of genotypes with climate conditions. Nested multinomial response models and nested by-value effect models with regions and climate Factor 1 as predictor variables. Total sampling area

Table 23 (continued)

LOCUS	Models with Factor 1 within regions (R=S: Southwest Oregon, R=C: Central Oregon)									
Allele	Significance levels of associations in model:				Estimated parameters					
	F 1 (Region	Region	F 1 (R=S)	F 1 (R=C)	F 1 (R=S)		F 1 (R=C)		Region	
	Prob	Prob	Prob	Prob	2	1	2	1	2	1
	p	p	p	p	copies	copy	copies	copy	copies	copy
IDH1-1	0.0960	0.1598	0.7706	0.0250	0.395	0.251	1.088	0.799	0.377	0.057
PGM1-1	0.0377	0.7916	0.6656	0.0093	1.033	0.924	1.750	1.360	-0.327	-0.401
PGM1-2	0.0170	0.4938	0.4081	0.0063	-0.635	-	-1.359	-	0.309	-
PGM1-3	-	-	-	-						
PGM1-4	0.1496	0.5520	0.7090	0.0482	-1.010	-0.011	-1.689	-0.100	-0.508	0.159
ACO1-1	0.7526	0.3283	0.6486	0.5938	0.126	0.249	0.102	-0.183	0.414	0.126
ACO1-2	0.1525	0.1394	0.0617	0.5684	-1.056	-0.142	-0.314	-0.207	-0.557	-0.041
ACO1-3	-	-	-	-						
ACO1-4	0.1107	0.9566	0.0620	0.3751	0.254	0.668	0.333	0.197	0.011	0.053
SKD1-1	0.8095	0.3637	0.4679	0.9623	-2.916	-2.995	-0.116	-0.161	-2.380	-2.280
SKD1-2	0.7631	0.3870	0.4823	0.8216	2.920	-0.010	0.092	-0.157	2.378	0.078
SKD1-3	0.4178	0.8056	0.4817	0.2635	-0.693	-	0.692	-	0.124	-
SKD2-1	0.1083	0.1565	0.8663	0.0261	0.226	0.216	0.909	0.482	0.134	-0.219
FDP2-1	0.9889	0.0323	0.8903	0.9540	-0.142	-	-0.017	-	1.224	-
UGP1-1	0.0330	0.1812	0.4003	0.0132	1.479	1.650	2.380	2.260	-1.208	-1.320
UGP1-2	-	-	-	-						
UGP1-3	0.1436	0.1939	0.4888	0.0664	-2.530	0.045	-2.770	0.176	-1.936	0.144
UGP2-1	0.6801	0.1000	0.8030	0.3936	-0.327	-0.364	-0.468	-0.534	-0.176	-0.513
UGP2-2	0.9464	0.8568	0.9513	0.7264	0.226	0.049	0.520	0.023	-0.136	0.077
UGP2-3	0.9174	0.5057	0.7203	0.8637	-1.441	-0.031	-0.949	-0.031	-0.991	-0.172
UGP2-4	0.0563	0.2766	0.2189	0.0394	3.116	-	1.910	-	2.056	-
UGP3-1	0.4000	0.6234	0.3158	0.4197	-0.373	-0.809	1.573	1.810	0.719	0.555
FUM2-1	0.7271	0.8608	0.4599	0.7813	-0.994	-0.917	-0.050	0.105	0.304	0.322
FUM2-2	0.9095	0.9339	0.9374	0.6464	0.197	-0.060	0.068	0.208	0.054	-0.054
FUM2-3	0.2183	-	0.0795	0.7081	7.106	3.262	0.744	-0.745	-	-
ADH2-1	0.1822	0.0290	0.1728	0.2560	-0.574	-0.855	0.100	0.423	0.677	0.297
ADH2-2	0.1927	0.0227	0.1862	0.2560	0.584	-0.260	-0.100	0.323	-0.686	-0.400
ADH2-3	-	-	-	-						
PGI2-1	0.8227	0.1962	0.4928	0.9479	2.530	2.560	-0.022	-0.139	-0.696	-0.290
PGI2-2	-	-	-	-						
PGI2-3	0.5857	0.0214	0.4184	0.5194	-0.335	-	-0.413	-	0.926	-
PGI2-4	-	-	-	-						
MDH1-1	-	-	-	-						
MDH1-2	0.1915	0.6227	0.3010	0.1350	1.377	-	-1.127	-	-0.479	-
MDH1-3	-	-	-	-						
MDH3-1	0.3156	0.6156	0.2556	0.3666	-0.781	-0.329	-0.667	-0.421	0.516	0.527
MDH3-2	0.0998	0.3879	0.1331	0.1533	2.761	0.533	0.704	0.601	-1.720	-0.159
MDH3-3	-	-	-	-						
MDH3-4	0.8841	0.9555	0.8075	0.6922	-1.000	-0.216	0.835	-0.327	0.198	0.093
MDH4-1										

area by several populations that had differentiated elsewhere). In the first two models, selection brings about differentiation while the last two imply stochastic processes. Real situations in nature will probably constitute combinations of two or more types.

Since statistical heterogeneity of values and geographic pattern are potentially independent, various outcomes from spatial autocorrelation analysis are possible. SOKAL and ODEN (1978b) discussed the implications of the various possible outcomes. Two outcomes are of particular importance:

A) Significant geographic heterogeneity accompanied by significant spatial patterns is the most common combination in geographic variation analysis. Three major situations could lead to such an outcome:

1) Migration will cause similarity between neighboring demes whereas distant demes differ for the variable studied

2) Occurrence of frequent local extinction followed by recolonization from neighboring colonies will cause similarity between neighboring demes

3) ***Demes track the environment closely,*** responding by adaptation, while the selective agents are patterned in space, either as small patches or arranged as gradients

The patterns may further be classified regarding their similarity among the different variables:

I) ***Similar surface patterns***: If patterns of variation in allele frequencies are very similar, their correlograms would by necessity be similar. SOKAL and WARTENBERG (1981) list three biological reasons for pattern similarity:

1) ***Functional associations*** between the variables in question
2) ***Migration*** with alleles studied tightly linked or diffusion between populations differing for the variables in question. If populations differ initially in many of their characteristics, diffusion patterns based on dispersal and migration of individuals from these populations will necessarily affect not just one but all the properties of these populations. Hence, this property will result in correlations between variables in which initial populations differ
3) ***The similar variables track the same environmental patterns***

II) ***Dissimilar surface patterns among variables*** are explained by these authors as due to:

1) ***Random processes*** such as genetic drift
2) ***Migration*** at different rates from differing source populations
3) ***The tracking of different environments by different variables***. Dissimilar patterns typically yield dissimilar correlograms, reflecting their different origin.

For the interpretation of patterns, a subtle but important distinction must be made between the variation pattern itself and its summarization by means of a correlogram (SOKAL and ODEN, 1978b). Similar patterns should yield similar correlograms. However, similar correlograms do not necessarily mean parallel patterns. Correlograms summarize the relationship among values of the space, but identical relationships may be realized from quite different patterns. It is easily seen, for example, that a north-south cline and a west-east cline would result in the same correlogram while the two variation patterns clearly differ. Spatial correlograms thus describe the underlying spatial relationships of a surface; they probably are closer guides to some processes that have generated the surface than the surface itself. Consequently, while identical variation patterns will result in identical correlograms, different variation patterns may or may not yield different correlograms. Conversely, different correlograms must be based on different patterns, but identical correlograms may result from identical or different patterns.

When allele frequencies in a given area show the same variation pattern and therefore identical correlograms, it would suggest a common response to selective agents that differ over the area. Different variation patterns yielding the same correlograms should occur in populations where migration strongly affects geographic variation of allele frequencies. Since a migrant individual must carry all its genes with it, autocorrelation should be the same for all loci and correlograms should resemble each other strongly. Different geographic variation patterns associated with different correlograms should most likely be found in populations in which differential selection among the loci predominates. Autocorrelation may, however, also differ among loci for historical reasons, such as invasion by populations differing in allele frequencies at some loci but not at others.

B) ***Significant geographic heterogeneities in the absence of patterns***, as a second combination, could occur in at least three situations:

1) ***Random processes*** such as genetic drift determines the character
2) Occurrence of ***frequent local extinction*** followed by establishment of new colonies by nearby founders

3) ***selective agents are*** either ***unpatterned*** or patterned such that the patterns are composed of components smaller than average interlocality distance or greater than the distance classes.

Results reported in *section 4.2* uncovered significant heterogeneities in allele frequencies among localities for several alleles and at different levels of sampling. As regards the total sampling area, significant geographic heterogeneities in allele frequencies were found between the two regions for 16 loci. With only two exceptions (*Pgi-1*, *Mdh-1*), one to three alleles at all these loci showed significant spatial patterns in the allele frequency surface over the total area. Moreover, considerable heterogeneity was found among the seed zones within the total sampling area. A high subpopulation differentiation was found for most of the loci exhibiting spatial structures. Hence, significant heterogeneities of allele frequencies are combined with significant spatial patterns for most of the alleles in question. According to SOKAL and ODEN (1978b), these combination may be due to three possible causes *i.e.* migration, local extinction or adaptation to the environment.

While frequent local extinction and recolonization seem rather unlikely due to the short immigration history of the species in the area, migration and adaptation are both possible explanations. Theoretically, migration may have come about from different sources at different rates during colonization. In this case, different variation patterns yielding similar correlograms would be expected. In other words, if migration was a major force in generating the patterns, a close resemblance among the correlograms would be expected. Since different variation patterns seem to be associated with quite different correlograms for many of the alleles in question, the observed patterns are more likely to be a result of differential selection *i.e.* of different alleles responding differently to the environment than the result of migration processes. It must be emphasized, however, that comparisons between spatial structures are not easily made, as there is no test available for the differences between correlograms (SOKAL and WARTENBERG, 1983). Consequently, correlograms can only be compared visually. Visual inspections of correlograms lead to the conclusion that correlograms differed substantially among the different alleles. A second valid explanation for the observed patterns would be a difference in the evolutionary history of the past. Observed differences in patterns could be the result of a colonization of the area (maybe the different regions) by populations (races) which differed in the frequencies of the alleles in question, but had similar frequencies for the other alleles.

If adaptation were a major force in generating the observed patterns, similar spatial structures should be found within each of the two regions as well as over the total sampling area since climate not only varies within but also differs substantially between the two regions. If, on the other hand, a different evolutionary history were responsible for the observed spatial structures, the patterns in the two regions are not expected to be similar.

Southwest Oregon showed distinct patterns for several allele frequencies combined with clearly differing correlograms. Since migration from different sources, as a possible explanation, is less probable (but not excluded) in this relatively small area, adaptation to the environment seems to be the most likely explanation for the observed patterns. Participation of historical events may not be ruled out completely, however. Most of the patterned alleles within Southwest Oregon were identical to the ones exhibiting a spatial structure over the total sampling area, arguing in favor of adaptation as a major force. Allele frequencies exhibiting spatial structures within Southwest Oregon but lacking such a structure over the total area are likely candidates for a historically caused pattern of variation.

In contrast, Central Oregon showed an extensive geographic differentiation of most allele frequencies but spatial patterning was scarce. Most allele frequencies seemed to vary randomly. Of the three possible situations leading to significant geographic heterogeneities in allele frequencies combined with a lack of spatial patterning (SOKAL and ODEN, 1978b), genetic drift and frequent local extinction followed by new recolonizations do not seem very likely. Extended population size combined with gene flow should prevent genetic drift while local extinction seem unlikely regarding the short immigration history in the area. Hence, the lack of geographic patterns is most likely caused by the spatial structure of the selective agents or by the methodological limitations to detect the patterns. Two possible explanations may be considered to explain the lack of spatial patterns. Selective agents are either unpatterned or patterned

such that the patterns are composed by components smaller than average interlocality distance or greater than the distance classes used for the analysis.

Climate conditions vary substantially among and within the regions. Climate, however, is only weakly associated with geographic location. The Mantel test of matrix associations clearly showed that climate distance and geographic distance are only weakly correlated, sharing about 13% to 15% of variance only. Since elevation largely influences temperature and precipitation, and since elevation varies substantially over quite small distances, especially in Central Oregon, climate conditions tend to vary locally and in relation to local topography. In Central Oregon, the same climate conditions tend to occur again in different geographic locations over the area and patchy structures are scarce (*Figure 27, p. 123*). Southwest Oregon, on the other hand, is characterized by a rather patchy spatial arrangement of climatic conditions *i.e.* geographically adjoining areas are characterized by similar climatic conditions over rather large distances. Climatic properties of the eight groups shown in *Figure 27, p. 123* will be described in detail in a later section.

Therefore, lack of patterns in Central Oregon is very likely a reflection of this rather fine-grained structure of climate conditions and of the methodological limitations for detecting such variation patterns. Consequently, the results of spatial autocorrelation within Central Oregon do not necessarily argue against adaptation as a likely explanation for the observed pattern. Selection and adaptation to climate conditions, which seem to be a plausible explanation for the patterns found within Southwest Oregon, apparently do not produce clear spatial patterns of variation within Central Oregon, because the selective agents are not patterned in gradients or in distinct patches of detectable size as is the case within Southwest Oregon.

In the overall analysis, 18 out of 26 alleles were characterized by a clinal variation pattern. In a regular gradient such as a cline there will be strong positive autocorrelation in the shorter distance classes, since near neighbors resemble each other. High positive autocorrelations can

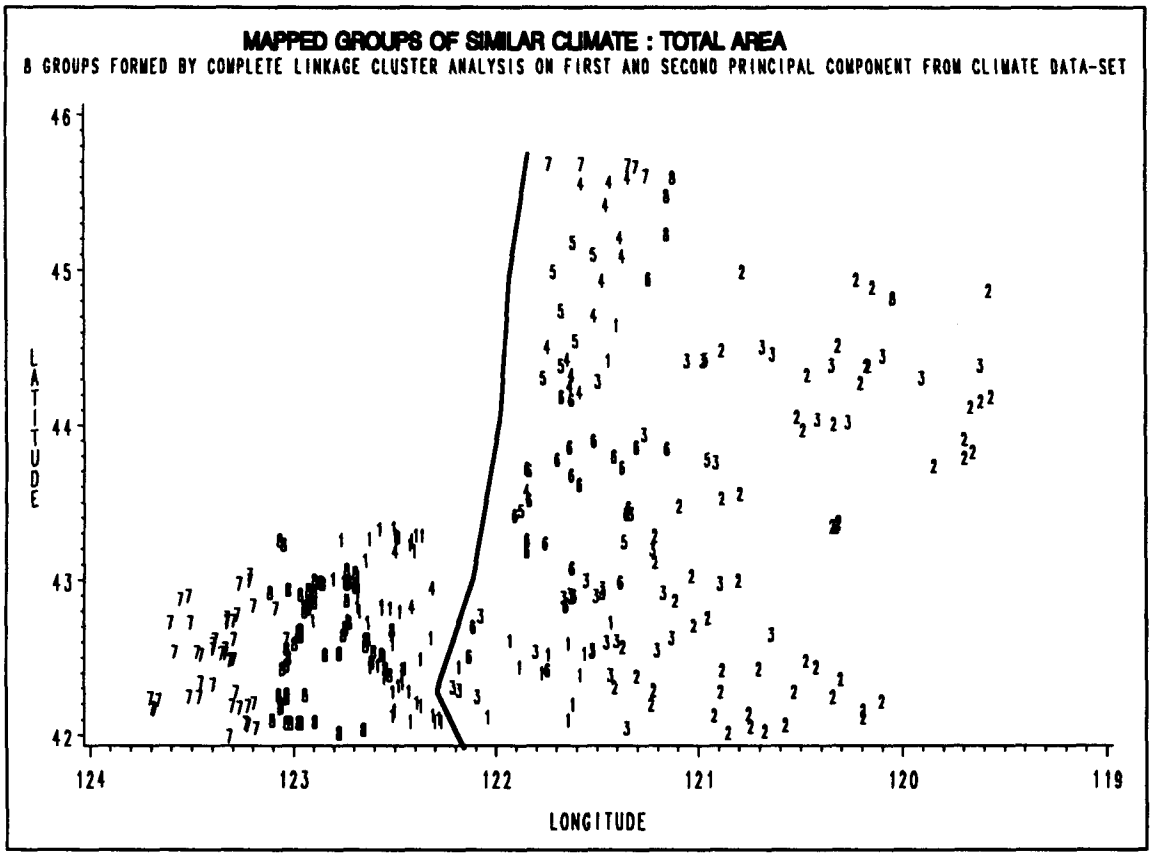


Figure 27: Mapped groups of similar climate conditions. 8 groups formed by complete linkage cluster analysis on climate Factor 1 and Factor 2

also be expected, however, if the study area is divided into two large contiguous patches each comprising numerous localities with homogeneous high values assigned to one patch while homogeneous low values are assigned to the other patch. In such a case, the number of tree pairs falling within one patch may exceed the number of trees falling into different patches thus producing a clinal pattern in the correlogram which in fact is a result of two big patches differing in the character studied. This latter is indeed the case in our situation. Thus, clinal variation patterns in the correlograms are the consequence of relatively similar frequencies within the regions relative to the overall variation, if both regions are combined. Consequently, alleles with clinal variation patterns primarily reflect an existing regional difference in the allele frequencies. Patch sizes which equal the size of the two regions support this assumption.

Significant positive autocorrelations in the first distance class *i.e.* trees which deviate in their allelic structure from the mean in the same direction within a rather short distance, were manifest for several alleles. Since individuals were sampled between 8 and 30 km apart, it is highly unlikely that these similarities in allele frequencies are due to relationships among individuals. Individuals, however, could be more similar within shorter distances because they may have been derived from the same founder at the same point in time. Alternatively, similarities in allelic structure may be the result of environmental conditions, being rather identical in the nearest neighborhood, which in turn could favor small scale patchy genetic structures as a result of adaptation to these conditions. In the latter case, alleles exhibiting such spatial structures are likely candidates for adaptive markers. Based on autocorrelation analysis, *Mnr1-1*, *Mnr1-2*, *Mnr2-1*, *Mnr2-2*, *Lap2-1*, *Lap2-2*, *Pep1-1*, *Pep3-1*, *Got1-3*, *G6p2-1*, *G6p2-3*, *Acp1-2*, *Acp1-4*, *Gdh1-1*, *Gdh1-2*, *Idh1-1*, *Pgm1-1*, *Aco1-3*, *Skd1-3*, *Skd2-1*, *Skd2-2*, *Fum1-3*, *Adh2-1*, *Adh2-2* and *Pgi2-3* are therefore likely candidates for adaptive markers.

No departures from random spatial arrangement were detected using Mantel test statistics. Associations among genetic and geographic distance were mostly not significant, suggesting that allele frequencies vary randomly over geographic space. Moreover, results did not reveal any substantial association between genetic and climatic variation. These results clearly contrast with the conclusions reached from spatial autocorrelation analysis. Several reasons may be responsible for these contrasting results. First of all, the two precedures are testing different patterns and results are therefore not entirely comparable. While spatial autocorrelation analysis is looking for clustering of single-locus genotypes in two dimensional space, Mantel test statistics are looking for associations between average genetic distance between 2 locations (over many loci) and spatial or climatic distance. Distance measures condense information into a composite index by summarizing and averaging over many variables. As may be seen from the results of spatial autocorrelation analysis, patterns of variation differ among alleles and among loci; such differences cannot be portrayed by a composite index. Moreover, performance of Mantel test statistic strongly depends on the patterns and on the geographic distance model used in the analysis (SOKAL, 1979a). If, for example, the nature of the process is a first order one *i.e.* only near neighbors affect the values of the variable at a given locality, then only a near-neighbor adjacency matrix will lead to significant associations. However, if the process works on a continuous surface between all pairs of localities, an ordinary geographic distance matrix should be used. If the observed pattern is characterized by low order, short distance positive spatial autocorrelation and a continuous distance matrix is used, a poor association between the two matrices will result since the pattern is far too complex to be a simple function of geographic distance. Furthermore, it appears, that Mantel test statistics is a far more conservative test than spatial correlograms. SOKAL et al. (1980), investigating geographic variation patterns of various traits in *Pemphigus populicaulis* with autocorrelation as well as Mantel test statistics, found that only a few of the spatial patterns resulting from correlograms yielded a significant Mantel test. The same conclusion was reached by JONES et al. (1980) in a study of patterns of morphological and molecular polymorphism in *Cepaea nemoralis* in Britain. Our results support this conclusion. Visual inspection of variation patterns, as well as results from autocorrelation analysis, clearly demonstrated the existence of distinct spatial structures for several alleles at many loci. Results of matrix comparisons, on the other hand, did not reveal any significant associations among genetic variation and geographic location. Not even the comparisons focusing on the 15 loci with clear spatial variation pattern produced any significant Mantel test statistic. It may be speculated that the observed patterns of variation are too complex to be reflected in Mantel test statistics.

It thus appears that spatial autocorrelation is better suited as a test for randomness of variation patterns and for describing geographic surfaces of allele frequency variation than matrix comparisons. Spatial autocorrelation analysis is able to uncover patterns of variation which remain undetected in Mantel test statistics. In spite of these limitations, matrix comparisons have the advantage of relating genetic variation to environmental variation. Even if the test is not very sensitive, the significant results seem to confirm a relationship between genetic variation and temperature. However, due to the low sensitivity, the amount of association is most likely underestimated.

Multinomial response models are looking for relationships between allelic makeup and climate at source locations. Consequently, results of multionomial response models may differ from results obtained by spatial autocorrelation analysis, since the former is relating allele frequencies to climate and the latter to spatial patterns. To the extent that climate and spatial patterns do not coincide, the results of the two analyses can be quite different. In Southwest Oregon, spatial clustering has been found for several alleles but only one allele appeared to be associated with climate (*Factor 1*). In Central Oregon, on the other hand, no clustering of alleles could be observed, while frequencies of several alleles appeared to be associated with climate (*Factor 1*). In Southwest Oregon, the spatial clustering of alleles coincides with the patterned distribution of environments. This coincidence may be interpreted as evidence for adaptation to the environments. If the clustering of alleles were in fact a response of the environmental clustering, then the frequencies of such alleles are expected to be associated with climate conditions. Since no such significant associations were found in the response models, no final conclusion can be drawn, however. As has been shown earlier, the range of climate variation in Southwest Oregon is less than that in Central Oregon. Associations between allele frequencies and climate may thus be more difficult to detect in Southwest Oregon. Alternatively, clustering of alleles may have been caused by other forces than differential selection.

In contrast, the significant multinomial response equations in Central Oregon and their consistency (regarding signs and magnitudes of parameters) with Southwest Oregon are compelling evidence that allozyme patterns of variation are partly the result of selection. Because climates are not spatially clustered in Central Oregon, clustering of alleles are not expected. Hence, lacking spatial structures do not argue against adaption to climate conditions. On the contrary, since climates are not spatially structured, the observed associations between allele frequencies and climate are hard to explain with other factors than selection.

Based on multinomial response models, associations between climate conditions and genotypic frequencies of single alleles may be postulated for several alleles at many loci. It appears that ponderosa pine is primarily adapted to temperature conditions at source location since all significant associations were dominated by climate *Factor 1*. No associations were found among genotypic frequencies and climate *Factor 2*. This result may be interpreted as a lack of adaptation of ponderosa pine to the water balance of the site. Alternatively, an existing adaptation to hydrological site conditions may not be reflected by the marker genes used in the analysis. A dominating influence of temperature was also reflected in the results of Mantel test statistics. Even if correlations between genetic and climatic distance were weak, correlations with *Factor 1* were significant, but non significant with *Factor 2*.

Results from multinomial response models indicate associations between genotypic frequencies and temperature for alleles at *Mnr-2*, *Pep-3*, *Mpi-1*, *Idh-1*, *Pgm-1*, *Aco-1*, *Skd-2*, *Ugp-1*, *Ugp2* and *Mdh-3*. It has been shown with nested models, that some of the associations found in the former analyses were most likely caused by indirect effects. Since climate differences among the regions may parallel other regional differences unrelated to climate, some of the association may in fact not be due to adaptation to temperature itself but may be caused by such parallel effects. Results from nested models suggest, that *Mnr-1*, *Acp-1*, *Adh-2*, *Lap-2* and *Fdp-2* are likely candidates for such indirect effects. Differentiation found at these loci most likely is due to other effects such as historical events. Unfortunately, effects may not be separated completely because climate has a strong regional effect such that all other regional effects may mimic adaptation to climate conditions.

5. 5 Summary

Several alleles at many loci show spatial variation patterns which clearly deviate from random spatial arrangement. Significant heterogeneities of allele frequencies, significant spatial patterns and observed associations between genotypic frequencies and climate for the same alleles suggest that adaptation to the environment is the most likely cause for the observed variation patterns. Although other causes can not be ruled out, patterns are most likely the result of differential selection for most of the alleles which exhibit a spatial pattern. Associations of allele frequencies with temperature could be confirmed by results from the various analyses using different analytical tools. Temperature seems to be an important selective agent responsible for adaptation. Genetic structure, on the other hand, seems unrelated to water balance of the site. Lack of adaptation of ponderosa pine to hydrological conditions or the inability of the marker genes to portray adaptation to water regime may be plausible explanations for these findings. Results do not allow a final conclusion about the adaptivness of alleles, however, since climate effects cannot be separated from other effects. Nested multinomial response models suggest that some of the associations among genotypic frequencies and temperature are most likely caused by indirect effects. According to the results from nested response models, *Mnr-2*, *Pep-3*, *Mpi-1*, *Idh-1*, *Pgm-1*, *Aco-1*, *Skd-2*, *Ugp-1*, *Ugp-2* and *Mdh-3* are most likely associated with temperature at source location. Associations of *Mnr-1*, *Acp-1*, *Adh-2*, *Lap-2* and *Fdp-2* with climate conditions, however, are most likely caused by other than climate effects. Such effects are indirectly related to climate conditions and thus mimic associations with temperature conditions, although temperature and genotypic frequencies are in fact not associated.

6. Multivariate patterns of variation based on allozyme scores and associations with climate conditions

Multivariate patterns of allozyme variation and possible associations of multilocus genetic structures with habitat conditions may be investigated using different multivariate statistical techniques, which provide a way of handling genetic variants at all 31 loci at the same time. Multivariate techniques offer a means of detecting and quantifying truly multivariate genetic patterns that arise from the correlation structure of the complete genetic data set of all 71 allozyme variables. When geographic patterns of more than one variable are available, one approach to summarize the multivariate information is to combine the information into few synthetic variables derived by a linear combination of the original variables, possibly without any loss of information. Numerous methods of ordination may be applied for this purpose. The reduced set of variables may then be regressed against location variables, obtaining different trend-surfaces of the response variables which illustrate the variation pattern of the multivariate data set. Likewise, derived linear combinations of the original variables may be regressed against environmental variables in order to assess associations between genotypes and environment. Instead of this two-step, indirect analysis (TER BRAAK, 1986), ordination and regression may be jointly applied in one analysis, summarizing the multivariate information of the response and predictor data sets into few synthetic variables (variates) and at the same time obtaining the best possible correlations among the variates of the response and the predictor data sets. This direct analysis may be achieved by canonical correlation analysis (also called canonical variate analysis). All methods described above generally assume continuity of the geographically distributed variables but in practice they can be applied to discontinuous distributions as well.

A different set of techniques is based on the philosophy of categorization. These methods search for homogeneous sets or clusters of samples either within the multivariate data sets of the response or of the predictor variables and may help in the establishment of artificial boundaries between discontinuous groups. Various multiple comparison methods or cluster analyses may be used to find such homogeneous sets of samples. Sets of samples derived from these methods may further be evaluated by discriminant analysis, in order to either describe differences among clusters in the multivariate data set or to derive functions which may be used to classify samples into the respective subsets based on their multivariate data set. These methods generally assume discontinuity of the geographically distributed variables but may also be applied to distinguish patterns within continuous distributions.

The various techniques all assume the existence of a "pattern" whether this be one of uniform change or a patchy distribution of differing homogeneous subsets of samples. The existence of patterns for several allele frequencies in the sampling area was shown in *sections 4 and 5.* Multivariate techniques will be used to further describe patterns on a multilocus level and to investigate associations between multilocus variables and environment. If such associations exist, environmentally related multilocus response surfaces may then be developed and used to illustrate patterns of adaptation. The response surfaces may also be used to illustrate relative seed transfer-risk among locations within the sampling area. Response variates may be scaled to represent certain degrees of relative transfer risks, thus producing contour maps which can serve as guidelines for the safe movement of genotypes in reforestation programs. Hence, the two main objectives that will be adressed in this section are:

1) to determine whether multilocus patterns of allozyme variation are associated with environments over the entire sampling area as well as within both regions
2) if 1) is true, to use these adaptive patterns to assess transfer-risks and to develop seed transfer guidelines.

In order to base results and interpretations on safe grounds, associations between genotypes and environment are examined using different multivariate techniques. Because every technique has its limitations or makes certain assumptions which may not be met by the data, this approach helps to add more certainty to results and interpretations.

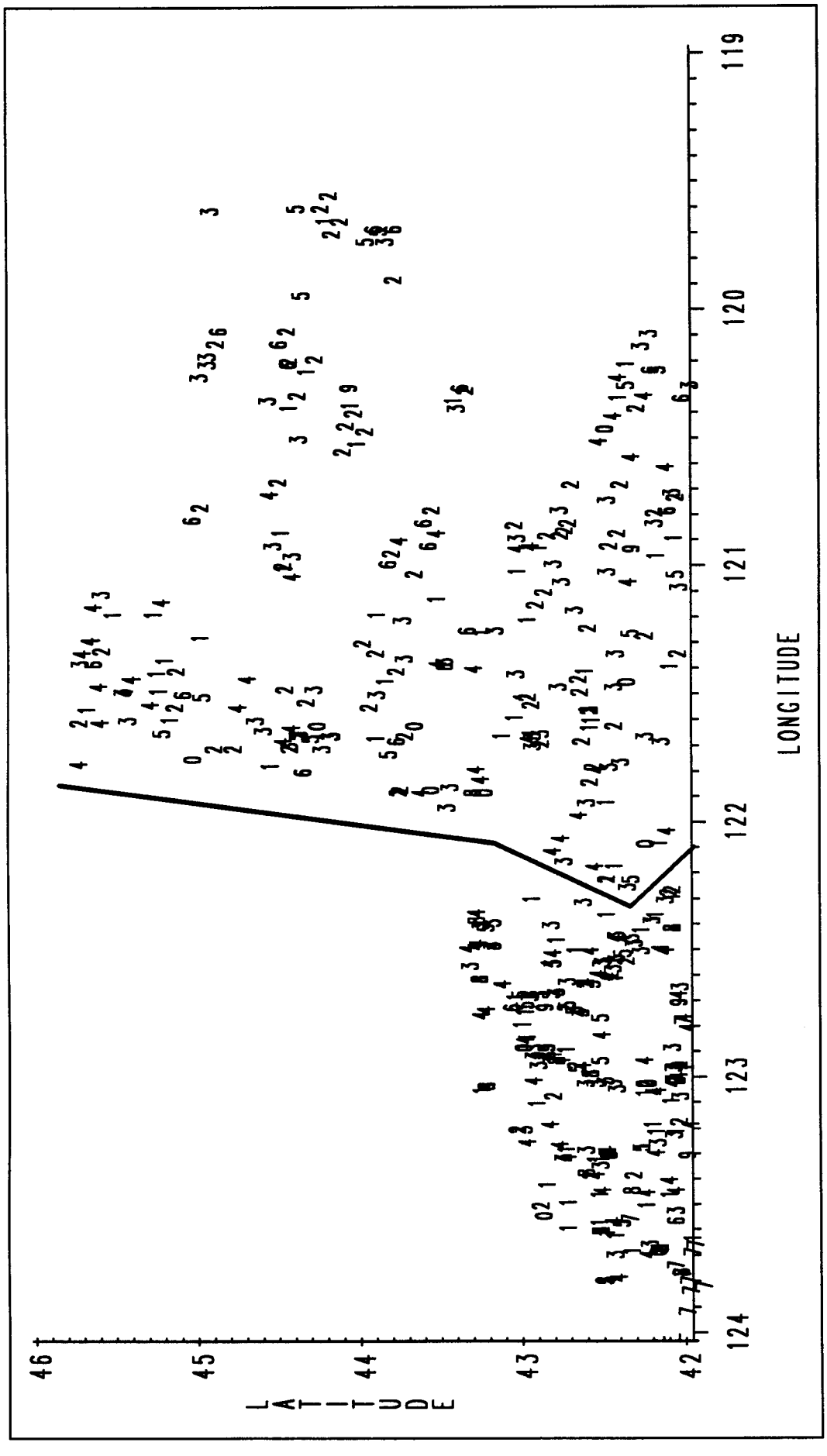


***Figure 28**: Mapped locations of genotypes found within each of 10 clusters of similar genotypes (0-9), determined by complete-linkage cluster analysis of genetic distance [Euclidean distance] among 488 individuals (based on 71 allozyme variables)*

6.1 Associations of multilocus genotypes with climate conditions, inferred by cluster and discriminant analysis

If multilocus genetic structure of ponderosa pine in Oregon were associated with habitat conditions, then similar multilocus genotypes should be found in similar environments. In order to analyze associations between genotypes and environment, cluster analysis may be used to find homogeneous sets of either similar genotypes or similar habitat conditions, accompanied by discriminant analysis which is used to describe differences among the homogeneous sets in terms of the opposite data set. Samples can be grouped into clusters of genetically similar individuals, based on their multilocus genotypes, and variation in environmental conditions analyzed in terms of within- to among-group variance, using discriminant analysis. Associations between the multilocus genetic structures and environmental conditions may be estimated based on the proportion of variation associated with group differences. Alternatively, locations of sample trees may be categorized according to their similarity in habitat conditions and differences in multilocus genetic variance among-groups investigated. Genetic variants may then be described in terms of their association with differences in habitat conditions

Ten subsets of individuals with similar multilocus genotypes were formed by cluster analysis based on all 71 allozyme variables. A complete linkage algorithm was used to fuse individuals to clusters based on either genetic distance by GREGORIUS (1974), genetic distance by NEI (1978) or Euclidean distance among all pairs of individuals. Complete linkage algorithm minimizes dissimilarities between neighbors thus producing very homogeneous clusters of genetically similar individuals.

Figure 28, p. 128, depicts the geographic distribution of the 10 groups based on among sample Euclidean distance and *Table 24* lists the number of individuals and the regional distribution found for each subset. Similar multilocus genotypes occurred all over the sampling area. With one exception, patchy geographic distributions were not apparent. Several subsets of genotypes occurred in both regions at about similar frequencies while other genotypes were more frequent in one of the regions. One subset (*group 7)* was found exclusively within Southwest Oregon, located only in the southwestern corner of the sampling area. This part of the area (seed zone 90), has already been shown to be genetically distinct in former analyses. Similar geographic distributions were observed for cluster analyses based on genetic distance by NEI (1978) or by GREGORIUS (1974).

To test for associations of multilocus genotypes with habitat conditions, a number of descriptive discriminant analyses were performed on the subsets of similar genotypes which were revealed by cluster analysis. Climate *Factor 1* and *Factor 2*, second order terms and cross-products were used as discriminating variables in the analyses. However, adding second order terms and cross-products did not improve nor change the results in any of the analyses and these terms were dropped from subsequent analyses. Significance levels of the discriminating variables were estimated using covariance-controlled partial *F* -ratios for each variable. In partial *F* - ratios, the variance which is already explained by the other variables is partitioned out of the variable of interest such that the *F* - ratio indicates the significance of the respective variable as discriminator after the impact of other variables has been removed.

Results for the 10 subsets formed by cluster analyses on genetic distance by GREGORIUS (1974) and on Euclidean distance are furnished in *Table 25, p. 130* for the total sampling area and the two regions separately. Results involving genetic distance by NEI (1978) are not shown

GROUP / REGION		1	2	3	4	5	6	7	8	9	10	TOTAL
SOUTHWEST	N	61	19	54	54	11	3	10	15	4	5	236
	%	56.48	22.35	48.21	58.06	47.66	13.64	100.00	93.75	66.67	38.12	48.36
CASCADES	N	47	66	58	39	12	19	0	1	2	8	252
	%	43.52	77.65	51.79	41.94	52.17	86.36	0.00	6.25	33.33	61.54	51.64

***Table 24**: Groups of similar genotypes and respective frequencies for the two regions*

	TOTAL SAMPLING AREA									
	Genetic distance by GREGORIUS [31 loci]					Euclidean distance [31 loci]				
Climate	CAN 1		CAN 2		F-ratio	CAN 1		CAN 2		F-ratio
variables	struc.	% red.	struc.	% red.	Prob. p	struc.	% red.	struc.	% red.	Prob. p
PCA - Factor 1	0.9314	7.43	0.3638	0.16	0.0001	0.9528	12.02	-0.3034	0.07	0.0001
PCA - Factor 2	-0.3638	1.13	0.9314	1.05	0.0372	0.3034	1.22	0.9528	0.67	0.1247
Eigenvalue	0.1177		0.0257			0.1774		0.0207		
Percent varianc	82.06		17.94			89.55		10.45		
Can. correlatio	0.2927		0.1102			0.3638		0.0861		
r square	0.0857		0.0121			0.1324		0.0074		
Significance	0.0001		0.2471			0.0001		0.4065		

	SOUTHWEST OREGON									
	Genetic distance by GREGORIUS [31 loci]					Euclidean distance [31 loci]				
Climate	CAN 1		CAN 2		F-ratio	CAN 1		CAN 2		F-ratio
variables	struc.	% red.	struc.	% red.	Prob. p	struc.	% red.	struc.	% red.	Prob. p
PCA - Factor 1	0.7877	3.25	0.6160	1.33	0.4753	0.8831	5.71	-0.4691	1.12	0.1667
PCA - Factor 2	-0.5244	1.44	0.8514	2.53	0.5593	0.5646	2.33	0.8253	3.46	0.3168
Eigenvalue	0.0557		0.0363			0.0790		0.0535		
Percent varianc	60.56		39.44			59.61		40.39		
Can. correlatio	0.2296		0.1870			0.2706		0.2254		
r square	0.0527		0.0350			0.0732		0.0508		
Significance	0.5684		0.5932			0.1781		0.3025		

	CENTRAL OREGON									
	Genetic distance by GREGORIUS [31 loci]					Euclidean distance [31 loci]				
Climate	CAN 1		CAN 2		F-ratio	CAN 1		CAN 2		F-ratio
variables	struc.	% red.	struc.	% red.	Prob. p	struc.	% red.	struc.	% red.	Prob. p
PCA - Factor 1	0.7239	2.13	0.6898	1.11	0.1035	0.9275	10.48	-0.3737	0.38	0.0001
PCA - Factor 2	-0.7914	2.54	0.6112	0.87	0.1915	0.2466	0.74	0.9690	2.55	0.0970
Eigenvalue	0.0868		0.0438			0.1856		0.0528		
Percent varianc	66.47		33.53			77.84		22.16		
Can. correlatio	0.2014		0.1524			0.3491		0.1649		
r square	0.0406		0.0232			0.1219		0.0272		
Significance	0.0796		0.3286			0.0002		0.1979		

CANONICAL DISCRIMINANT ANALYSIS ON GROUPS OF INDIVIDUALS WITH SIMILAR GENOTYPES:

- Grouping variables:
  10 clusters of similar genotypes formed by complete linkage cluster analysis on among sample matrix of:
  - genetic distance by GREGORIUS (= proportion of alleles not shared by each other), 31 loci
  - Euclidean distance:

$$ED_{jk} = \sqrt{\sum_{i=1}^{p} (x_{ij} - x_{ik})^2}$$

  with: x=allele-frequencies, p=total number of alleles [N=71], j and k=jth and kth individual

- Discriminating variables:
  First and second principal component scores (Factor 1 and Factor 2) from climate data set. Factor 1 represents temperature regime, Factor 2 water balance of the site

- struc: structure coefficients:
  simple bivariate correlations between variables and the canonical functions

- % red: redundancy: Proportion of variation in respectiv variable contributing to group differences

- F-ratio: covariance controlled partial F-ratio and associated level of significance for dicriminating variables

Table 25: *Association of genotypes with climate: Canonical discriminant analysis on groups of similar genotypes using climate variables as discriminators. Groups of similar genotypes formed by complete linkage cluster analysis of genetic distance among individuals [genetic distance by GREGORIUS (1974), Euclidean distance]. Discriminating variables: Climate Factor 1 and Factor 2*

because they were very similar to results involving genetic distance by GREGORIUS (1974). The 10 subsets used for the analyses within the two regions were formed based on the regional samples only.

The first variate generated by canonical discriminant analyses performed on the total sampling area was highly significant ($p < 0.0001$) for both methods of group formation. The variates accounted for 82% or 89% of total variance in climate variables. Canonical R^2, however, indicated that only 8.6% (GREGROIUS distance) or 13.2% (Euclidean distance) of the variation in the first canonical function were explained by differences in group means. Thus, the maximum among-group variance of climate variables was small compared to the variance found within the groups. It may be concluded from this result that similar multilocus genotypes are growing under highly variable habitat conditions. Based on the ratio of among to within-group variance, only 8% to 13% of all similar genotypes were found in similar habitats.

Structure coefficients, redundancies and significance levels of *F* - ratios showed a significant contribution of climate *Factor 1* to group separation. However, the proportion of variance in climate *Factor 1* associated with group differences was only 7.4% (GREGORIUS distance) or 12% for the subsets based on Euclidean distance. Climate *Factor 2* did not contribute to differences among the subsets. Although significant ($p = 0.037$) in the analysis involving subsets based on GREGORIUS distance, contribution of climate *Factor 2* was very low in both analyses (1.1% or 1.2%). Hence, multilocus genotypes were significantly but only weakly associated with temperature. Of all similar genotypes (Euclidean distance), 12% grew under similar temperature conditions while no associations were found with moisture characteristics of the site.

Discriminant analysis, performed on the samples from Southwest Oregon only, produced no significant canonical variates for both methods of group formation, suggesting that genotypes are not associated with habitat conditions (*Table 25, p. 130*). Even if the small sample size ($N = 189$) is partly responsible for the lack of significance, resulting correlation coefficients (5% to 7%) and redundancies (1.4% to 5.7%) clearly indicated only weak associations.

The first variate from discriminant analyses performed on the Central Oregon sample was highly significant ($p = 0.0002$) for the groups based on Euclidean distance while it was not significant for the groups based on GREGORIUS (1974) distance (*Table 25, p. 130*). Canonical R^2 of the significant variate, which accounted for 78% of climate variance, indicated that 12.2% of the variation of the first variate was associated with differences among the groups. Structure coefficients, redundancies and *F* - ratios showed a significant ($p < 0.0001$), but weak association among multilocus genotypes and climate *Factor 1*. Ten percent of similar genotypes were growing under similar temperature conditions, while no significant association was found with climate *Factor 2*.

Climate conditions vary considerably within the sampling area. Similar climate conditions tend do reoccur over the area in different locations and at different elevations. *Figure 29, p. 132* illustrates climate conditions of the seed zones, plotted as seed zone mean values of both climate factors. The axis of climate *Factor 1*, depicting temperature regime, clearly separates the two regions. Axis 2, illustrating hydrological conditions, separates areas with wetter conditions from drier habitats. It can be seen that 8 to 10 types of specific habitat conditions, encompassing 2 to 8 seed zones each, emerge when both factors are combined. Similar habitat conditions occur again at different locations within the sampling area. For example, cool and dry conditions (*aggregate number 4*) are found in the northern as well as in the southern parts of Central Oregon.

A similar but even more complex picture of climate conditions within the sampling area emerges from complete linkage cluster analysis, grouping similar climate conditions based on climate factors *Factor 1 and Factor 2* into 8 homogeneous groups of habitat conditions. While Southwest Oregon is characterized by a patchy structure of 3 geographically distinct habitats, conditions are more variable within Central Oregon (*Figure 30, p. 133*). Climate conditions vary on quite small geographic distances and similar conditions re-occur over the area, especially in the Cascade Mountains of Central Oregon. Climate becomes more homogeneous in the high desert country to the east. *Table 26, p. 134* characterizes the habitat conditions of the 8 habitats. The 3 groups found within Southwest Oregon (1, 7, 8) are all characterized by mild winter

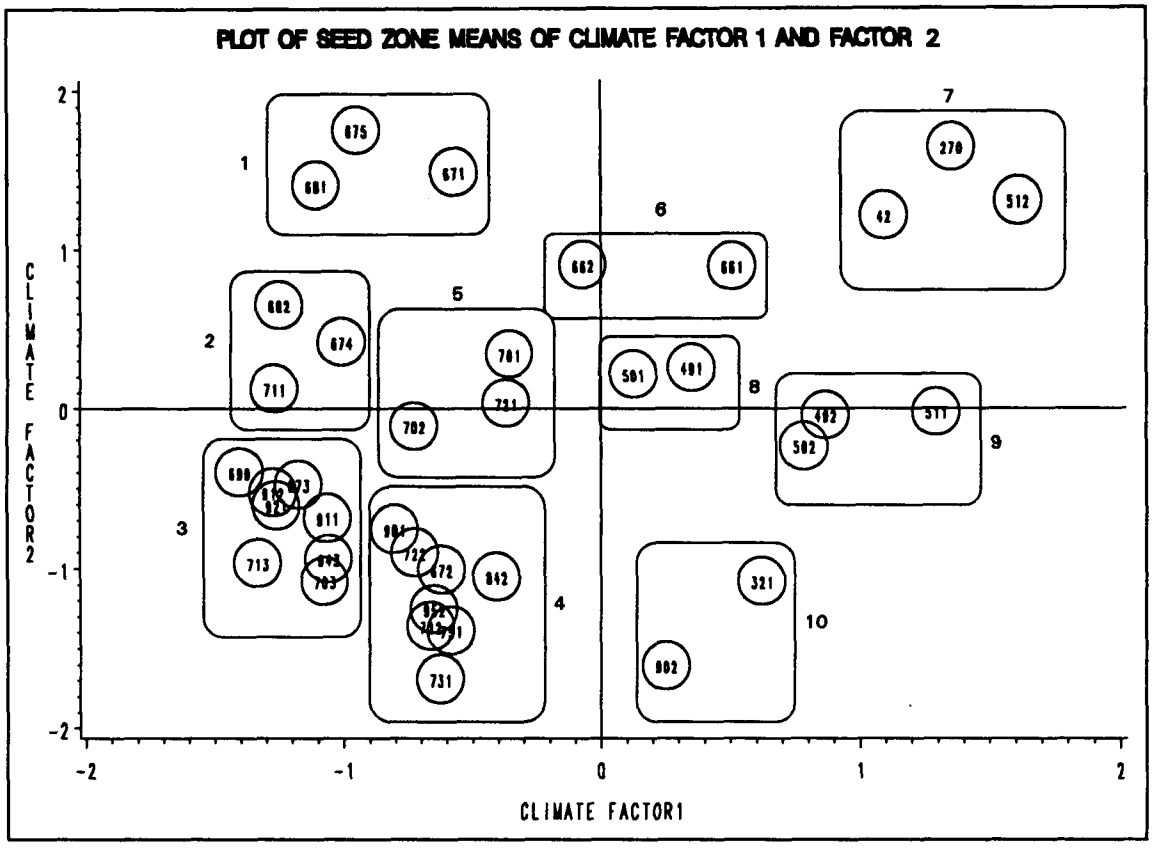


Figure 29: Plot of seed zone means of climate Factor 1 and Factor 2. Negative values on Factor 1 mean cold, positive values mean warmer temperatures. Negative values on Factor 2 mean dry, positive values mean moist habitat conditions

temperatures, a high average yearly temperature between 8.3 and 10.8 degrees Celsius, hot summers and low precipitation in July. The area towards the coast (7) is warm and wet (especially in winter), the central part (8) is warm and relatively dry (especially in summer) and the eastern part (1) is somewhat cool and wet. The area of seed zones 661 and 662 has warm and wet climate (7), showing similar conditions as found towards the coast in Southwest Oregon. High elevation locations in the Cascades are characterized by cold and wet conditions (5, 6) with exceptions in the southern part. The high desert has cold climate and is divided into a group with moderately moist (3) and a group (2) with dry habitat conditions.

If associations among multilocus genotypes and habitat conditions exist, then different genotypic structures should be found among the trees growing in different habitat conditions. In order to test for such differences and to describe eventual differences, descriptive discriminant analyses was performed on the 8 groups of habitats using genotypic scores as discriminating variables. Results are presented in *Table 27, p. 135* for the total sampling area as well as for the two regions separately.

The first variate generated by canonical discriminant analysis performed on total sampling area was highly significant ($p < 0.0001$), accounting for 47% of total variation in scores of 45 allozyme variables. Variables with structure coefficients < 0.10 were dropped from the analysis in order to improve the ratio of samples to variables. Canonical R^2 indicated that 32.8% of variation in the first canonical variate (representing 47% of total variance) could be accounted for by differences in habitat conditions. The first variate was dominated by several alleles at different loci. Highest redundancies, *i.e.*, largest proportions of variation contributing to group differences (in brackets) were apparent for *Idh1-1* (7.25), *Skd1-1* (4.95), *Gdh1-1* (3.75), *Fdp2-1* (3.30), *Pep3-1* (3.35), *Gdh1-2* (3.25), *Pgm1-1* (3.0), *Lap2-1* (3.0), *Acp1-2* (2.60), *Pgm1-2* (2.60), *Mnr1-1* (2.60), *Mnr1-2* (2.50), *Pep3-2* (2.45) and *Mnr2-1* (2.40). Thus, canonical variate 1, explaining

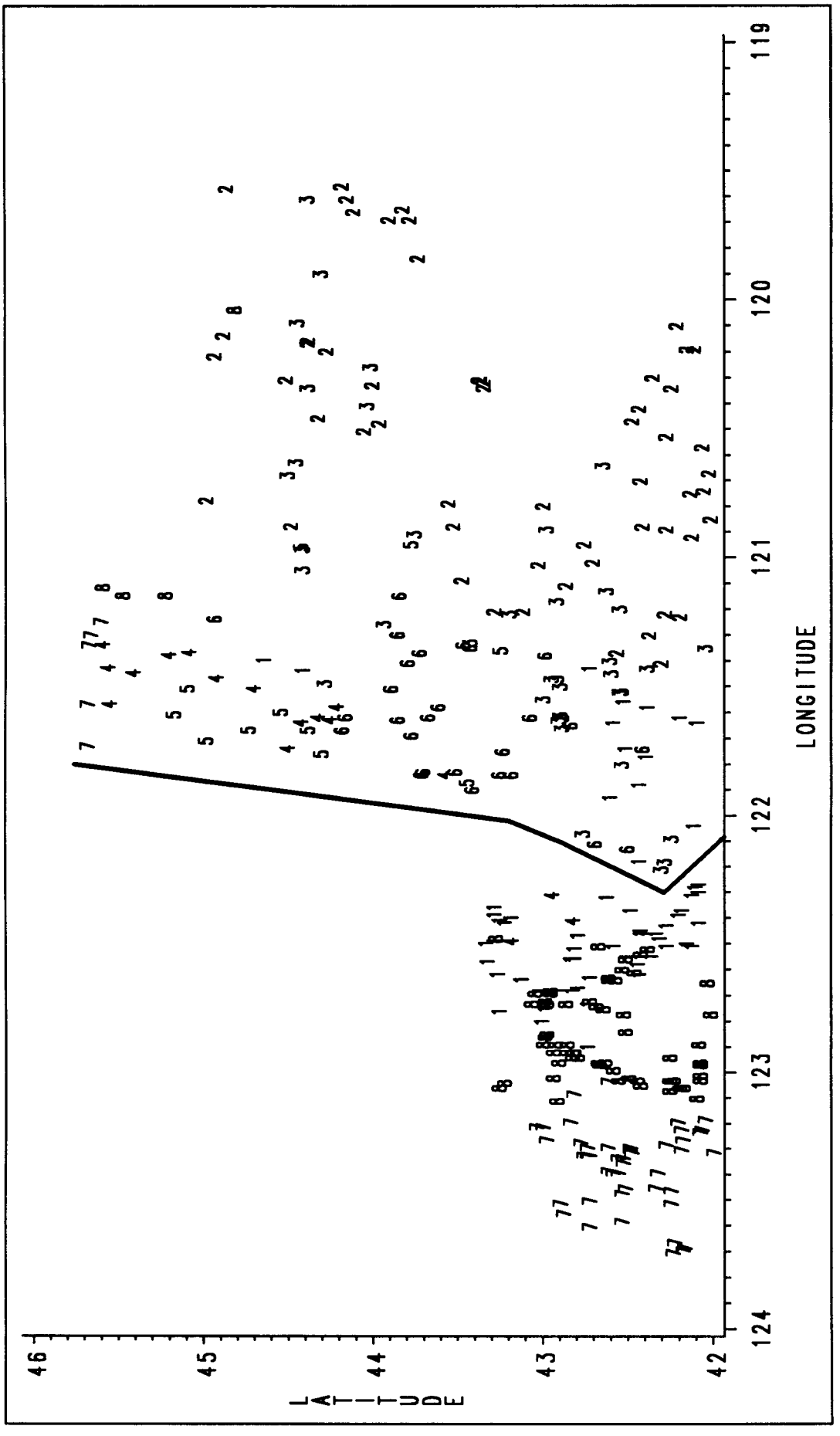


***Figure 30**: Mapped groups of similar climate: Total sampling area. 8 groups formed by complete linkage cluster analysis on climate Factor 1 and Factor 2*

Groups / Climate Variables	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
PCA Factor 1	0.128	-0.763	-1.057	-0.373	-1.346	-1.097	1.342	1.064
PCA Factor 2	-0.105	-1.177	-0.080	1.311	2.242	1.062	1.053	-0.688
APJA	166	51	148	284	340	265	378	151
APJU	16	14	22	28	43	28	10	7
APYE	1111	436	1054	1954	2425	1777	2209	917
AVJA	0.05	-3.8	-3.6	-0.8	-2.64	-3.0	3.3	2.3
AVJU	17.1	15.9	14.6	15.7	13.21	14.1	18.7	19.0
AVYE	8.3	5.8	5.1	7.0	4.5	5.0	10.8	10.4
AVMIJA	-4.4	-10.1	-9.4	-4.9	-7.5	-8.6	-0.6	-1.8
AVMIJU	8.1	4.9	4.4	7.5	4.7	4.1	9.8	9.4
AVMIYE	1.5	-2.3	-2.4	1.2	-1.5	-2.3	4.2	3.4
AVMAJA	4.6	2.4	2.2	3.0	1.9	2.3	8.4	7.2
AVMAJU	26.2	27.0	24.9	24.2	22.1	24.1	27.7	28.7
AVMAYE	14.5	13.0	12.5	13.1	11.0	12.3	16.9	16.9
Samples N	81	75	50	22	14	32	58	79

Factor 1/Factor 2: Principal components from climate data set. Factor 1 represents a temperature, Factor 2 a water bilan effect.

AP: Average precipitation [mm]
AV : Average daily temperature [Celsius]
AVMI: Average daily minimum temperature [Celsius]
AVMA: Average daily maximum temperature [Celsius]

JA: January
JU: July
YE : annual

Table 26: Climatic characteristics of the 8 groups as shown in Figure 30, p. 133: Total sampling area

the single maximum among-group variance (32.8%) was dominated by the same loci which showed highly significant heterogeneity among the two regions, which exhibited a spatial pattern and were related to temperature in the multinomial response models.

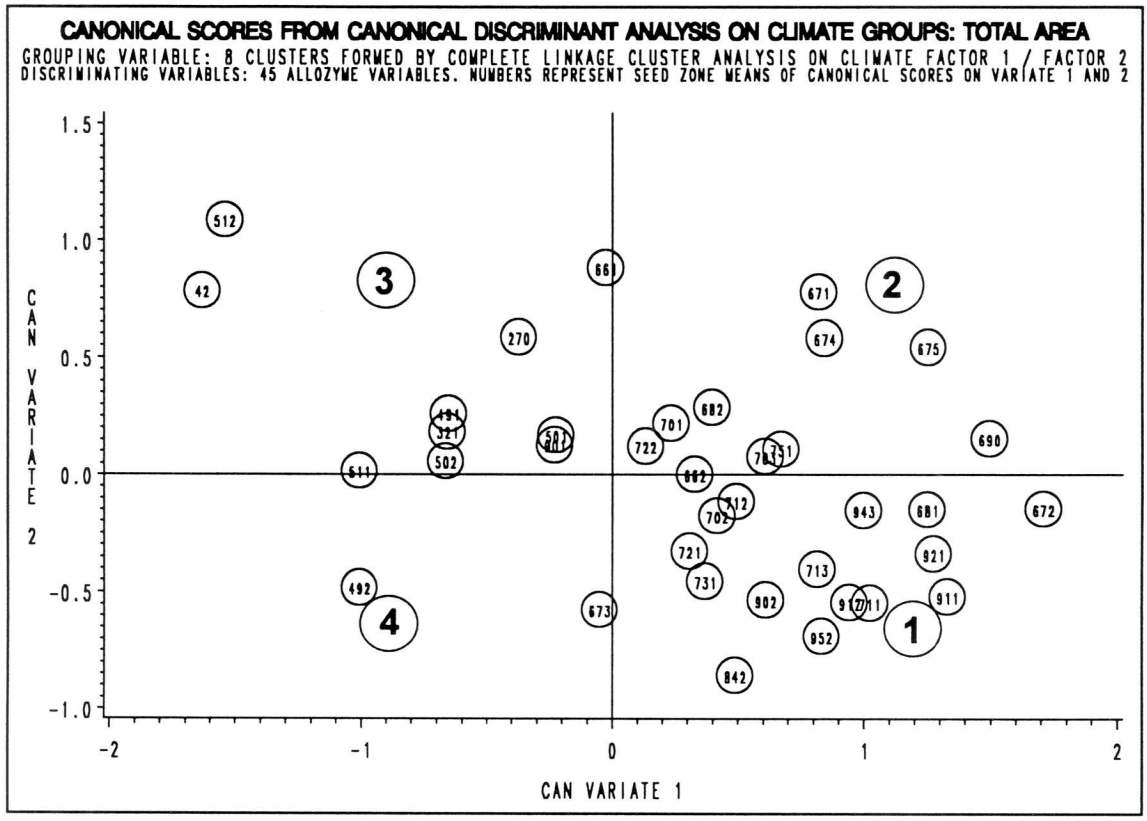
Plotting mean canonical scores of the seed zones for the first two canonical axis separated the two areas completely on axis 1 (*Figure 31, p. 136*), reflecting either a real association with differences in climate conditions of the two regions or being, at least partly, a consequence of indirect effects (such as a different evolutionary history) associated with the two regions.

Variation on axis 1 seemed to reflect associations with temperature while variation on axis 2, although not significant ($p = 0.35$), showed a trend of association with moisture. On average, positive values on axis 1 combined with positive values on axis 2 were associated with cold and relatively moist climate since they predominated under such habitat conditions (*Figure 29, p. 132, aggregates 1, 2 and 5*). Positive values on variate 1 combined with negative values on variate 2 were associated with cold climate and drier habitat conditions (*aggregates 2 and 4*). Negative values on axis 1 combined with positive values on axis 2 were related to mild and wet habitat conditions (*aggregates 6 and 7*) while zero or negative values on axis 2 were associated with drier habitats as found in aggregates 8 and 9. Exceptions of this trend were apparent for seed zones 321, 703, 711, 722, 751, 681 and 690.

Derived linear discriminant functions used to predict group membership of individuals based on their allozyme scores (45 variables), produced a reasonably high classification accuracy of 44.7% (*Table 28, p. 137*). Individuals could be assigned to their initial habitat groups with a precision that was on average 32.2% higher than what might be expected by chance alone (prior probabilities). Chance-corrected classification estimates which account for unequal group size such as Tau statistics (KLECKA, 1980) or Kappa statistics (COHEN, 1960) showed an improvement of 37% or 35% over chance. Classification rates differed between the groups, however. High precision was reached for the habitats number 4, 5 and 6, which are characterized by cold and moist conditions and which predominate in the high elevations of the Cascade Range. Moderately high precision was reached for individuals growing in habitat number 7 which has mild and moist conditions; these conditions predominate towards the coast in Southwest Oregon. Trees growing in habitats number 1, 2, 3 and 8 seemed genetically less distinct.

Locus	TOTAL AREA				SOUTHWEST				CENTRAL OREGON			
	CAN 1		CAN 2		CAN 1		CAN 2		CAN 1		CAN 2	
Allele	struc.	% red.	struc.	% red.	struc.	% red.	struc.	% red.	struc.	% red.	struc.	% red.
MNR1-1	-0.2808	2.59										
MNR1-2	0.2769	2.52	-0.1803	0.50								
MNR1-3					0.1281	0.67						
MNR2-1	0.2691	2.38	-0.1764	0.48					0.2424	1.16		
MNR2-2	-0.2431	1.94	0.1320	0.27					-0.1581	0.49	-0.1373	0.37
LAP2-1	0.3028	3.01			0.1006	0.41						
LAP2-2	-0.1656	0.91										
LAP2-3	-0.1168	0.45	0.1817	0.50	-0.1999	1.62						
LAP3-2					0.1365	0.76						
PEP1-1	0.1797	1.06	-0.2253	0.78								
PEP1-2	-0.1257	0.53	0.1911	0.56								
PEP2-1	0.1512	0.75										
PEP2-2	-0.1175	0.45			0.1527	0.95						
PEP3-1	-0.3182	3.33					-0.1706	0.50	-0.2938	1.70	0.2057	0.83
PEP3-2	0.2717	2.43	-0.2147	0.71					0.1420	0.40	-0.2061	0.83
PEP4-1	0.1150	0.43	-0.1319	0.27					0.2735	1.47	0.2201	0.96
MPI1-1									0.3064	1.85	0.3923	3.03
MPI2-1	-0.1512	0.75	0.1747	0.48								
GOT1-1									0.1557	0.48	0.1039	0.21
GOT1-2	0.1485	0.72	0.1727	0.48	-0.3511	5.00	-0.2933	1.37				
GOT1-3									-0.1984	0.78	-0.2134	0.90
GOT2-1					0.2550	2.64	-0.2689	1.24				
GOT2-2					-0.2829	3.24	0.2957	1.50				
GOT3-2	-0.1549	0.79	-0.1713	0.48								
GOT3-3					0.1697	1.17			-0.1490	0.44	0.1240	0.30
G6P2-1	0.2030	1.35	-0.3351	1.72	-0.1197	0.58			-0.1485	0.44		
G6P2-2	-0.1229	0.51	0.1161	0.21			0.1897	0.62				
G6P2-3	-0.1689	0.94	0.3593	1.98					0.1731	0.60		
ACP1-1	0.2003	1.35										
ACP1-2	-0.2823	2.62	-0.1247	0.24								
ACP1-4	0.1410	0.65	0.3775	2.18								
GDH1-1	0.3373	3.74										
GDH1-2	-0.3137	3.23										
IDH1-1	-0.4693	7.25			-0.1240	0.62	0.3549	2.15	-0.2921	1.68		
PGM1-1	-0.3029	3.01			0.1481	0.89			-0.3466	2.37		
PGM1-2	0.2799	2.57	-0.1077	0.18	-0.3618	5.31	-0.2099	0.75	0.2303	1.05		
ACO1-2	0.1112	0.41							0.2832	1.58		
ACO1-3	0.1041	0.36	0.2147	0.71								
SKD1-1	-0.1897	1.18			0.1558	0.98			-0.2053	0.83	0.1954	0.75
SKD1-2	0.1922	1.21			-0.1906	1.47			0.2725	1.46	-0.1081	0.23
SKD2-1	-0.3883	4.95			0.1067	0.46	0.2370	0.96	-0.2276	1.02	-0.1781	0.63
FDP2-1	-0.3177	3.32										
UGP1-1	-0.1552	0.79			0.1431	0.83			-0.3970	3.11	0.3489	2.40
UGP1-2	0.1378	0.62			-0.1246	0.63			0.2707	1.44	-0.1582	0.49
UGP2-1	0.1021	0.34	0.1694	0.44	-0.1408	0.80						
UGP3-1					0.2325	2.20	0.1027	0.18			-0.1587	
FUM2-1					-0.1367	0.76						
ADH2-1	-0.2205	1.61	0.1258	0.24	-0.2104	1.80	0.1133	0.22	0.1311	0.34		
ADH2-2	0.2267	1.69	-0.1185	0.21	0.2211	2.00	-0.1256	0.27	-0.1311	0.34		
PGI2-1	0.1625	0.87										
PGI2-2					0.1535	0.96						
PGI2-3	-0.1355	0.61	0.1389	0.30	-0.1248	0.63	0.3454	2.04				
MDH1-1					0.3222	4.20			-0.1747	0.60	-0.2084	0.86
MDH1-2					-0.1894	1.45	0.1967	0.66				
MDH3-1	0.1533	0.77	-0.2575	1.01								
MDH3-2	-0.1427	0.67	0.2495	0.95	0.1942	1.53	-0.2038	0.71				
MDH3-3	-0.1084	0.39			-0.2028	1.67			0.1526	0.46		
MDH4-1	0.2075	1.41	0.2258	0.78	-0.2647	2.84			0.2848	1.60		
Eigenvalue	0.6674		0.1808		1.0482		0.4749		0.4461		0.2488	
% variance	47.4		12.85		37.56		17.20		31.00		17.10	
Can. corr	0.5732		0.3912		0.6367		0.4136		0.4440		0.4440	
r square	0.3286		0.1530		0.4054		0.1711		0.1971		0.1971	
Signif.	0.0001		0.3525		0.0005		0.3773		0.0533		0.6137	

Table 27: *Canonical discriminant analysis of groups of similar climate with allozymes as discriminating variables: Total sampling area, Southwest Oregon and Central Oregon. Only structure coefficients greater than 0.10 are shown*



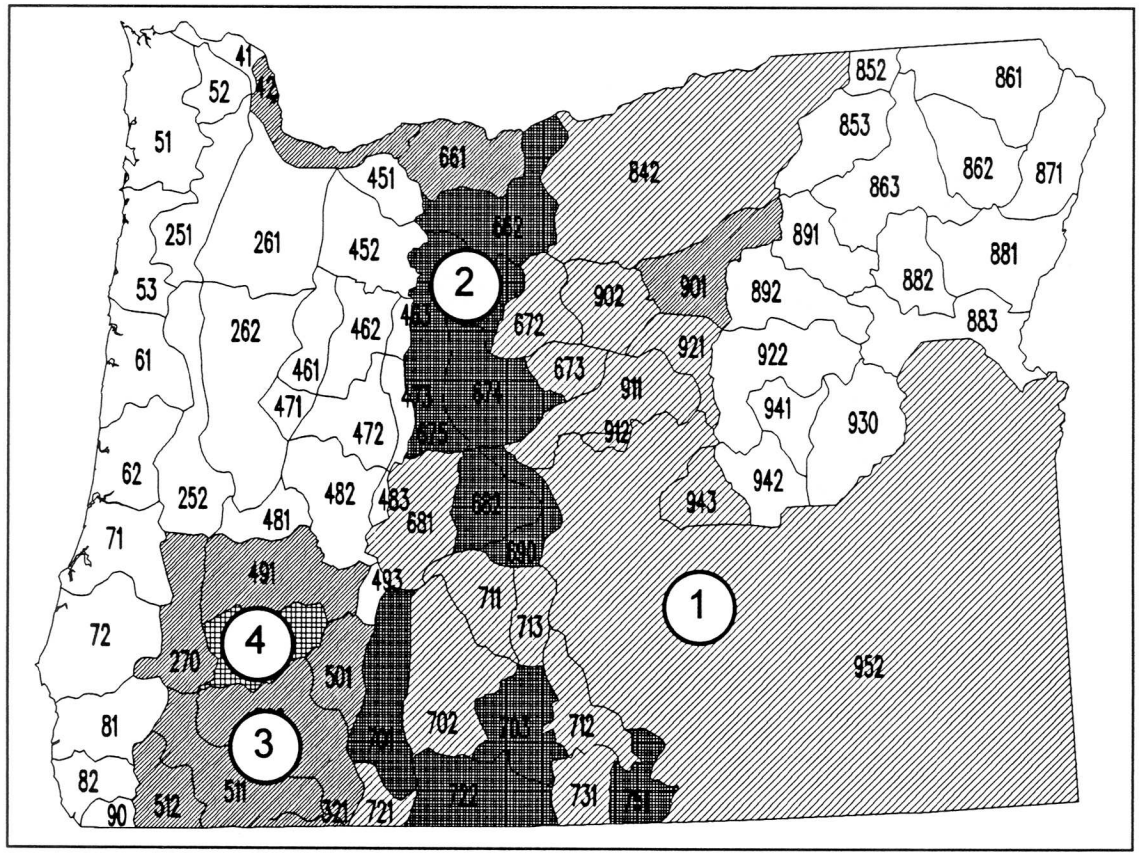


Figure 31: *Plot of canonical scores from canonical discriminant analysis on climate groups and mapped illustration of the four quadrants: Total sampling area*

From actual group	Percent trees classified into group number:								
Number	1	2	3	4	5	6	7	8	N
1	32.10	9.88	8.64	4.94	2.47	9.88	16.05	16.05	81
2	8.00	36.00	12.00	2.67	12.00	17.33	8.00	4.00	75
3	2.00	20.00	36.00	12.00	10.00	10.00	8.00	2.00	50
4	13.64	9.09	9.09	54.55	4.55	9.09	0.00	0.00	22
5	0.00	7.14	7.14	0.00	64.29	21.43	0.00	0.00	14
6	9.38	6.25	6.25	9.38	3.13	53.13	9.38	3.13	32
7	20.69	5.17	5.17	8.62	1.72	1.72	44.83	12.07	58
8	21.52	10.13	7.59	3.80	1.27	2.53	16.46	36.71	79
Prior probabilities [%]	12.50	12.50	12.50	12.50	12.50	12.50	12.50	12.50	411
Error rate [%]	67.90	64.00	64.00	45.45	35.71	46.88	55.17	63.29	55.30
Higher than prior prob [%]	19.60	23.50	23.50	42.05	51.79	40.63	32.33	24.24	32.20
Tau-statistic (Klecka, 1980)									0.37
Kappa-statistic (Cohen, 1960)									0.35
Crossvalidation:									
Error rate %	76.54	84.00	82.00	95.45	78.57	68.75	74.14	78.48	79.74
Higher than prior prob [%]	10.90	3.50	5.50	-7.95	8.93	18.70	13.36	9.32	7.76
Tau-statistic									0.10
Kappa-statistic									0.07

***Table 28**: Canonical discriminant analysis of climate groups: Total sampling area. Classification results based on linear discriminant function and cross validation*

Classification results of 32% to 37% higher than chance paralleled the estimated association (R^2 = *0.328, Table 27, p. 135)* among genotypes and habitat conditions. Results from cross-validation, however, showed a substantially reduced classification rate of only 7% to 10% over chance (Tau- and Kappa-statistic, *Table 28*), suggesting violations of assumptions for discriminant analysis and indicating that the derived canonical functions are not robust estimates. Consequently, estimates of among-group to within-group variance may be distorted and biased. According to classification results obtained by cross-validation, trees growing in habitats numbered 6, 7, 8, 5 and 1 seemed genetically most distinct while the other 3 habitats showed low classification results.

Mapping the 10 subsets of habitat types (*Table 29*), evaluated by cluster analysis on both climate factors within Southwest Oregon only, revealed a relatively small scale, irregular distribu-

Groups / Climate Variables	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
PCA Factor 1	1.461	0.363	0.000	0.781	0.828	1.345	1.500	1.343	-0.463	1.550
PCA Factor 2	1.606	-0.200	0.302	-0.316	-0.850	-0.573	0.276	1.178	1.061	-1.196
APJA	455	159	192	179	109	184	304	403	269	121
APJU	10	16	23	10	8	5	7	9	34	3
APYE	2619	1061	1290	1126	719	1054	1702	2319	1706	664
AVJA	4.1	0.9	0.2	1.7	1.8	3.1	3.5	3.4	-0.8	3.5
AVJU	18.2	17.5	16.7	18.2	18.6	19.2	19.2	18.4	15.8	20.0
AVYE	11.0	8.9	8.1	9.8	9.9	11.0	11.0	10.8	6.9	11.4
AVMIJA	0.0	-3.2	-3.9	-2.4	-2.2	-0.9	-0.5	-0.6	-4.9	-0.7
AVMIJU	9.4	8.5	8.3	8.6	9.3	9.5	9.6	9.2	7.7	10.0
AVMIYE	4.3	2.3	1.8	2.7	3.0	3.8	4.0	3.9	1.0	4.2
AVMAJA	9.4	5.3	4.3	6.6	6.6	8.4	9.4	8.9	3.0	8.7
AVMAJU	27.4	26.6	25.3	27.9	28.0	29.3	29.2	27.7	23.9	30.4
AVMAYE	17.2	15.0	14.0	16.1	16.2	17.7	17.7	16.9	12.7	18.5
Samples N	15	32	22	24	27	24	10	24	4	7

Factor 1/Factor 2: Principal components from climate data set. Factor 1 represents a temperature, Factor 2 a water bilan effect.

AP:	Average precipitation [mm]	JA:	January
AV:	Average daily temperature [Celsius]	JU:	July
AVMI:	Average daily minimum temperature [Celsius]	YE:	annual
AVMA:	Average daily maximum temperature [Celsisus]		

***Table 29**: Climatic characteristics of the 10 groups as shown in Figure 32, p. 138: Southwest Oregon*

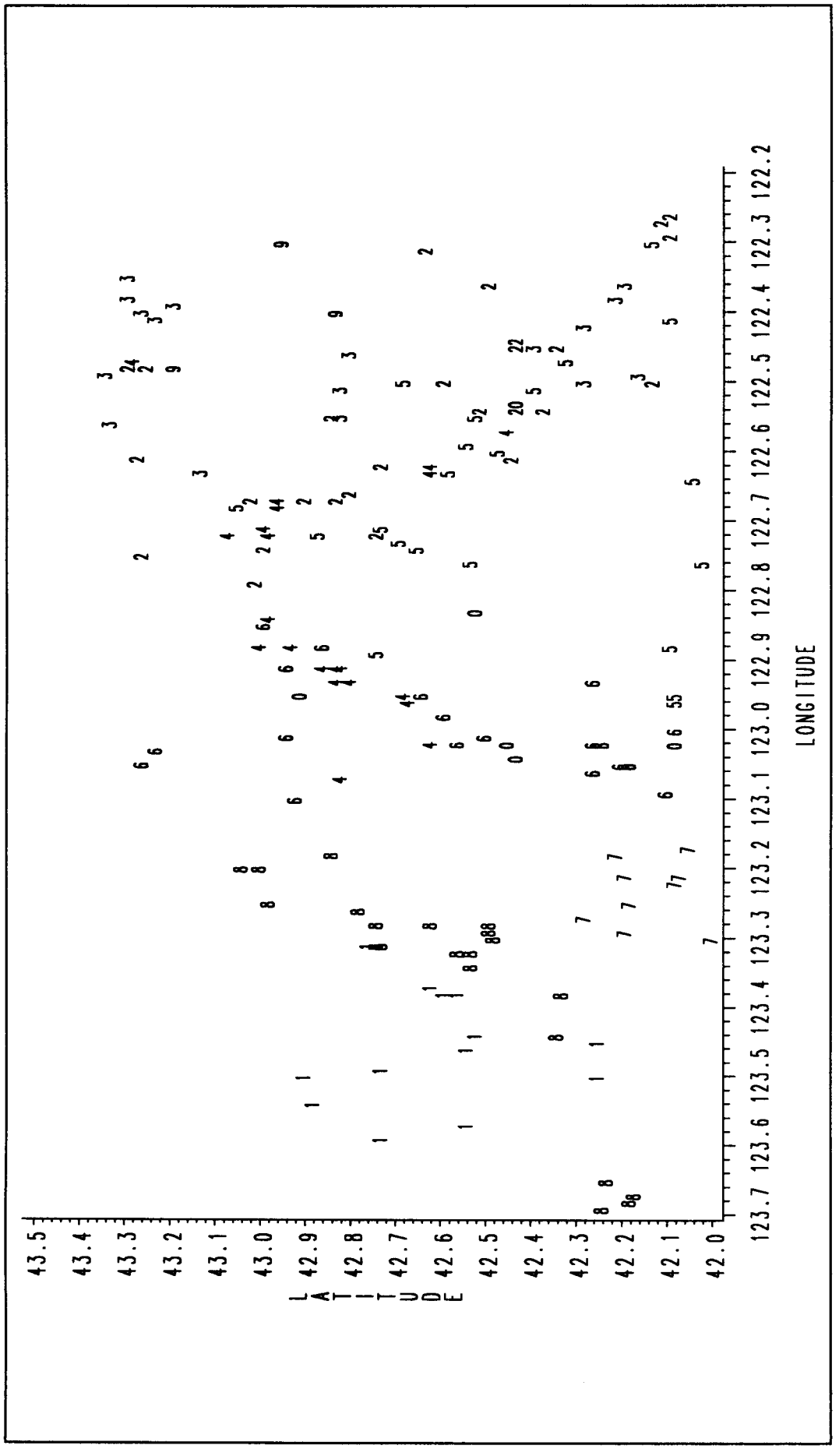


Figure 32: *Mapped groups of similar climate: Southwest Oregon. 10 groups formed by complete linkage cluster analysis on climate Factor 1 and Factor 2*

tion of habitats (*Figure 32, p. 138*). Although the distribution of habitats appeared to be patchy at the area-wide scale (*Figure 30, p. 133*), patterns of climate were more complex when looked at from within the area.

At the regional scale, a patchy distribution of habitats is still observable, more in the western than in the eastern part of the region. The western part towards the coast is characterized by mild and wet climate, with small local differences in water conditions (*groups 1, 7 and 8*). A gradient of decreasing moisture availability to the south and especially to the center of the area is apparent, *groups 7, 6, 5 and 10* exhibiting low to negative values on climate *Factor 2*. The northeastern part is mostly cool and moderately moist while the southeastern part is moderately warm and relatively dry.

Discriminant analysis, performed on the samples of Southwest Oregon only, produced a highly significant canonical variate ($p = 0.0005$), accounting for 37.5% of total variation in scores of the 32 allozyme variables which exhibited structure coefficients greater than 0.10 (*Table 27, p. 135*). Variables were dropped in order to improve the ratio of samples to variables. Canonical R^2 indicated that 40.5% of variation in the first variate could be accounted for by differences among habitat groups. This variate was dominated by several alleles at different loci. Highest redundancies (in brackets) were apparent for *Pgm1-2* (5.31), *Got1-2* (5.0), *Mdh1-1* (4.2), *Got2-2* (3.25) and *Mdh4-1* (2.85).

Plotting mean scores on the first two discriminating axis revealed no straightforward, easy-to-explain patterns (*Figure 33, p. 140*). Classification based on linear discriminant function produced an average rate of 51.6% correctly classified individuals, an accuracy which was 36% to 42% above chance (*Table 30, p. 140*). Cross validation indicated, however, that canonical functions were far from being robust. Classification accuracy was reduced to 3% to 5% above chance.

Environmental conditions within Central Oregon, evaluated by cluster analysis within Central Oregon only (*Table 31, p. 141*), may be characterized as a small scale, irregular distribution of habitats (*Figure 34, p. 142*). Conditions may change abruptly over short distances due to the highly variable topography. A somewhat local patchy structure was apparent for habitat *number 7* which occurred only within seed zone 661, for habitat *number 1* which was restricted to the southwestern part of the zone and for habitats *number 9 and 10* which were concentrated in a zone in the center of the east side of the Cascades. All other habitats showed an irregular distribution, occurring at different locations all over the area.

Discriminant analysis, performed on the 10 subsets of habitats within Central Oregon, produced a significant canonical variate ($p = 0.053$) which accounted for 31% of variance of the 25 allozyme variables (*Table 27, p. 135*). Variables which exhibited a structure coefficient smaller than 0.10 were dropped from the analysis in order to improve the ratio of samples to variables. Canonical R^2 revealed that 20% of variation in the first variate was due to differences among habitat types. The variate was dominated by several alleles at many loci. Highest redundancies (in brackets) were found for *Ugp1-1* (3.1), *Pgm1-1* (2.4), *Mpi1-1* (1.85), *Pep3-1* (1.70), *Idh1-1* (1.68) and *Mdh4-1* (1.60).

Plotting mean scores on the first two canonical variates revealed a complex pattern which showed only weak trends of associations with habitats (*Figure 35, p. 143*). Positive values on variate 1 combined with positive values on variate 2 were mostly associated with cold to moderately cold climate under variable moisture conditions (the exception being zone 701). Positive values on axis 1 combined with negative values on axis 2 were associated with moderate to mild climate and dry habitat conditions (the exception being zone 675). No obvious associations were apparent for zones with negative values on axis 1 since they were found under all temperature and all moisture conditions.

Classification based on linear discriminant functions produced an average rate of 48.3% correctly classified individuals into their initial subsets (*Table 32, p. 144*). The classification accuracy was 31% to 41% higher than what might be expected by chance alone. Classification accuracy was thus higher than the estimated association of $R^2 = 0.20$. However, estimated ca-

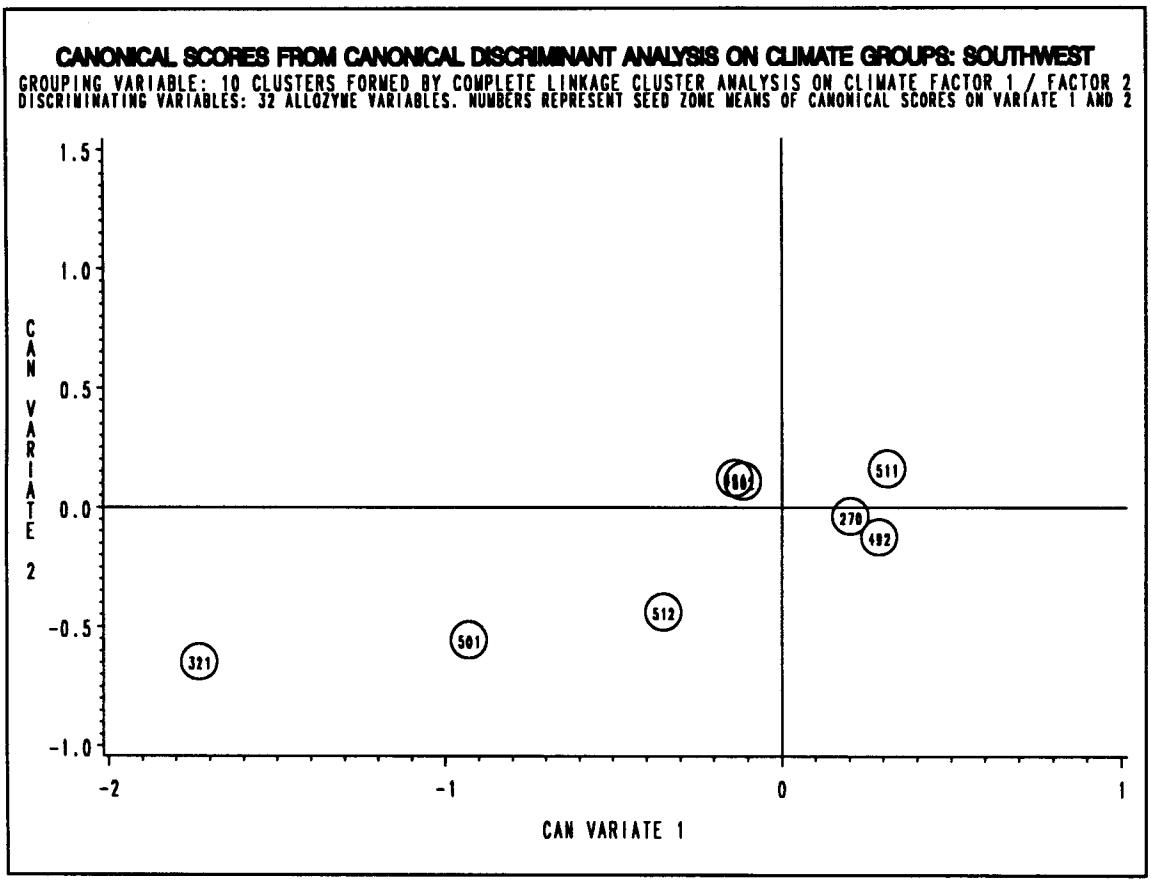


Figure 33: *Plot of canonical scores from canonical discriminant analysis on climate groups: Southwest Oregon*

From actual group	Percent trees classified into group number:										
Number	1	2	3	4	5	6	7	8	9	10	N
1	40.00	13.33	6.67	13.33	0.00	13.33	6.67	6.67	0.00	0.00	15
2	6.25	25.00	6.25	18.75	6.25	9.38	18.75	9.38	0.00	0.00	32
3	4.55	4.55	40.91	13.64	9.09	13.64	4.55	0.00	0.00	9.09	22
4	4.17	8.33	8.33	45.83	16.67	0.00	4.17	8.33	0.00	4.17	24
5	7.41	0.00	11.11	3.70	48.15	3.70	3.70	7.41	7.41	7.41	27
6	8.33	8.33	8.33	12.50	0.00	50.00	8.33	4.17	0.00	0.00	24
7	10.00	10.00	10.00	0.00	0.00	20.00	40.00	10.00	0.00	0.00	10
8	4.17	12.50	4.17	8.33	16.67	8.33	4.17	33.33	4.17	4.17	24
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	75.00	25.00	4
10	0.00	0.00	0.00	0.00	14.29	0.00	0.00	0.00	0.00	85.71	7
Prior prob. [%]	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	189
Error rate [%]	60.00	75.00	59.09	54.17	51.85	50.00	60.00	66.67	25.00	14.29	51.61
Higher than p. prob [%]	30.00	15.00	30.91	35.83	38.42	40.00	30.00	23.33	65.00	75.71	38.39
Tau-statistic (Klecka, 1980)											0.36
Kappa-statistic (Cohen, 1960)											0.42
Crossvalidation:											
Error rate %	100.00	93.75	68.36	79.17	88.89	75.00	100.00	83.33	100.00	85.71	89.22
Higher than p. prob [%]	10.00	-3.75	21.64	10.83	1.11	15.00	-10.00	6.67	-10.00	4.29	0.78
Tau-statistic											0.03
Kappa-statistic											0.05

Table 30: *Canonical discriminant analysis of climate groups: Southwest Oregon. Classification results based on linear discriminant function and cross validation*

Groups / Climate Variables	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
PCA Factor 1	-0.318	-0.763	-0.551	-1.458	-0.598	-1.153	0.983	0.845	-1.217	-2.350
PCA Factor 2	-0.451	-1.177	0.063	-0.236	1.230	2.131	0.824	-0.912	0.594	1.886
APJA	153	52	187	107	284	349	272	127	215	247
APJU	9	15	16	27	26	38	18	8	29	60
APYE	980	436	1230	874	1915	2392	1893	808	1485	2107
AVJA	-2.4	-3.9	-2.3	-4.6	-1.5	-2.2	1.2	0.0	-3.7	-5.6
AVJU	16.6	15.9	15.7	13.7	15.1	13.6	19.7	19.7	14.1	11.4
AVYE	6.7	5.8	6.3	4.1	6.3	4.9	10.6	9.9	4.8	2.4
AVMIJA	-8.1	-10.1	-7.8	-10.4	-6.1	-7.0	-1.6	-3.7	-10.0	-11.9
AVMIJU	6.3	4.8	5.7	3.5	6.4	4.9	13.5	11.8	3.3	1.3
AVMIYE	-1.0	-2.3	-1.2	-3.1	0.0	-1.2	6.1	4.1	-3.3	-5.0
AMAJA	3.1	2.4	2.9	1.5	2.7	2.2	3.7	3.4	2.4	1.1
AVMAJU	26.9	27.0	25.7	24.1	24.1	22.6	26.3	27.9	25.0	21.6
AVMAYE	14.4	14.0	13.7	11.5	12.8	11.4	15.5	16.0	12.8	9.9
Samples N	12	75	34	24	32	18	7	6	11	3

Factor 1/Factor 2: Principal components from climate data set. Factor 1 represents a temperature, Factor 2 water bilan effect.

AP:	Average precipitation [mm]	JA:	January
AV:	Average daily temperature [Celsius]	JU:	July
AVMI:	Average daily minimum temperature [Celsius]	YE:	annual
AVMA:	Average daily maximum temperature [Celsisus]		

Table 31: Climatic characteristics of the 10 groups as shown in Figure 34, p. 142: Central Oregon

nonical functions do not seem to be very robust. A substantially reduced classification accuracy of 2.2% to 10% above chance is a strong evidence for violations of assumptions or an insufficient sample size.

6.2 Associations of multilocus genotypes with climate conditions, inferred by canonical trend surface analysis

Multivariate patterns of allozyme variation associated with climate conditions may be described using canonical trend surface analysis (LEE, 1969; WARTENBERG, 1985; WESTFALL and CONKLE, 1992). Canonical correlation analysis is the multivariate equivalent of multiple regression, the difference being that more than one dependent variable may be included in the analysis. Specifically, canonical correlation describes the association between two sets of variables by simultaneously quantifying and comparing their variation patterns. Multivariate trend-surfaces are derived from the two variance-covariance matrices (one for the dependent, one for the independent data set), such that the correlation is maximized between the linear composite values for the trend surfaces and the linear composite values for the other set of variables. This method is called canonical trend surface analysis (WARTENBERG, 1985). Trend surface analysis removes local variations and anomalies, leaving just the overall trends. Hence, the method should prove adequate to quantify and describe the main associations and to illustrate the important multivariate patterns of variation in relation to effects specified in a model of habitat variables. Canonical correlation analysis was described in detail in *section 3.6.4.4.*

Results of canonical trend surface analysis, relating multi-locus genotypic frequencies to climate conditions at source location for the total sampling area, are furnished in *Table 33, p. 145.*

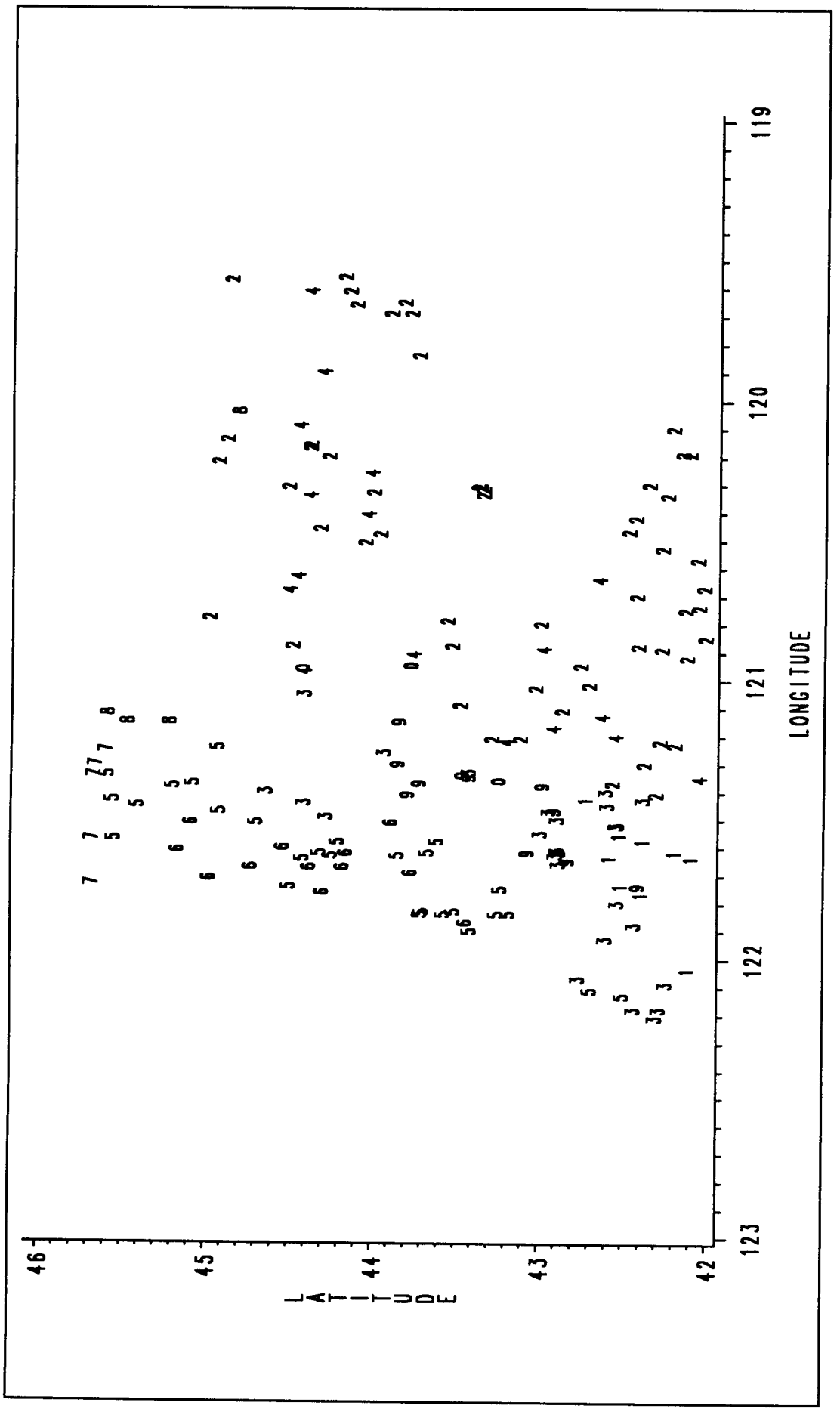
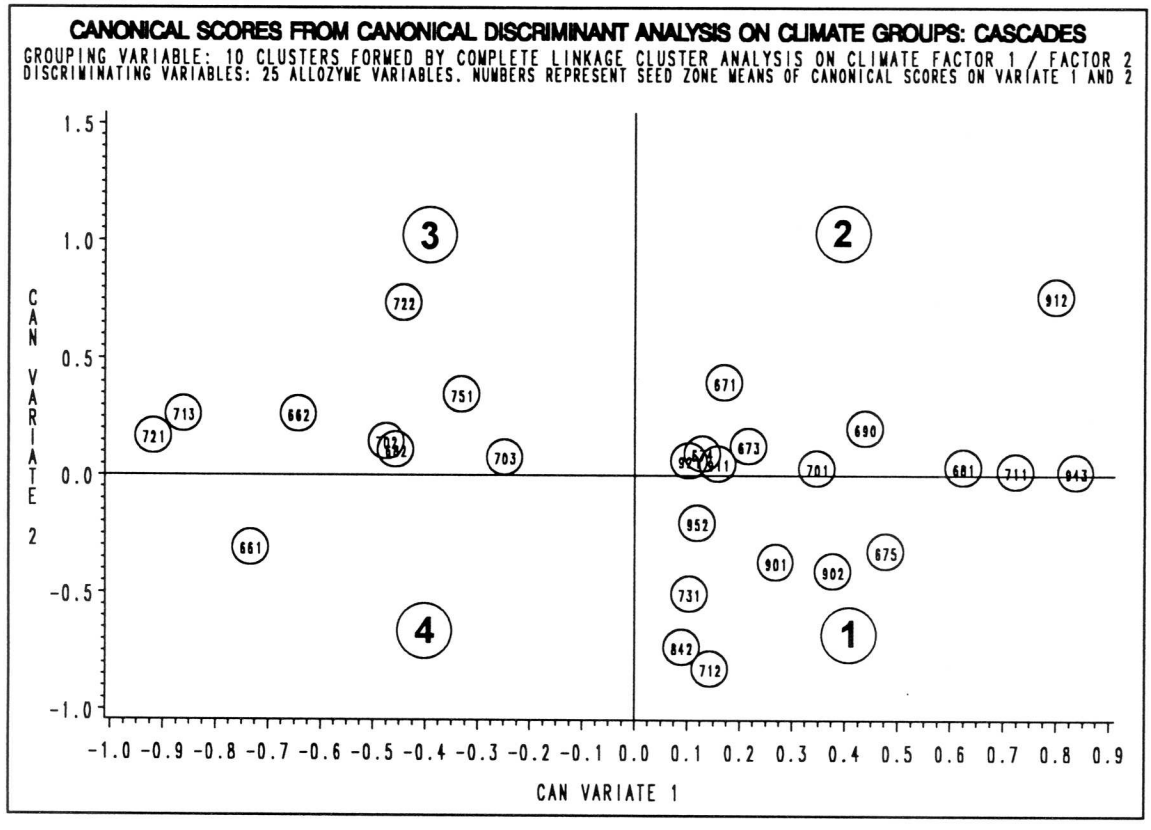


Figure 34: *Mapped groups of similar climate: Central Oregon. 10 groups formed by complete linkage cluster analysis of climate Factor 1 and Factor 2*



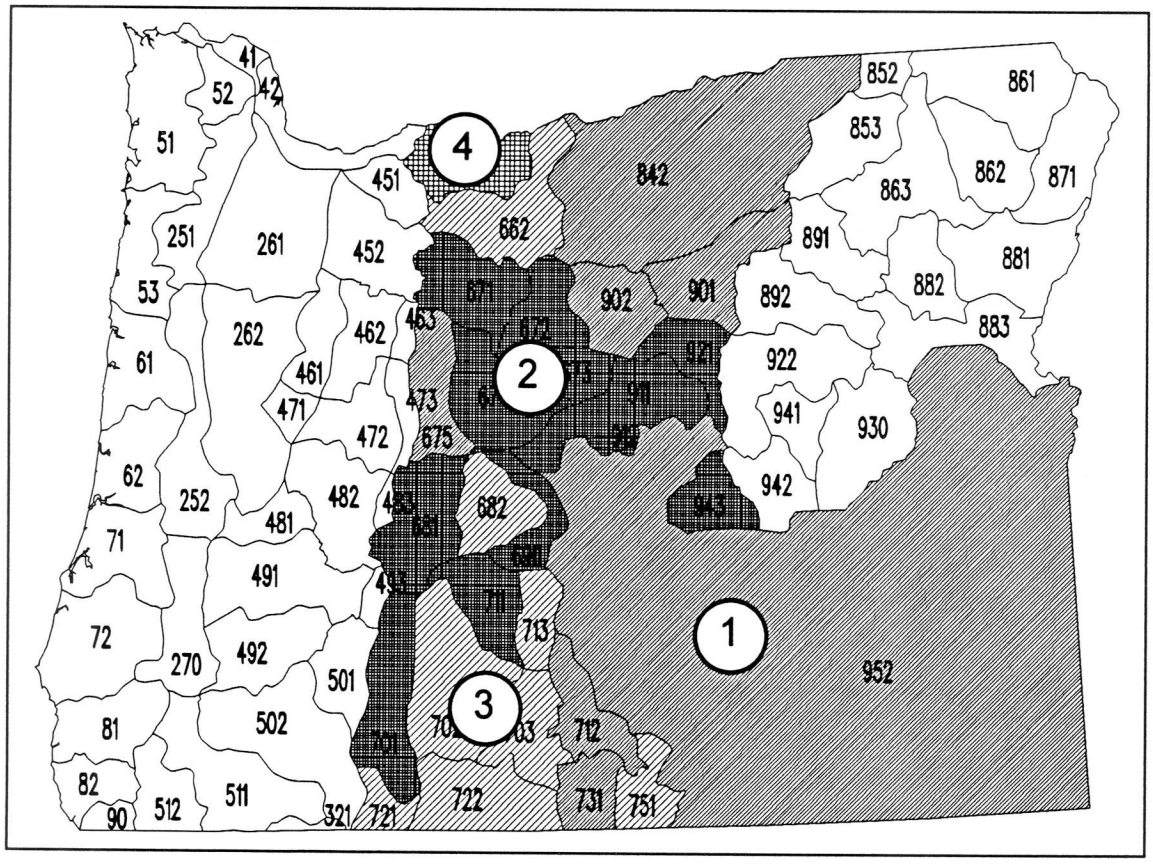


Figure 35: *Plot of canonical scores from discriminant analysis of climate groups and mapped illustration of the four quadrants: Central Oregon*

From actual group	Percent trees classified into group number:										
Number	1	2	3	4	5	6	7	8	9	10	N
1	50.00	0.00	0.00	0.00	0.00	8.33	8.33	25.00	8.33	0.00	12
2	5.33	30.67	8.00	5.33	8.00	13.33	4.00	6.67	14.67	4.00	75
3	5.88	17.65	20.59	11.76	5.88	8.82	5.88	8.82	11.76	2.94	34
4	4.17	12.50	4.17	45.83	4.17	4.17	0.00	12.50	12.50	0.00	24
5	6.25	6.25	6.25	15.63	28.13	18.75	3.13	3.13	12.50	0.00	32
6	5.56	11.11	0.00	5.56	5.56	61.11	5.56	5.56	0.00	0.00	18
7	0.00	0.00	0.00	0.00	0.00	0.00	66.67	16.67	16.67	0.00	6
8	33.33	0.00	16.67	0.00	0.00	0.00	0.00	50.00	0.00	0.00	6
9	0.00	9.09	18.18	0.00	0.00	9.09	0.00	0.00	63.64	0.00	11
10	0.00	0.00	0.00	0.00	0.00	33.33	0.00	0.00	0.00	66.67	3
Prior prob. [%]	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	189
Error rate [%]	50.00	69.33	79.41	54.17	71.88	38.89	33.33	50.00	36.36	33.33	51.67
Higher than p. prob [%]	40.00	20.67	10.59	35.83	18.12	51.11	56.67	40.00	53.64	56.67	38.33
Tau-statistic (Klecka, 1980)											0.31
Kappa-statistic (Cohen, 1960)											0.41
Crossvalidation:											
Error rate %	91.67	78.67	100.00	87.50	81.25	88.89	50.00	100.00	100.00	100.00	87.80
Higher than p. prob [%]	-1.67	11.33	-10.00	2.50	8.75	1.11	40.00	-10.00	-10.00	-10.00	2.20
Tau-statistic											0.05
Kappa-statistic											0.10

Table 32: *Canonical discriminant analysis of climate groups: Central Oregon. Classification results based on linear discriminant function and cross validation*

A second-order polynomial of climate variables *Factor 1 and Factor 2* was used as independent model (*Table 33, p. 145*). As the dependent data set, 41 allozyme variables were retained. Alleles with structure-correlations smaller than 0.05 with the trend surface equation (canonical variate) were dropped from the analysis in order to reduce the number of variables in the model (WESTFALL and CONKLE, 1992).

Canonical analysis of allozyme scores with the climate model produced one highly significant ($p < 0.0001$) canonical variate which accounted for 65% of the patterned variation. The Hotelling-Lawley trace, representing the multivariate equivalent of the ratio of variance due to the model to the error variance, adjusted for covariances among the variables, indicated that the model explained 51.5% of total variance of the allozyme data set. However, since the first and only significant variate accounted for 65% of total trace only, 33.5% of total variance in the allozyme data set could be significantly explained by the specified model of climate variables.

The first canonical variate of the trend surface was associated with many alleles at several loci. Of the 41 alleles, 22 had a correlation (structure coefficient, *Table 33, p. 145*) higher than 0.20 with the first canonical variate *i.e. Mnr1-1, Mnr1-2, Mnr2-1, Mnr2-2, Lap2-1, Pep3-1, Pep3-2, G6p2-1, G6p2-3, Acp1-1, Acp1-2, Gdh1-1, Gdh1-2, Idh1-1, Pgm1-1, Pgm1-2, Skd2-1, Fdp2-1, Adh2-1, Adh2-2, Mdh3-1* and *Mdh3-2*. Especially high structure coefficients of 0.30 and higher were apparent for 8 alleles *i.e.* for *Idh1-1* (0.497), *Skd2-1* (0.427), *Gdh1-1* (0.346), *Pgm1-1* (0.339), *Pep3-1* (0.337), *Mnr2-2* (0.333), *Gdh1-2* (0.323) and *Lap2-1* (0.310).

High positive scores on variate 1 thus indicate high combined frequencies of alleles at *Mnr1-1, Mnr2-2, Pep3-1, G6p2-3, Acp1-2, Gdh1-2, Idh1-1, Pgm1-1, Skd2-1, Fdp2-1, Adh2-1* and *Mdh3-2* while low scores stand for high combined frequencies of *Mnr1-2, Mnr2-1, Lap2-1, Pep3-2, G6p2-1, Acp1-1, Gdh1-1, Pgm1-2, Adh2-2* and *Mdh3-1*.

The linear composite in the climate data set was dominated by climate *Factor 1* with a structure coefficient of 0.984; the first canonical variate in the independent data domain thus represented 96.6% of the variance of climate *Factor 1*. Climate *Factor 2* was only very weakly related with the first variate (1.2% of variance) while the quadratic term of *Factor 2* showed a higher association with the linear composite (7.3%), indicating that the relationship with moisture conditions of the site is weak and mostly non-linear.

Since climate *Factor 1* and scores on variate 1 are positively and strongly correlated, the relat-

ALLOZYME DATA SET			
Locus	CANONICAL VARIATE 1		
Allele	Struc. coeff	Crossloadings	Redundancy %
MNR1-1	0.2737	0.1749	3.06
MNR1-2	-0.2530	-0.1617	2.61
MNR2-1	-0.2958	-0.1890	3.57
MNR2-2	0.3293	0.2104	4.43
LAP2-1	-0.3099	-0.1980	3.92
LAP2-2	0.1959	0.1251	1.57
PEP1-1	-0.1633	-0.1044	1.09
PEP1-2	0.1213	0.0775	0.60
PEP2-1	-0.1662	-0.1062	1.13
PEP2-2	0.1672	0.1068	1.14
PEP3-1	0.3375	0.2156	4.65
PEP3-2	-0.2676	-0.1710	2.92
PEP4-1	0.1399	0.0894	0.80
MPI2-1	0.1390	0.0888	0.79
GOT3-2	0.1267	0.0810	0.66
G6P2-1	-0.2507	-0.1602	2.56
G6P2-2	0.1440	0.0920	0.85
G6P2-3	0.2157	0.1378	1.90
ACP1-1	-0.2039	-0.1303	1.70
ACP1-2	0.2680	0.1712	2.93
GDH1-1	-0.3463	-0.2213	4.90
GDH1-2	0.3232	0.2065	4.26
IDH1-1	0.4971	0.3176	10.09
PGM1-1	0.3389	0.2165	4.69
PGM1-2	-0.2914	-0.1862	3.47
ACO1-1	-0.1183	-0.0756	0.57
SKD1-1	0.1047	0.0669	0.45
SKD1-2	-0.1074	-0.0686	0.47
SKD2-1	0.4271	0.2729	7.45
FDP2-1	0.2716	0.1735	3.01
UGP1-1	0.1524	0.0974	0.95
UGP1-2	-0.1482	-0.0947	0.90
UGP2-1	-0.1594	-0.1019	1.04
FUM2-1	-0.1123	-0.0718	0.52
ADH2-1	0.2020	0.1290	1.67
ADH2-2	-0.2055	-0.1313	1.72
PGI2-1	-0.1436	-0.0918	0.84
PGI2-3	0.1445	0.0923	0.85
MDH3-1	-0.2031	-0.1298	1.68
MDH3-2	0.2208	0.1411	1.99
MDH4-1	-0.1629	-0.1041	1.08

Canonical correlation analysis:

- Dependent data-set:
  41 allozyme variables. Variables wit structure coefficients <0.10 dropped from analysis.
- Independent data set:
  Full second order (quadratic) model with climate Factors 1 and 2 [PCA]
- Structure coefficients:
  simple bivariate correlations betwee the variables and the canonical vari ates of the same data domain
- Crossloadings:
  simple bivariate correlations betwee the variables in one data domain and the canonical variates of the other domain
- Redundancy:
  Percent of variation in the respective variable accounted for by the canoni cal variate of the opposite domain
- The canonical correlation coefficien r is the correlation between the two variates of both data domains
- H-L-trace: Hotelling-Lawley-trace, multivariate equivalent of ratio of the variance due to the model to the error variance, adjusted for co-variances among variables (Tr)
- Model variance: aggregate variance of the model [Tr/(1+Tr)]
- Sign. variance: proportion of model variance of signif. variates only pt*Tr/(1+Tr)

CLIMATE DATA SET			
	CANONICAL VARIATE 1		
	Struc. coeff	% variance	
Factor 1	0.6288	39.54	0.9842	96.86
Factor 2	-0.0701	0.50	-0.1097	1.20
Factor 1 square	-0.0205	0.04	0.0322	0.10
Factor 2 square	-0.1727	3.00	-0.2704	7.31
F1 x F2	0.0952	0.90	0.1491	2.22

Eigenvalue	0.6897
% trace (pt)	65.10
unb. can. corr. r	0.5894
r square	0.3474
Significance	0.0001
H-L-trace (Tr)	1.060
Model variance	0.515
Sign. variance	0.335

Table 33: *Canonical correlation analysis: Allozyme and climate variables. Total sampling area*

Terms	Degrees of freedom	Type I Sum of Squares	R-Square	F-Ratio	Prob > F
Linear	2	164.118	0.400	137.000	0.0000
Quadratic	2	2.234	0.005	1.864	0.1563
Crossproduct	1	1.005	0.003	1.678	0.1959
Total model	5	167.357	0.408	55.868	0.0000
Variables	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
PCA-Factor1	3	155.124	51.708	86.307	0.0000
PCA-Factor2	3	3.497	1.166	1.946	0.1216
Residual	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
Lack of Fit	329	203.942	0.6198	1.217	0.1513
Pure Error	76	38.701	0.5092		
Total Error	405	242.643	0.5991		

***Table 34**: Evaluation of the fitted trend-surface: Relative importance of linear, quadratic and cross-product terms and of independent variables in the canonical model, test for lack-of-fit of the retained climate model: Total sampling area*

ionships between variate 1 and allele frequencies above are the same for climate *Factor 1* and allele frequencies.

The relative importance of the model terms and eventual lack-of-fit in the model was tested using least-squares regression procedures to estimate the quadratic response surface (PROC RSREG, © SAS Institute Inc., Cary, NC, USA)) and to describe the fitted trend surface of the response data set in terms of model fit and importance of independent variables in the model. Model fit was calculated by taking advantage of the locations having two samples per location. Results are given in *Table 34.*

The overall canonical model was highly significant ($p = 0.0000$) and adequately describes the data since lack-of-fit in the model was not significant ($p = 0.1513$). Climate *Factor 1* which was highly significant ($p = 0.0000$) clearly contributed most to the model sums of squares while *Factor 2* showed no significant ($p = 0.1216$) contribution. The relationships between *Factor 1* and the response surface was linear; both quadratic terms in the model as well as the interaction term were not significant. Inspection of the residuals, plotted against the canonical scores, revealed no bias in the model. Moreover, residuals did not significantly deviate from normality.

The plot of canonical scores on the first two response vectors, illustrated as mean scores for the seed zones, clearly separated the two regions on canonical variate 1 (*Figure 36, p. 147*). While variate 2 was not significant (it was only used to illustrate variate 1 in a two-dimensional plot), variate 1 was clearly related to temperature. Hence, axis 1 depicts associations of multilocus frequencies with the temperature regime at source location. All seed zones of Southwest Oregon showed positive scores on variate 1 while all zones from Central Oregon had negative mean scores. Zones 661, 842 and 662 had a mean score of about zero, indicating their close similarity to zones in Southwest Oregon. Within the two areas, patterns were rather complex, indicating a rather small scale geographic distribution of multilocus genetic variation, especially in Central Oregon.

To further aid the visual description of climatically associated multilocus genotypic frequencies in geographic space, original canonical scores of the first response variate were regressed against the geographic coordinates latitude and longitude, using local regression procedures. The complex pattern of multilocus genotypic frequencies could thus be visualized in geographic space by plotting the "predicted" scores of the response surface by latitude and longitude.

The trend surface of predicted scores on the first allozyme variate, resulting from canonical trend surface analysis based on total sampling area, is shown in the upper half of *Figure 37, p. 148.* A rather complex pattern of variation in multilocus genotypic frequencies was apparent.

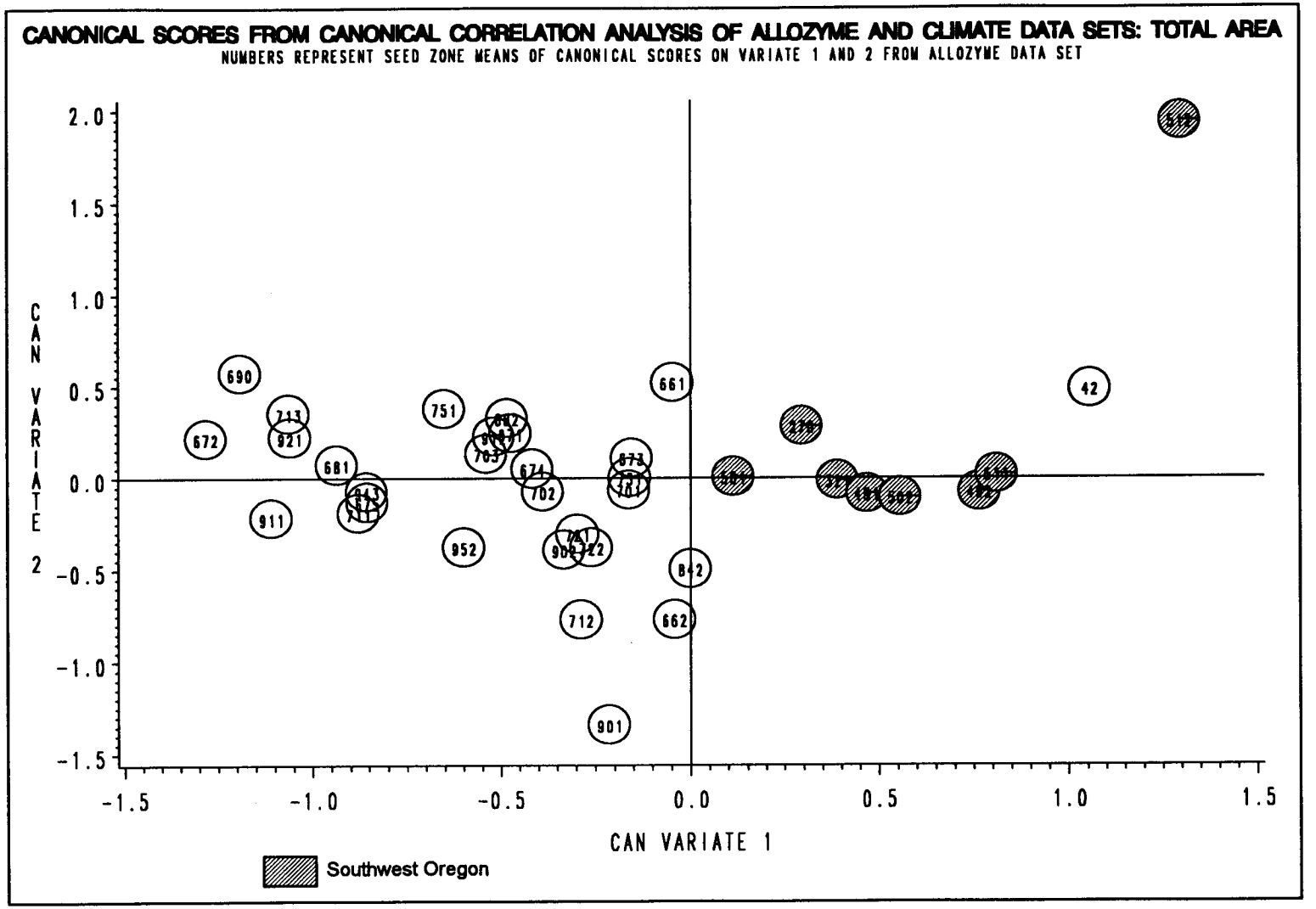


Figure 36: *Plot of canonical scores from canonical correlation analysis of allozyme and climate data sets: Total sampling area. Numbers represent seed zone mean scores on canonical variate 1 and 2.*

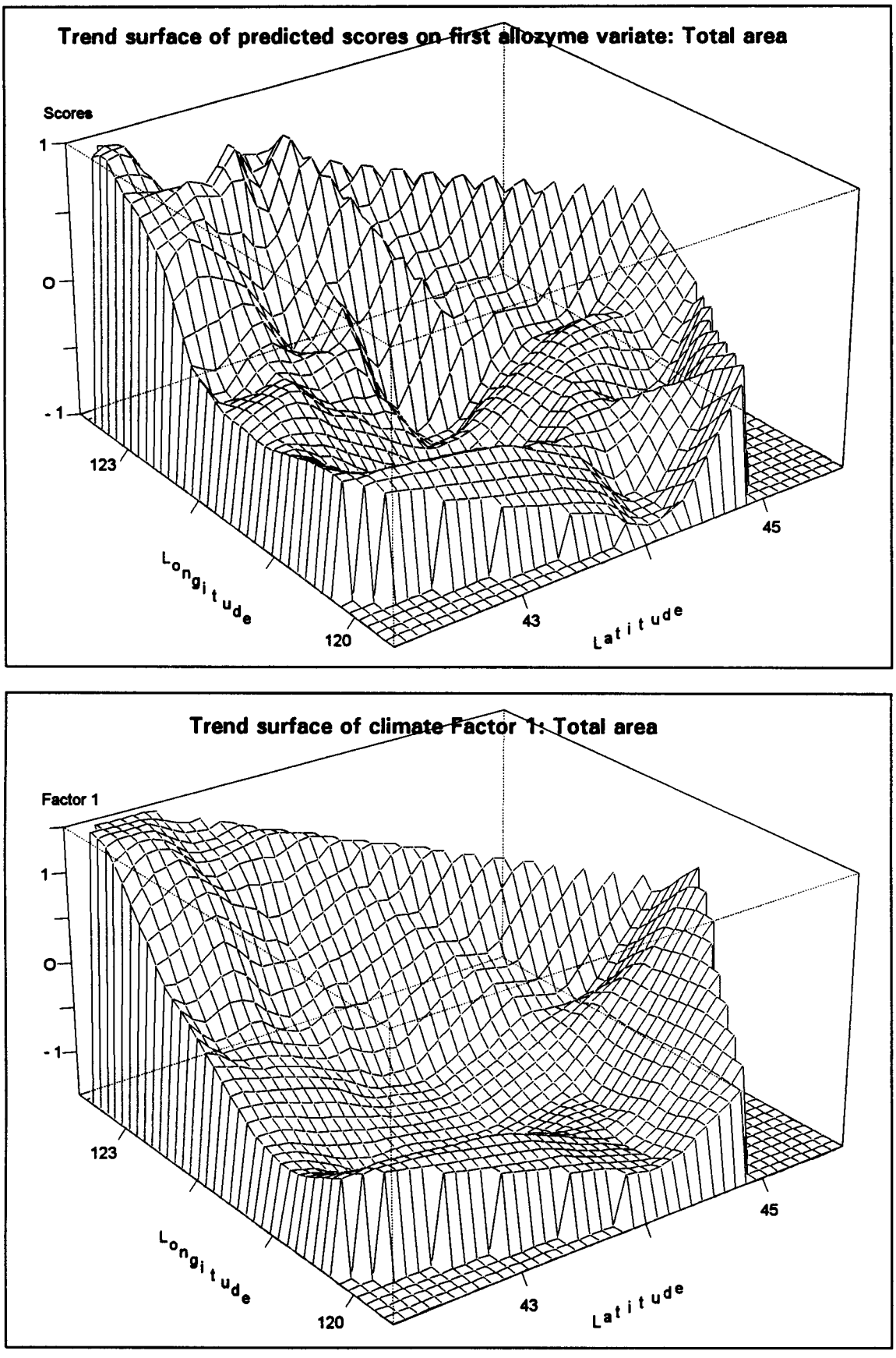


***Figure 37**: Upper: Trend surface of predicted scores of first allozyme variate for total sampling area. Based on canonical correlation analysis of allozymes and climate as reported in Table 33, p. 145. Lower: Trend surface of actual values of climate Factor 1 (temperature)*

Southwest Oregon showed a more or less clinal change in west-east direction, ranging from high positive values of about 1 in the most western locations near the Pacific coast to values close to zero near the Cascade crest. The pattern on the east side of the Cascade Range was more complex. In Central Oregon, multilocus frequencies varied on quite small geographic distances. In general, however, the northern and southern parts of the area were characterized by similar, moderately negative scores while the central part of the area was clearly distinguished by moderate to high negative values. The variation pattern resembled a double cline running in opposite directions from north to south, resulting in minimum values in the central parts of Oregon.

The canonical trend surface of multilocus genotypic frequencies strongly resembled the variation pattern of climate *Factor 1,* emphasizing the association of multilocus frequencies with temperature conditions at source location. Predicted scores not only paralleled the overall trend of variation in temperature but rather matched the pattern of variation in great detail, even on a very local scale (*Figure 37, p. 148, lower half).*

A different and more complex pattern of multivariate frequency distribution was apparent from canonical trend surface analysis restricted to Southwest Oregon samples only. Results are presented in *Table 35, p. 150.* Although all alleles with a structure coefficient smaller than 0.05 were dropped from the analysis, retaining only the most important 32 allozyme variables in the final model, the first canonical variate, accounting for 43% of the trace, was only significant at $p = 0.0508$, indicating that associations among multilocus genotypic frequencies and climate conditions in this area are weak. While the model explained 56% of total variance, a proportion of 24.6% of total variation in the 32 allozyme variables was explained by the first and only moderately significant variate.

The first variate was associated with several alleles at many loci. Correlations above 0.20 were apparent for 12 alleles *i.e. Lap2-3, Got1-3, Got2-1, Got2-2, Gdh1-1, Gdh1-2, Pgm1-2, Skd1-1, Skd1-2, Ugp1-2, Adh2-2* and *Mdh3-2.* Highest associations with variate 1 were apparent for *Got2-1* (0.371), *Got2-2* (0.375) and *Skd1-2* (0.278). High positive scores on variate 1 thus indicate high combined frequencies of *Lap2-3, Got2-2, Gdh1-1, Pgm1-2, Skd1-2* and *Ugp1-2* while low scores stand for high combined frequencies at *Got1-3, Got2-1, Gdh1-2, Skd1-1, Adh2-2* and *Mdh3-2.*

Climate variables were only weakly related to the first variate of the climate data set, supporting the conclusion that associations are only weak. The variate was dominated by the quadratic term of *Factor 2* and by the linear term of *Factor 1.* However, the first variate represented only 20% of variation in the squared values of *Factor 2* and 12.5% of variance in *Factor 1.* Nevertheless, this result clearly indicates that the moisture characteristics of the site have a higher influence on multilocus frequencies than the temperature regime. Scores on variate 1 were positively related with *Factor 2* and negatively with *Factor 1.* High combined frequencies of *Lap2-3, Got2-2, Gdh1-1, Pgm1-2, Skd1-2* and *Ugp1-2* are thus, at least in tendency, related to moist and cooler climate while higher combined frequencies of *Got1-3, Got2-1, Gdh1-2, Skd1-1, Adh2-2* and *Mdh3-2* predominate on the drier and warmer sites of the area.

Based on least-squares regression (*Table 36, p. 151*), the overall canonical model was highly significant ($p = 0.0000$) and adequately described the data, since lack-of-fit was not significant ($p = 0.1176$). Both climate factors contributed highly significantly to the model sums of squares. Inferred from the sums of squares, *Factor 1* had a slightly higher overall effect than *Factor 2.* Linear and quadratic effects were both highly significant ($p = 0.0013$ *or* $p = 0.0000$) while the interaction term had no significant influence in the model. Inspection of the residuals, plotted against the canonical scores, revealed no bias in the model. In addition, residuals did not significantly deviate from normality.

The plot of canonical scores on the first two response vectors, illustrated as mean scores for the seed zones, is shown in *Figure 38, p. 151.* The clinal pattern which was apparent in the overall analysis, was replaced by a more complex pattern, suggesting a more locally, ecotypic distribution of multilocus frequencies if the variation pattern is looked at from within the region.

A better illustration of the complexity of the pattern of multilocus frequency distribution in

ALLOZYME DATA SET			
Locus	CANONICAL VARIATE 1		
Allele	Struc. coeff	Crossloadings	Redundancy %
MNR1-3	-0.1554	-0.0937	0.88
LAP2-1	-0.1431	-0.0862	0.74
LAP2-2	0.1032	0.0622	0.39
LAP2-3	0.2455	0.1480	2.19
LAP3-1	0.1500	0.0904	0.82
PEP2-1	0.1496	0.0902	0.81
PEP2-2	-0.1417	-0.0854	0.73
PEP3-1	-0.1124	-0.0678	0.46
PEP3-2	-0.1691	-0.1019	1.04
MPI2-1	0.1946	0.1173	1.38
MPI2-2	-0.1825	-0.1100	1.21
GOT1-1	0.1449	0.0874	0.76
GOT1-2	0.1616	0.0974	0.95
GOT1-3	-0.2074	-0.1250	1.56
GOT2-1	-0.3711	-0.2237	5.00
GOT2-2	0.3747	0.2259	5.10
ACP1-3	-0.1196	-0.0721	0.52
ACP1-4	0.1231	0.0742	0.55
GDH1-1	0.2137	0.1288	1.66
GDH1-2	-0.2240	-0.1351	1.82
PGM1-2	0.2531	0.1526	2.33
SKD1-1	-0.2388	-0.1440	2.07
SKD1-2	0.2787	0.1680	2.82
SKD2-1	-0.1328	-0.0800	0.64
UGP1-2	0.2340	0.1410	1.99
FUM2-1	0.1634	0.0985	0.97
ADH2-1	0.1958	0.1180	1.39
ADH2-2	-0.2164	-0.1305	1.70
MDH1-1	-0.1527	-0.0920	0.85
MDH3-1	0.1466	0.0884	0.78
MDH3-2	-0.2430	-0.1465	2.15
MDH4-1	0.1659	0.1000	1.00
Factor 1		-0.2136	4.56
Factor 2		0.0242	0.06
Factor 1 square		-0.0884	0.78
Factor 2 square		0.2677	7.17
F1 x F2		0.0583	0.34
Eigenvalue	0.5708		
% of trace (pt)	43.36		
unb. can. corr. r	0.4928		
r square	0.2429		
Significance	0.0508		
H-L-trace (Tr)	1.3164		
Model variance	0.5683		
Sign. variance	0.2464		

Canonical correlation analysis:

- Dependent data-set:
  32 allozyme variables. Variables with structure coefficients <0.10 dropped from analysis.

- Independent data set:
  Full second order (quadratic) model with climate Factor 1 and 2 [PCA]

- Structure coefficients:
  simple bivariate correlations between the variables and the canonical variates of the same data domain

- Crossloadings:
  simple bivariate correlations between the variables in one data domain and the canonical variates of the other domain

- Redundancy:
  Percent of variation in the respective variable accounted for by the canonical variate of the opposite domain

- The canonical correlation coefficient r is the correlation between the two variates of both data domains

- H-L-trace: Hotelling-Lawley-trace, multivariate equivalent of ratio of the variance due to the model to the error variance, adjusted for covariances among variables (Tr)

- Model variance: aggregate variance of model Tr/(1+Tr)

- Sign. variance: proportion of model variance of signif. variates only pt*Tr/(1+Tr)

CLIMATE DATA SET		
CANONICAL VARIATE 1		
Struc. coeff	% variance	
-0.3544	12.56	
0.0402	0.16	
-0.1467	2.15	
0.4442	19.73	
-0.0967	0.94	

Table 35: Canonical correlation analysis: Allozyme and climate variables. Southwest Oregon

Terms	Degrees of freedom	Type I Sum of Squares	R-Square	F-Ratio	Prob > F
Linear	2	9.019	0.048	6.895	0.0013
Quadratic	2	59.285	0.315	45.325	0.0000
Crossproduct	1	0.014	0.000	0.021	0.8857
Total model	5	68.317	0.363	20.892	0.0000
Variables	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
PCA-Factor1	3	50.796	16.932	25.891	0.0000
PCA-Factor2	3	40.944	13.648	20.869	0.0000
Residual	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
Lack of Fit	155	106.550	0.6874	1.466	0.1176
Pure Error	28	13.131	0.4688		
Total Error	183	119.682	0.6541		

***Table 36**: Evaluation of fitted trend-surface: Relative importance of linear, quadratic and cross-product terms and of independent variables in the canonical model, test of lack-of-fit of fitted surface: Southwest Oregon*

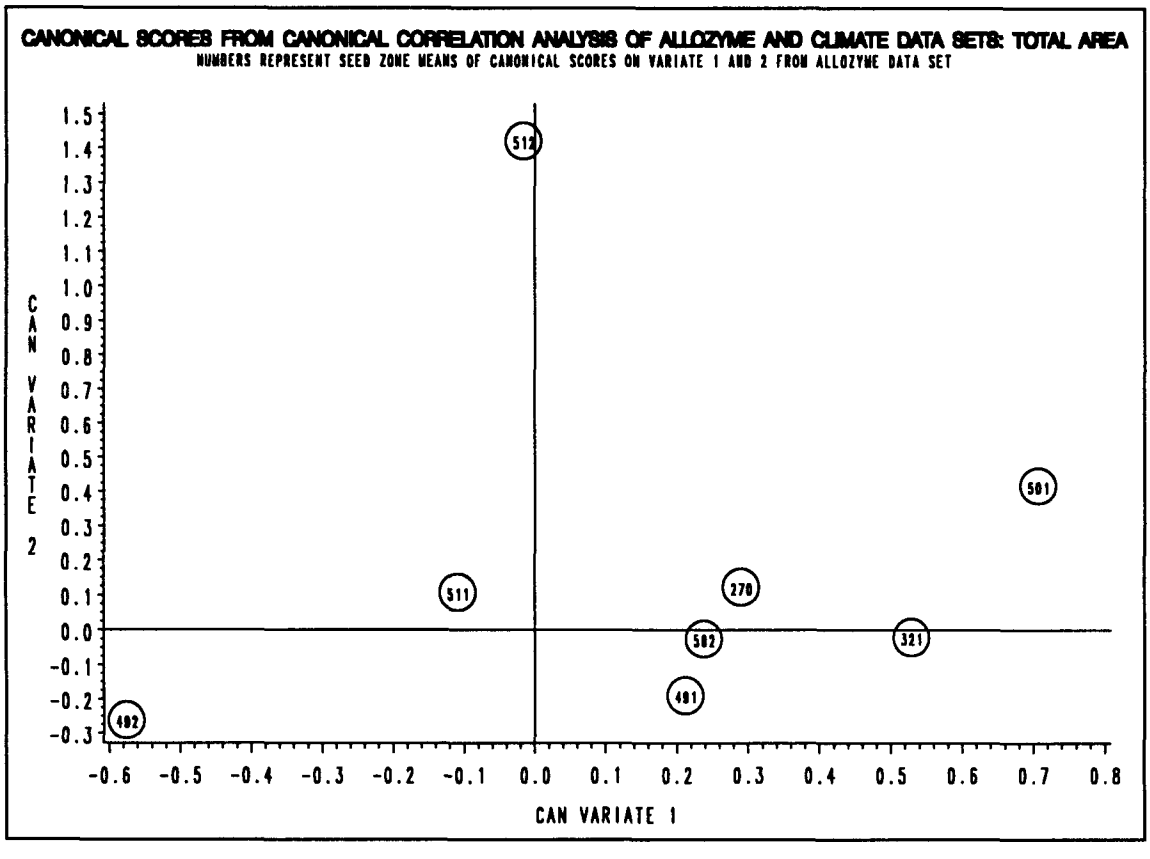


***Figure 38**: Plot of canonical scores from canonical correlation analysis of allozyme and climate data sets: Southwest Oregon. Numbers represent seed zone mean scores on canonical variate 1 and 2.*

Southwest Oregon is provided by the trend surface of predicted scores, shown in the upper half of *Figure 39, p. 152*. The pattern was in fact highly variable at rather small geographic distances. In general, however, the variation pattern resembled a double cline running in opposite direction from west to east, resulting in minimum values in the central parts of the area.

The trend surface of predicted scores on the first allozyme variate clearly resembled the variation pattern of the squared values of climate *Factor 2*. Although association in the canonical model were rather weak, an overall resemblance of the two variation patterns was apparent

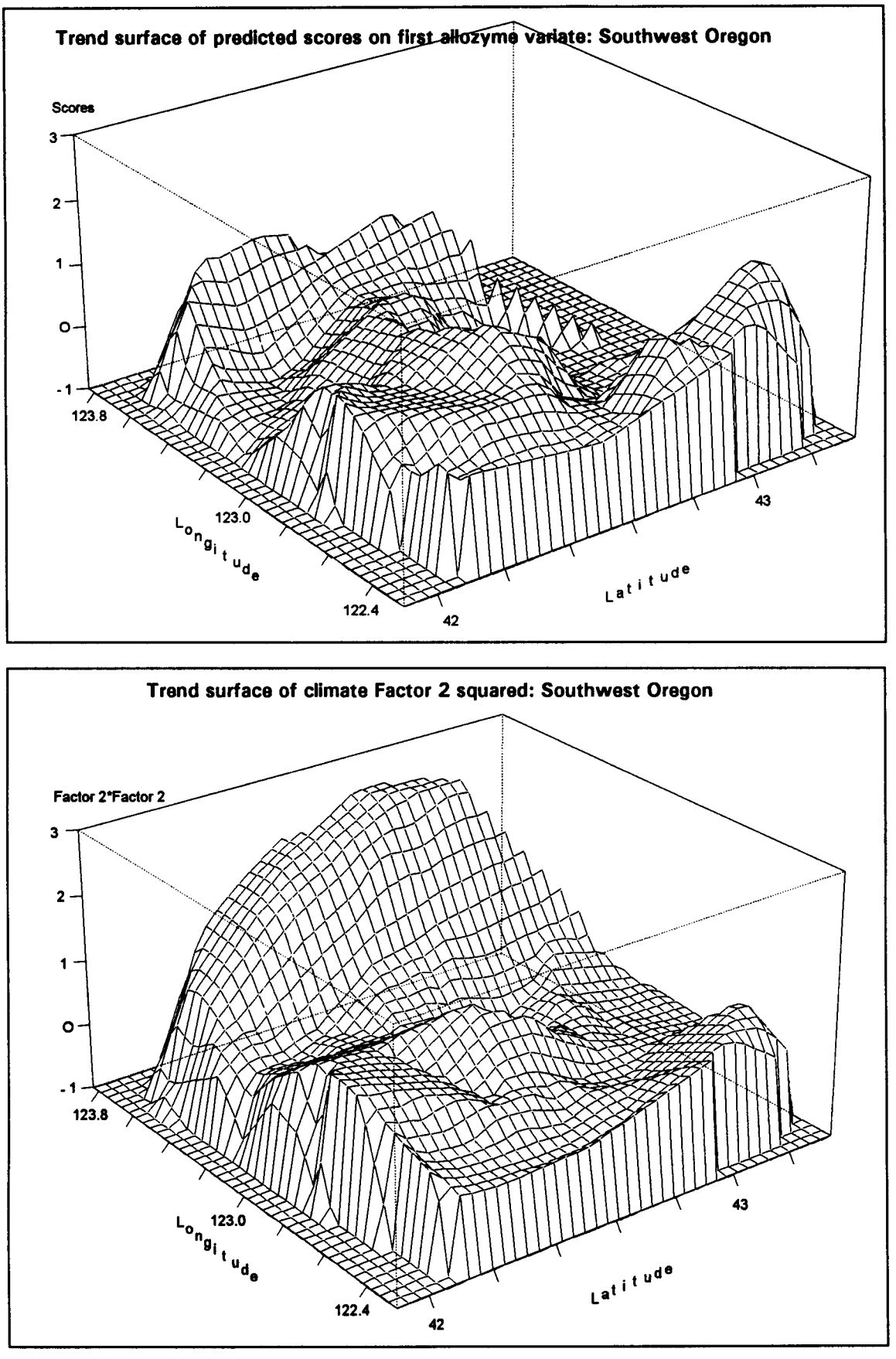


Figure 39: *Upper: Trend surface of predicted scores of first allozyme variate for Southwest Oregon. Based on canonical correlation analysis of allozymes and climate as reported in Table 35, p. 150. Lower: Trend surface of squared values of climate Factor 2 (water balance)*

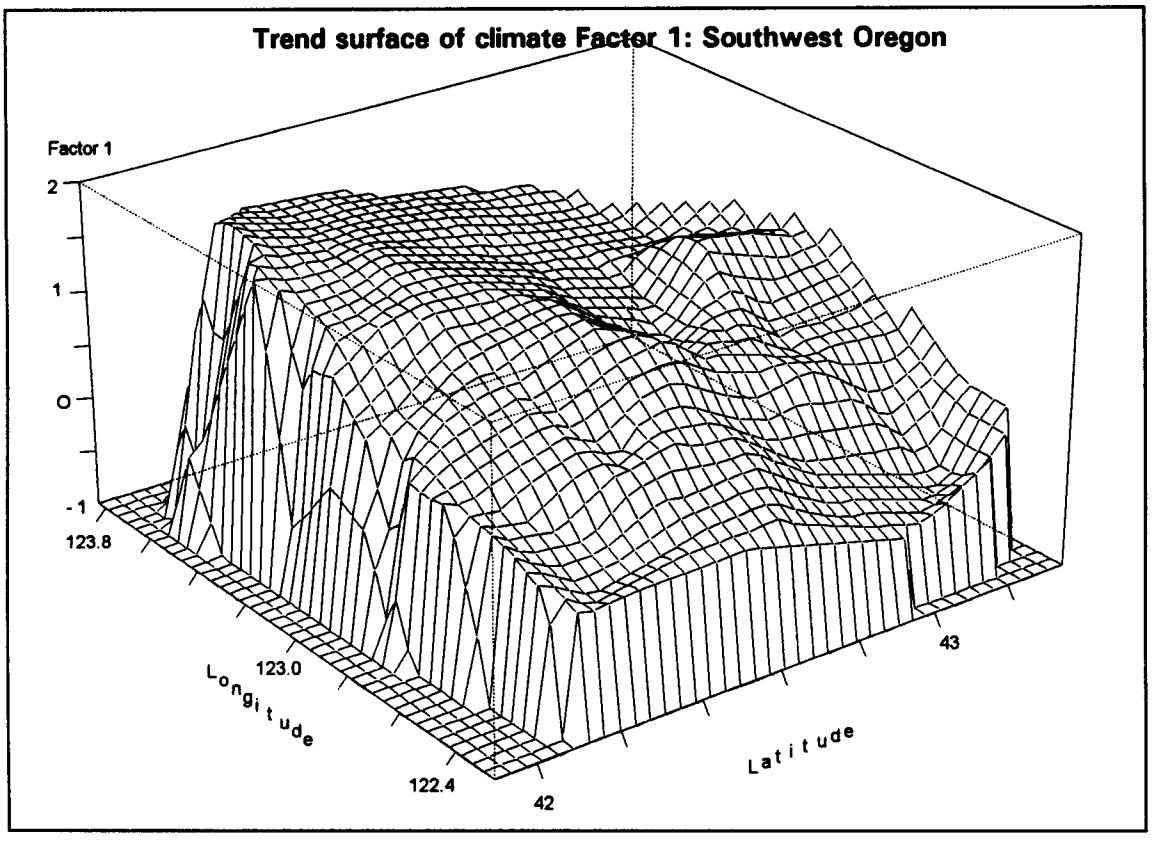


***Figure 40**: Trend surface of actual values of climate Factor 1 (temperature): Southwest Oregon*

(*Figure 39, p. 152, lower half*), indicating that the areas with low negative scores paralleled areas with comparatively dry site conditions while high positive scores were associated with moist conditions. *Figure 40* depicts the variation pattern of temperature in Southwest Oregon. Although climate *Factor 1* significantly contributed to the model sums of squares, there was only a minor resemblance between the trend surface of predicted scores of the first allozyme variate and the pattern of variation in temperature.

Canonical trend surface analysis of the Central Oregon sample produced a highly significant ($p = 0.0081$) first variate which accounted for 43% of the patterned variation (*Table 37, p. 154*). The Hotelling-Lawley trace indicated that the model explained 54% of total variance. Since only 43% of the trace was accounted for by the significant vector, a proportion of 23.3% of the total variance in the retained 33 allozyme variables could be significantly explained by the climate model.

The first canonical variate of the allozyme domain was associated with many alleles at several loci. 13 alleles had a correlation higher than 0.20 with the first variate *i.e. Mnr1-2, Mnr2-1, Mnr2-2, Pep3-1, Pep3-2, Pep4-1, Mpi1-1, Idh1-1, Pgm1-1, Pgm1-2, Skd2-1, Ugp1-2* and *Mdh3-2*. Especially high structure coefficients were apparent for *Idh1-1* (0.401), *Pep3-1* (0.385), *Pgm1-1* (0.361), *Skd2-1* (0.355) and *Pgm1-2* (0.317).

High positive scores on variate 1 thus stand for high combined frequencies of *Mnr2-2, Pep3-1, Pep4-1, Idh1-1, Pgm1-1, Skd2-1* and *Mdh3-2* while low scores represent high combined frequencies at *Mnr2-1, Pep3-2, Mpi1-1, Pgm1-2* and *Ugp1-2*.

The first variate of the independent data domain was dominated by the linear and the quadratic term of *Factor 1*. The variate represented 75.5% of variance of *Factor 1* or 59% of variance of the squared values of *Factor 1*. Climate Factor 2 was only weakly related with the first variate (12% or 21%), the interaction term, however, was represented by 35% of its variance.

ALLOZYME DATA SET			
Locus	CANONICAL VARIATE 1		
Allele	Struc. coeff	Crossloadings	Redundancy %
MNR1-1	0.1438	0.0834	0.70
MNR1-2	-0.2111	-0.1224	1.50
MNR2-1	-0.2102	-0.1219	1.49
MNR2-2	0.2504	0.1452	2.11
PEP1-1	-0.1534	-0.0890	0.79
PEP1-2	0.1742	0.1010	1.02
PEP2-1	-0.1112	-0.0645	0.42
PEP2-2	0.1733	0.1005	1.01
PEP3-1	0.3851	0.2234	4.99
PEP3-2	-0.2631	-0.1526	2.33
PEP4-1	0.2804	0.1626	2.64
MPI1-1	-0.2900	-0.1682	2.83
MPI1-2	0.1331	0.0772	0.60
MPI2-1	0.1362	0.0790	0.62
MPI2-3	-0.1547	-0.0897	0.80
GOT1-2	-0.1074	-0.0623	0.39
GOT3-2	-0.1125	-0.0653	0.43
ACP1-4	0.1295	0.0751	0.56
IDH1-1	0.4012	0.2327	5.41
PGM1-1	0.3608	0.2093	4.38
PGM1-2	-0.3166	-0.1836	3.57
ACO1-3	-0.1837	-0.1065	1.13
SKD2-1	0.3548	0.2058	4.23
UGP1-1	0.1879	0.1090	1.19
UGP1-2	-0.2021	-0.1172	1.27
UGP3-1	-0.1602	-0.0929	0.86
FUM2-2	0.1724	0.1000	1.00
ADH2-1	-0.1134	-0.0658	0.43
ADH2-2	0.1134	0.0658	0.43
PGI2-3	-0.0869	-0.0504	0.25
MDH3-1	-0.1780	-0.1032	1.07
MDH3-2	0.2567	0.1489	2.22
MDH4-1	-0.1519	-0.0881	0.78
Factor 1		0.5036	25.36
Factor 2		-0.2020	4.08
Factor 1 square		-0.4454	20.30
Factor 2 square		-0.2664	14.33
F1 x F2		0.3420	11.75
Eigenvalue	0.5069		
% of trace (pt)	43.11		
unb. can. corr. r	0.4790		
r square	0.2294		
Significance	0.0081		
H-L trace (Tr)	1.176		
Model variance	0.5404		
Sign. variance	0.2330		

Canonical correlation analysis:

- Dependent data-set:
  33 allozyme variables. Variables with structure coefficients <0.10 dropped from analysis.
- Independent data set:
  Full second order (quadratic) model with climate Factor 1 and 2 [PCA]
- Structure coefficients:
  simple bivariate correlations between the variables and the canonical variates of the same data domain
- Crossloadings:
  simple bivariate correlations between the variables in one data domain and the canonical variates of the other domain
- Redundancy:
  Percent of variation in the respective variable accounted for by the canonical variate of the opposite domain.
- The canonical correlation coefficient r is the correlation between the two variates of both data domains
- H-L-trace: Hotelling-Lawley-trace, multivariate equivalent of ratio of the variance due to the model to the error variance, adjusted for covariances among variables (Tr)
- Model variance: aggregate variance of model (Tr/(1+Tr)
- Sign. variance: proportion of model variance of signif. variates only

CLIMATE DATA SET		
CANONICAL VARIATE 1		
Struc. coeff	% variance	
0.8683	75.40	
-0.3483	12.13	
-0.7680	58.98	
-0.4594	21.10	
0.5897	34.77	

***Table 37**: Canonical correlation analysis: Allozyme and climate variables: Central Oregon*

Terms	Degrees of freedom	Type I Sum of Squares	R-Square	F-Ratio	Prob > F
Linear	2	60.136	0.272	44.285	0.0000
Quadratic	2	11.293	0.051	8.316	0.0003
Crossproduct	1	2.915	0.013	4.293	0.0395
Total model	5	74.344	0.336	21.899	0.0000
Variables	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
PCA-Factor1	3	56.188	18.729	27.586	0.0000
PCA-Factor2	3	14.316	4.772	7.028	0.0002
Residual	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
Lack of Fit	167	108.760	0.6512	0.842	0.7879
Pure Error	49	37.895	0.7733		
Total Error	216	146.656	0.6789		

Table 38: Evaluation of fitted trend-surface: Relative importance of linear, quadratic and cross-product terms and of independent variables in the canonical model, test of lack-of-fit of fitted surface: Central Oregon

Inferred from least-squares regression (*Table 38*), the overall canonical model was highly significant ($p = 0.0000$) and adequately described the data since lack-of-fit was not significant ($p = 0.7879$). Both climate factors were highly significant in contributing to the model sums of squares ($p = 0.0000$ *or* $p = 0.0002$), the contribution of *Factor 1* being 4 times the contribution of *Factor 2*, however. Linear, quadratic and cross-product terms all significantly ($p= 0.0000$, $p = 0.0003$, $p = 0.0395$) contributed to the model, according to the sums of squares, their importance being different, however. Inspection of the residuals, plotted against the canonical scores, showed no bias in the model and residuals were normally distributed.

Plotting mean canonical scores of seed zones of the first two vectors showed a clear separation of the northern and southern zones from the zones in the center of the area as regards the scores on axis 1 (*Figure 41, p. 156*). Northern and southern zones all had positive scores while the central zones all had negative scores (the exception being zone 682).

A more detailed picture of the multilocus frequency distribution in geographic space is provided by the trend surface of predicted scores, shown in the upper half of *Figure 42, p. 157*. The pattern was rather complex, showing roughly a double cline running in opposite direction from north to south, combined with three clines and two minima in the west-east direction.

The trend surface of scores of the first allozyme variate strongly resembled the variation pattern of climate *Factor 1*, supporting the dominant association of multilocus frequencies with temperature. Variation in the first allozyme variate paralleled temperature not only in tendency but matched the variation even on a quite small local scale. Areas with low canonical scores *i.e.* with high combined frequencies at *Mnr2-1*, *Pep3-2*, *Mpi1-1*, *Pgm1-2* and *Ugp1-2* were associated with low winter and yearly mean temperatures which prevail in the center of the sampling area. High positive scores on variate 1, standing for high combined frequencies of *Mnr2-2*, *Pep3-1*, *Pep4-1*, *Idh1-1*, *Pgm1-1*, *Skd2-1* and *Mdh3-2* were however associated with higher winter, but also higher summer temperatures found in the northern and southern parts of the area.

6.3 Proportion of model variance caused by indirect effects, inferred by canonical and partial canonical analysis

The estimated association among multilocus frequencies and climate conditions at source locations reported in *section 6.2* may have been caused in part by indirect effects which parallel climate variation. Since climate and location are correlated, spatial and environmental variation are in part redundant, thus proportions of the association attributed to climate variation may in fact be caused by pure spatial effects. As has been discussed in detail in *section 4.4*, historical effects may especially have influenced the spatial structure of allozyme variation in the two re-

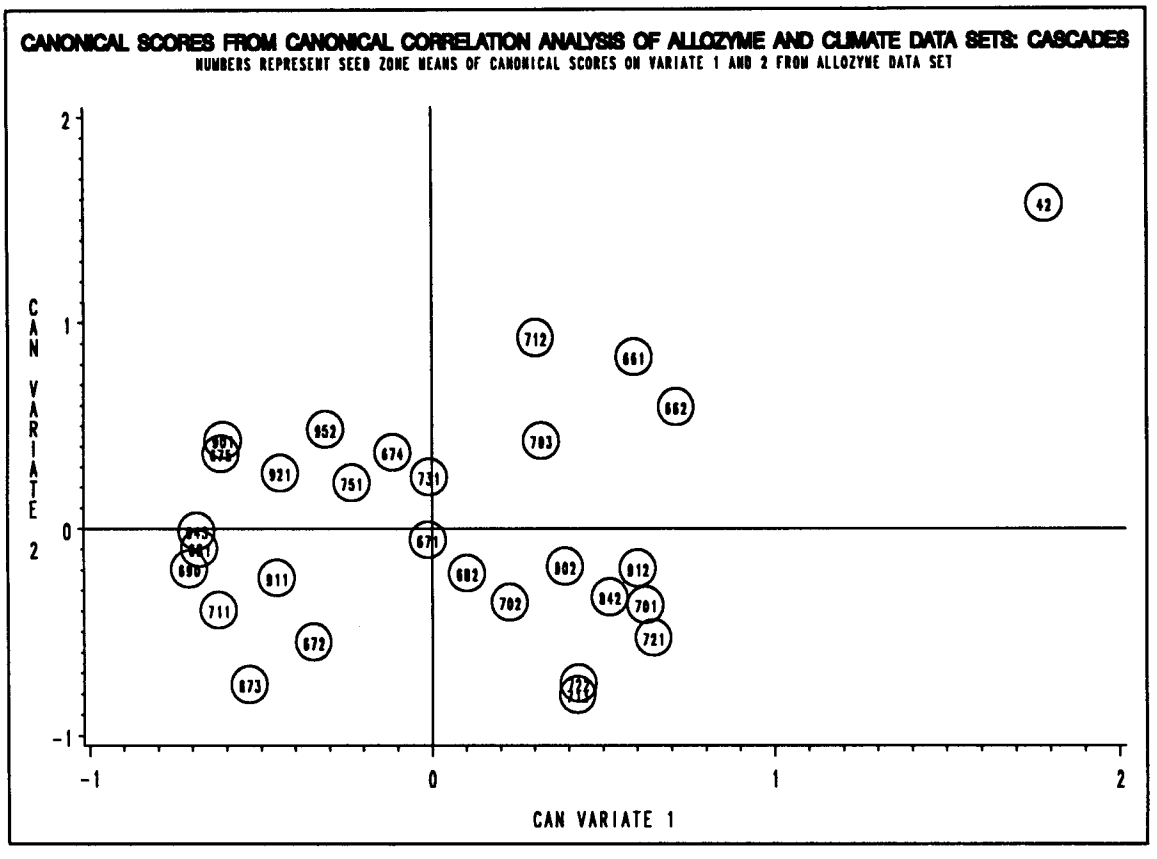


Figure 41: Plot of canonical scores from canonical correlation analysis of allozyme and climate data sets: Central Oregon. Numbers represent seed zone mean scores of canonical variates 1 and 2.

gions. Since regional differences in genetic structures parallel the regional differences in climate conditions, both effects may have contributed directly or indirectly to the total variance in the canonical model.

Basically, three possible explanations may be envisioned for the genotypic differentiation between the two regions. A different evolutionary history *i.e.* migration from two different refugia may be the major source of differentiation. Alternatively, observed differences may be primarily a consequence of adaptation to the different climate conditions which are found in the two areas. As a third possibility, patterns of variation may be a result of a combination of both historical as well as environmental effects. The spatial structuring of genotypic frequencies thus poses the problem of the relative contribution of different factors whose interaction may result in an overlaid effect in space (BORCARD, 1992).

Consequently, estimated associations between multilocus frequencies and climate conditions reported in *section 6.2* may partly be due to other effects than climate, especially in the analysis based on total sampling area. In order to describe true associations among genotypic frequencies and climate conditions, total variance in the canonical model must be partitioned into components which are due to climate effects and components which are caused by such indirect effects.

Had the base populations in the two regions come from a different origin, significant spatial patterns of variation would be expected as a consequence of migration from different sources and at different rates (*section 5.4*). In this case, spatial patterns of variation would most likely be independent of environmental variation. Moreover, if the genotypic differentiation between the two regions were caused by a separate evolutionary history, strong associations between genotypic frequencies and geographic locations should exist (caused by migration) even under identical climatic conditions. Consequently, spatial associations which are independent of environmental associations should indicate such historical effects. Therefore, in order to separate

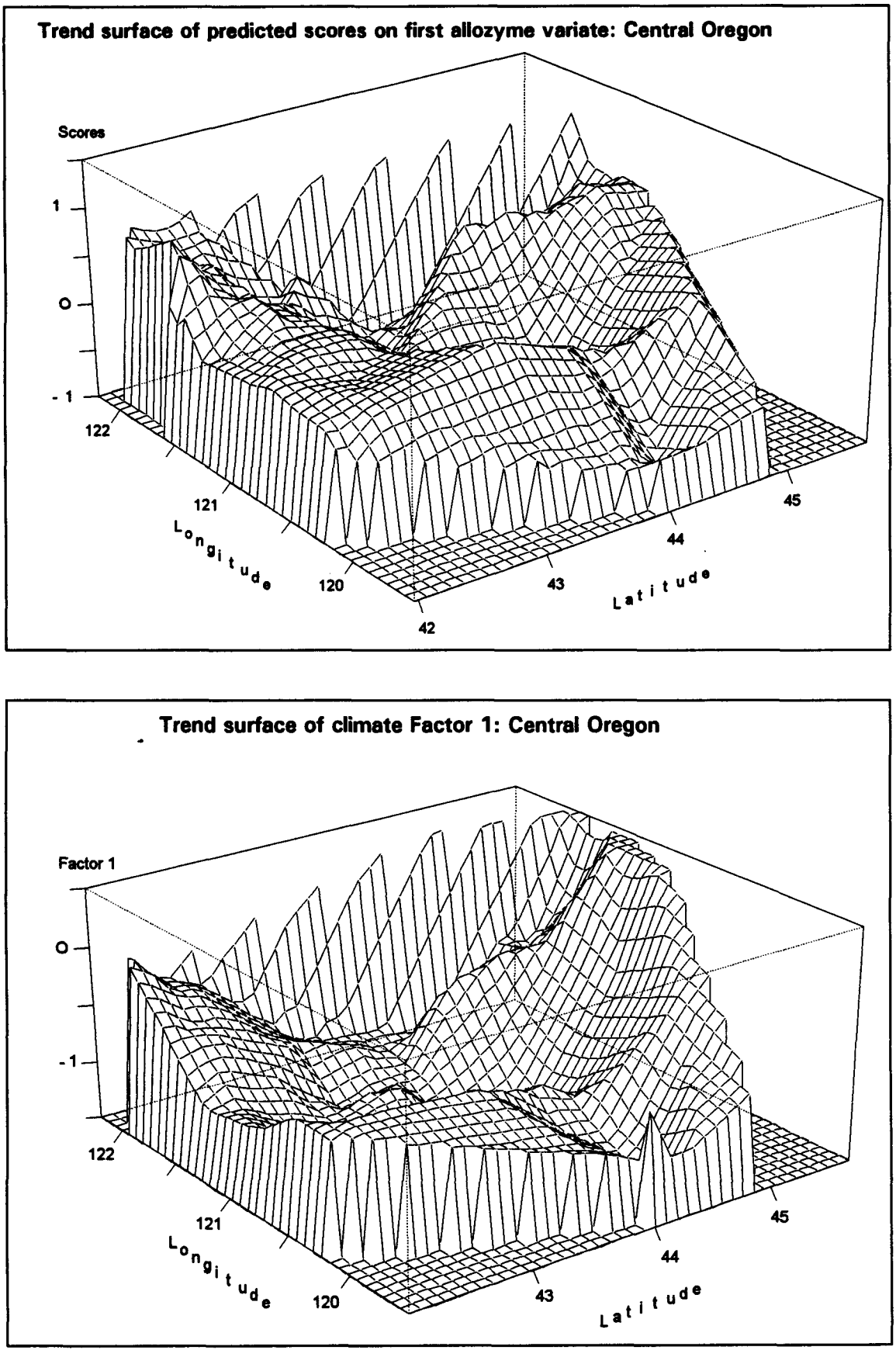


***Figure 42**: Upper: Trend surface of predicted scores of first allozyme variate for Central Oregon. Based on canonical correlation analysis of allozymes and climate as reported in Table 37, p. 154. Lower: Trend surface of actual values of climate Factor 1 (temperature)*

historical from environmental effects, the total variance in the canonical model must be partitioned into components which are due to either pure spatial, pure environmental or interaction effects.

Following the basic methodology which was proposed by BORCARD et al. (1992), total variance in the canonical model was therefore partitioned into the following four components:

a) *the non-spatial environmental variation* in the model. This is the fraction of the variation which can be explained independently of any spatial structure (*pure environmental variation*)
b) *the spatial variation in the model that is shared by the environmental variation.* This common variation is partly a consequence of the associations of the dependent variables with spatially structured environmental conditions, but a certain part may also be non-causal or even historical (*interaction of spatial and environmental variation*)
c) *the spatial variation* in the model that is not shared by environmental *variation (pure spatial variation*, due to historical effects or recent drift)
d) *the unexplained variation.* This is the fraction of variation in the model that is neither explained by spatial coordinates nor by environmental conditions.

Certain proportions of the shared variation (*b*) may be due to other causes than spatially structured environmental conditions. Historical effects such as migration are primarily not expected to have any relationship with environmental variation thus interaction should either be due to spatially structured environments or due to non-causal effects. Historical effects may not be ruled out completely, however, since migration may have been related to certain environmental conditions during colonization, certain habitats providing better habitat conditions for the establishment of ponderosa pine and faster colonization than others. Nevertheless, since clear spatial structures of habitat conditions were common in the sampling area (*section 6.1*), spatially structured environments are the most likely cause for the common variation expressed in the interaction term.

In order to partition total model variance into these four components, different full and partial canonical models were analyzed. In partial canonical correlation, the specified variables are partialled out of the model effects before canonical correlation analysis is performed. The 5 models analyzed and presented in *Table 39, p. 159* were all second order models including the following independent effects or variance components:

1. *Full model* with climate *Factor 1*, *Factor 2*, latitude and longitude, representing all four possible components (a-d).
2. *Climate model* with climate *Factor 1* and *Factor 2*, representing components *a*, *b* and *d*
3. *Location model* with latitude and longitude, representing components *b*, *c* and *d*
4. *Partial climate model* with climate *Factor 1* and *Factor 2*, location effects are partialled out. This model represents the components *a* and *d*
5. *Partial location model* with latitude and longitude, climate effects are partialled out. This model represents components *c* and *d*
6. *Interaction of climate and location*, representing component *b*. The interaction was calculated by subtracting the variance accounted for by model 2 from that of model 4 (or by model 3 from that of model 5). Minor differences between the two methods of calculation were observed due to rounding and deviations from multivariate normality
7. *Unexplained variance*, representing component *d*. This proportion is the difference between the full model variance ($T_r/(1+T_r)$) and one hundred percent ($1+T_r$)

The canonical eigenvalues of the different models were used as measures for the amount of variation accounted for by the explanatory variables specified in the models. To transform eigenvalues into percentages of total variation (to get the variation explained by the model), individual eigenvalues were divided by $1+T_r$ from the full model (# 1). Percentage relative contribution was calculated relative to the eigenvalue of the full model which was set to 100 %.

Partitioning of total model variance into spatial and environmental effects						
#	Model and Effects used in canonical analysis	Models terms in canonical analysis [all second order]	Number of model terms	Sum of Eigenvalue [H-L-trace	% variance of	
					Full model [2.2923]	relative contribution %
1	Full model with climate and location	Factor1/Factor2 Latitude/Longitude	10	1.2923	56.40	100.0
2	Climate model	Factor1/Factor2	5	1.0600	46.25	82.0
3	Location model	Latitude/Longitude	5	0.5799	25.30	44.9
4	Climate model with location effects partialled out	Partial climate Factor1/Factor2	5	0.6674	29.10	51.6
5	Location model with climate effects partialled out	Partial location Latitude/Longitude	5	0.2325	10.15	18.0
6	Interaction Climate*Location	Model 2 - Model 4 Model 3 - Model 5	- -	0.3926 [0.3474]	17.10 [15.15]	30.4 [26.28]
7	Unexplained variance	-	-	-	43.60	-
Partitioning of the significant part of variance into spatial and environmental effects						
#	Model and Effects used in canonical analysis	Models terms in canonical analysis [all second order]	Number of model terms	Sum of Eigenvalue [H-L-trace	% variance of	
					Full model [2.2923]	relative contribution %
1	Full model with climate and location	Factor1/Factor2 Latitude/Longitude	10	0.7446	32.48	100.0
2	Climate model	Factor1/Factor2	5	0.6897	30.09	92.6
3	Location model	Latitude/Longitude	5	0.5103	22.26	68.5
4	Climate model with location effects partialled out	Partial climate Factor1/Factor2	5	0.2625	11.45	35.2
5	Location model with climate effects partialled out	Partial location Latitude/Longitude	5	0.0551	2.40	7.4
6	Interaction Climate*Location	Model 2 - Model 4 Model 3 - Model 5	- -	0.4272 [0.4552]	18.64 [19.86]	57.4 61.1
7	Unexplained variance				67.52	-

Table 39: *Decomposition of climatic and spatial (historical) effects in the canonical trend surface analysis on total sampling area as reported in Table 33, p. 145. Estimates of the variance components due to climate and location effects using full and partial canonical models. Upper: Partitioning of total model variance [sum of eigenvalues, Hotelling-Lawley trace] into spatial and environmental effects Lower: Partitioning of the variance of significant variates only*

Results of the decomposition of model variance into the four components of interest are furnished in *Table 39*. The table in the upper half presents the proportions found in total model variance. Since not all this variation is significant, however, the more interesting results of decomposition are shown in the lower half of the table which lists the proportions of the four components relative to the significant part of model variance only.

Using the same 41 allozyme variables as in the analysis in *section 6.2*, the full model significantly explained 32.5% of total variance. The climate model, on the other hand, accounted for 30.1% of the significant variance. The difference of 2.4% thus representing the contribution of pure spatial effects that are not redundant with the climate variables. This difference equaled the result of model 5, where climate was partialled out, leaving only the pure spatial effects in the model. Of the significantly explained model variance of 32.5%, 11.5% were due to pure

climate effects, 18.6% were caused by the shared variation of climate and location (caused by spatially structured climate conditions) and only 2.4% were due to pure spatial effects. Expressed as a percentage of contribution to the significant variance of the full model, pure climate effects contributed 35.2% to the sum of all eigenvalues, the interaction of climate and location 57.4% and pure spatial or historical effects only 7.4%. Hence, from the total sum of eigenvalues, 92.6% were due to direct climate effects while pure spatial or historical effects accounted for a proportion of 7.4% only.

Assuming the same proportions as above, it may be estimated that from the significant variance of 33.5% in the climate model reported in *section 6.2*, 11.8% were caused by pure climate effects which are independent from a spatial structure (similar conditions occurring again at different locations), 19.2% were due to the interaction between climate and location (spatially structured environments) and only 2.5% had a pure spatial or historical basis. True associations between multilocus frequencies and climate conditions therefore may be estimated to be 31%.

6.4 Relative seed transfer risks, inferred from the canonical trend surface of multilocus genotypic frequencies

Canonical scores of the trend surface represent the relative position in the joint multivariate space defined by the two sets of variables. Hence, the trend surface of canonical scores represents the pattern of variation in multilocus genotypic frequencies which is maximally associated with the climatic conditions at the source locations, thus providing a picture of the adaptive pattern of allozyme markers. Accordingly, canonical scores may be used to estimate relative seed transfer risks between any pair of locations situated on the trend surface.

Relative transfer risk may be estimated for seed transfer among seed zones. Mean canonical scores of the zones and the variance in scores found within the zones may be used to calculate transfer risk. Alternatively, since contour plots of predicted scores indicate regions of similarity in environmentally related multilocus frequencies, the trend surface itself may be used to delineate areas of given relative transfer risks. Intervals between contours may be scaled such that the areas encompassed by two contours represent a relative transfer risk that is smaller than a certain value. Contour maps constructed in this way thus delineate areas within which seed may be transferred without exceeding the given value of transfer risk. The reasoning behind relative seed transfer risk estimates and the methods used to calculate transfer risks were described in detail in *section 3.6.5*. Contours on the trend surface, representing certain relative transfer risks, were mapped onto the official seed zone map of Oregon. Maps constructed in this way provide a means of comparing the existing seed zones with the resulting zones based on transfer risk estimates as inferred from multilocus allozyme frequency distributions. Contour lines representing a given relative transfer risk were derived from the canonical trend surface using S-PLUS statistical software (© StatSci, Statistical Sciences Inc., Seattle, WA, USA). Coordinates of contour lines were then transferred into a geographic information system (arcinfo) which produced overlay maps in true geographic coordinates.

Relative transfer-risk between all pairs of seed zones in the total sampling area, based on mean canonical scores of the first allozyme variate, are furnished in *Table 40, p. 161*. To facilitate analysis, the few individuals sampled in zone 42 were added to zone 661, the individuals of zone 90 were combined with zone 512. Estimated transfer risks varied widely; ranging from 1% to 90% of mismatch between the frequency distributions at source locations and potential planting sites. Risks were especially high when transferring individuals from seed zones in Southwest Oregon to zones in Central Oregon. To better illustrate relative transfer risk among zones, an average linkage cluster analysis was performed, clustering the seed zones into hierarchical groups based on the matrix of transfer risks among zones. The resulting dendrogram and the respective geographic illustration of seed zones with similar transfer risk are presented in *Figure 43, p. 162*

Three main groups are apparent in the dendrogram presented in *Figure 43, p. 162*. A clear separation of the two regions is observable, indicating that the average relative transfer risk between the two regions is about 50% mismatch. Within Southwest Oregon, seed zone 512, inclu-

ZONE	270	491	492	501	502	511	512	661	662	671	672	673	674	675	681	682
270	0.00		SOUTHWEST OREGON							CENTRAL OREGON						
491	0.11	0.00														
492	0.26	0.15	0.00													
501	0.04	0.14	0.29	0.00												
502	0.18	0.07	0.08	0.21	0.00											
511	0.27	0.16	0.01	0.30	0.09	0.00										
512	0.52	0.43	0.30	0.55	0.37	0.29	0.00									
661	0.15	0.25	0.40	0.11	0.32	0.40	0.63	0.00								
662	0.19	0.29	0.43	0.15	0.35	0.44	0.66	0.04	0.00							
671	0.30	0.40	0.52	0.26	0.46	0.53	0.73	0.15	0.12	0.00						
672	0.55	0.63	0.72	0.52	0.67	0.73	0.86	0.43	0.40	0.29	0.00					
673	0.17	0.28	0.42	0.14	0.34	0.43	0.65	0.02	0.01	0.13	0.41	0.00				
674	0.29	0.39	0.52	0.26	0.45	0.53	0.72	0.15	0.11	0.00	0.30	0.13	0.00			
675	0.49	0.57	0.68	0.46	0.62	0.68	0.83	0.36	0.32	0.22	0.08	0.34	0.22	0.00		
681	0.53	0.61	0.71	0.51	0.66	0.72	0.85	0.41	0.38	0.27	0.02	0.39	0.27	0.06	0.00	
682	0.36	0.45	0.57	0.33	0.51	0.58	0.76	0.22	0.18	0.07	0.23	0.20	0.07	0.15	0.21	0.00
690	0.66	0.73	0.80	0.64	0.76	0.81	0.90	0.56	0.53	0.43	0.16	0.54	0.44	0.24	0.18	0.37
701	0.14	0.25	0.39	0.11	0.31	0.40	0.63	0.01	0.04	0.16	0.44	0.03	0.16	0.37	0.42	0.23
702	0.29	0.39	0.52	0.26	0.45	0.53	0.72	0.15	0.11	0.00	0.30	0.13	0.00	0.22	0.27	0.07
703	0.39	0.49	0.60	0.36	0.54	0.61	0.78	0.26	0.22	0.11	0.19	0.23	0.11	0.11	0.17	0.04
711	0.55	0.63	0.72	0.52	0.67	0.73	0.86	0.43	0.40	0.29	0.00	0.41	0.30	0.08	0.02	0.23
712	0.24	0.34	0.48	0.21	0.41	0.49	0.69	0.10	0.06	0.06	0.35	0.07	0.05	0.27	0.32	0.13
721	0.24	0.34	0.47	0.20	0.40	0.48	0.69	0.09	0.05	0.07	0.35	0.06	0.06	0.28	0.33	0.13
722	0.22	0.32	0.46	0.18	0.38	0.46	0.68	0.07	0.03	0.08	0.37	0.05	0.08	0.29	0.35	0.15
731	0.18	0.29	0.42	0.15	0.35	0.43	0.65	0.03	0.00	0.12	0.40	0.01	0.12	0.33	0.38	0.19
751	0.39	0.48	0.60	0.36	0.54	0.60	0.78	0.25	0.21	0.10	0.20	0.23	0.10	0.12	0.18	0.03
842	0.09	0.20	0.35	0.06	0.27	0.35	0.59	0.06	0.09	0.21	0.48	0.08	0.21	0.41	0.46	0.27
911	0.61	0.68	0.76	0.58	0.72	0.77	0.88	0.50	0.47	0.37	0.08	0.48	0.37	0.16	0.10	0.30
912	0.29	0.39	0.52	0.26	0.45	0.52	0.72	0.14	0.11	0.01	0.30	0.12	0.01	0.22	0.28	0.08
921	0.57	0.64	0.73	0.54	0.69	0.74	0.86	0.45	0.41	0.31	0.02	0.43	0.31	0.10	0.04	0.25
943	0.51	0.59	0.69	0.48	0.64	0.70	0.84	0.38	0.35	0.24	0.06	0.36	0.25	0.03	0.03	0.18
952	0.45	0.54	0.65	0.42	0.59	0.65	0.81	0.32	0.28	0.17	0.13	0.30	0.18	0.04	0.10	0.11

ZONE	690	701	702	703	711	712	721	722	731	751	842	911	912	921	943	952
690	0.00															
701	0.56	0.00														
702	0.44	0.16	0.00													
703	0.34	0.26	0.11	0.00												
711	0.16	0.44	0.29	0.19	0.00											
712	0.48	0.10	0.06	0.16	0.35	0.00										
721	0.49	0.09	0.06	0.17	0.35	0.01	0.00									
722	0.50	0.08	0.08	0.19	0.37	0.03	0.02	0.00								
731	0.53	0.04	0.12	0.22	0.40	0.06	0.05	0.04	0.00							
751	0.35	0.26	0.10	0.01	0.20	0.16	0.16	0.18	0.22	0.00						
842	0.60	0.05	0.21	0.31	0.48	0.15	0.14	0.13	0.09	0.30	0.00					
911	0.08	0.50	0.37	0.27	0.08	0.42	0.42	0.44	0.47	0.27	0.54	0.00				
912	0.44	0.15	0.01	0.12	0.30	0.05	0.06	0.07	0.11	0.11	0.20	0.37	0.00			
921	0.14	0.45	0.31	0.21	0.02	0.36	0.37	0.39	0.42	0.22	0.49	0.06	0.32	0.00		
943	0.21	0.39	0.24	0.14	0.05	0.30	0.30	0.32	0.35	0.14	0.43	0.13	0.25	0.07	0.00	
952	0.28	0.33	0.18	0.07	0.12	0.23	0.24	0.25	0.29	0.07	0.37	0.20	0.18	0.14	0.07	0.00

Table 40: *Relative seed-transfer risk among seed zones, based on original canonical scores from the trend surface (first allozyme variate): Total sampling area*

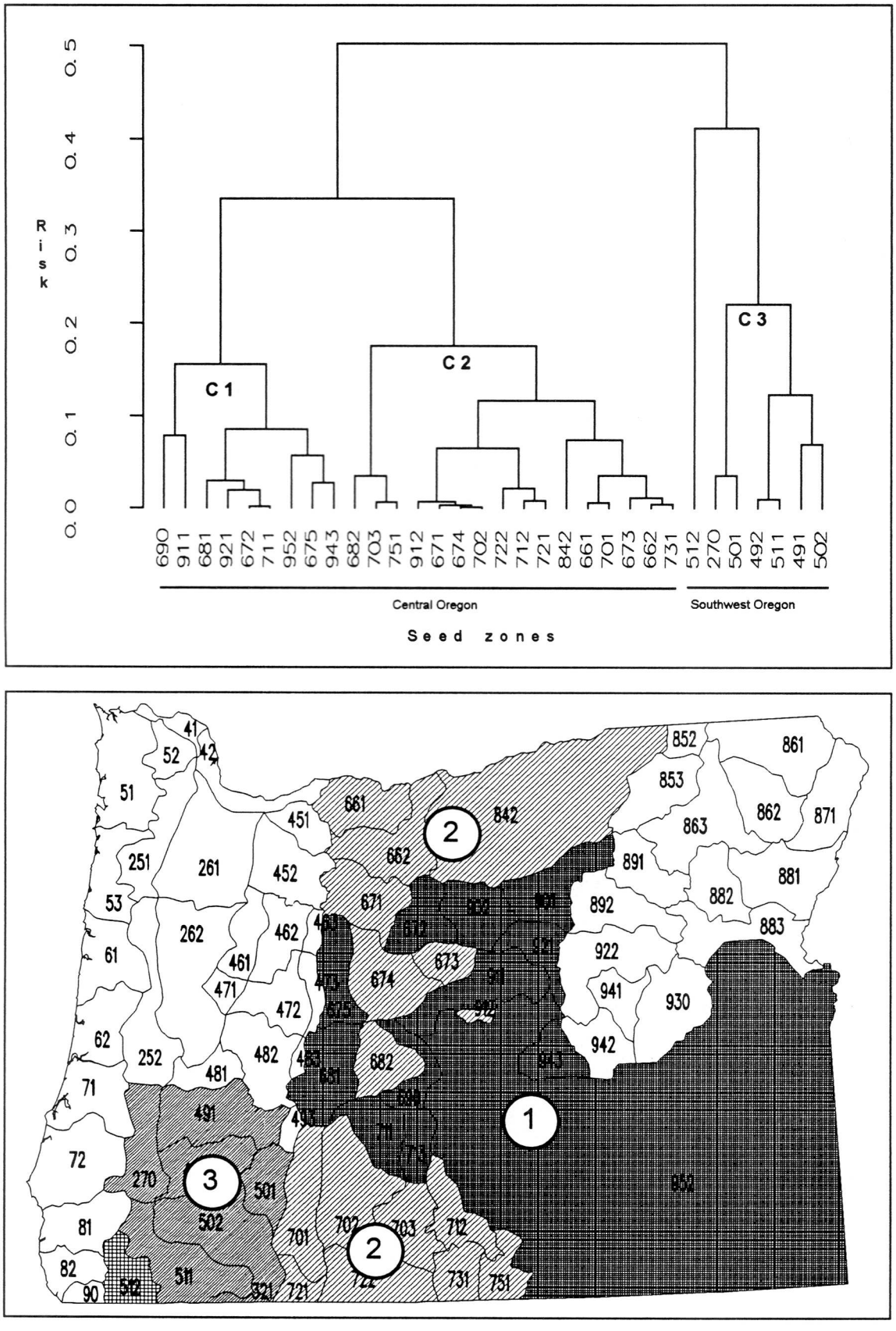


Figure 43: *Group formation of similar zones, based on relative seed-transfer risks among zones: Total sampling area. Upper: Dendrogram resulting from average linkage cluster analysis based on matrix of transfer risks among zones (see Table 40, p. 161). Lower: Mapped groups of seed zones with an average transfer risk of less than 30% within groups*

ding all individuals of zone 90, was quite distinct, exhibiting an average relative transfer risk of about 40% with all other southwestern zones. In addition, zones 270 and 501 differed by an average risk of about 22% from the zones in the center of this region. The zones of Central Oregon were clustered into two main groups. The central seed zones were clearly separated from the northern and the southern zones, the exception being zones 674, 673, 682 and 912. Transfer from the northern zones to the southern zones resulted in an average relative transfer risk of less than 20% while moving individuals from northern or southern zones to the center resulted in an average mismatch of 35%. Based on relative transfer risks, zones 661 and 662 could no longer be regarded as transition zones with penetrating maritime climate influence since relative transfer risk was higher with most of the southwestern zones than with most of the zones in Central Oregon.

An even better illustration of relative transfer risk in the sampling area is provided by the maps of relative seed transfer risk (*Figure 44, p 164*). In the upper half of *Figure 44,* distances between contours were scaled to represent a relative transfer risk of 20% between two neighboring isolines; 20% transfer risk equaled a difference of 0.36 units in canonical scores on the first variate between the neighboring contours. In Southwest Oregon, the central area, especially encompassing seed zones 511, 502, 492 and 491, was separated from the western as well as from the eastern parts. In contrast, no separation in a north-south direction was apparent. Contours did not match existing seed zone boundaries. Contours within Central Oregon showed a more complex zonation of transfer risk. Three isolated areas with especially low scores were apparent . The first such zone was made up mainly of the seed zones 711, 713, 690, 682 and 681. Another area with low scores encompassed mainly seed zones 912, 921 and parts of 952 and 943. A third and small area of negative scores was found in the southeastern part of zone 662 and the northern part of 672. Contours did not match seed zone boundaries. The three areas with low scores were surrounded by an extended zone of moderately negative scores, encompassing several southern (712, 703, 702, 711) as well as northern zones (675, 674, 673, 672, 671, 662, 902, 911). Finally, a small band at the northern limit of the area (661, 842, 901) was joined with a band situated in the highest altitudes near the Cascade crest and with a zone at the southern border of the sampling area (751, 731, 722, 702, 721, 701).

In the lower half of *Figure 44, p 164,* contours were scaled to represent a relative transfer risk of 30% between two neighboring contours. A mismatch of 30% was achieved when the difference in canonical scores between two locations on the trend surface exceeded a value of 0.55. In Southwest Oregon, a large central zone was separated from two small bands in the western and eastern part of the area, while no separation in the north-south direction was found. Central Oregon was separated into two main areas within which transfers were possible with a relative risk smaller than 30%. The central area was clearly separated from the northern and southern parts which were linked together by a band of the highest elevation sites near the Cascade crest.

Mean relative transfer risks among seed zones, estimated from the original canonical scores based on the Southwest Oregon sample only (*Table 35, p. 150*), are furnished in *Table 41, p. 165* and are illustrated in a dendrogram in *Figure 45, p. 165.* For the risk analysis, individuals of zone 42 were added to zone 661 of the Central Oregon sample while individuals of zone 90 were combined with the individuals of zone 512. Average relative transfer risks ranged from 2% to 50% of mismatch between the frequency distributions at source locations and potential planting sites. A rather high transfer risk was apparent for individuals from seed zone 492 which showed an average relative transfer risk of more than 30% compared with all the other zones. An average risk of more than 20% with all the other zones was also apparent for seed zone 501. Average transfer risks were below 15% among all the remaining zones.

Maps of contours from the predicted trend surface, scaled to represent 20% or 30% of relative transfer risks, are shown in *Figure 46, p. 166.* These maps clearly illustrate that seed transfer is much more complex within Southwest Oregon than what might be inferred from the results applying seed zone mean values. A highly complex pattern with changes in transfer risks of up to 90% over very short geographic distances and within the same seed zones was apparent.

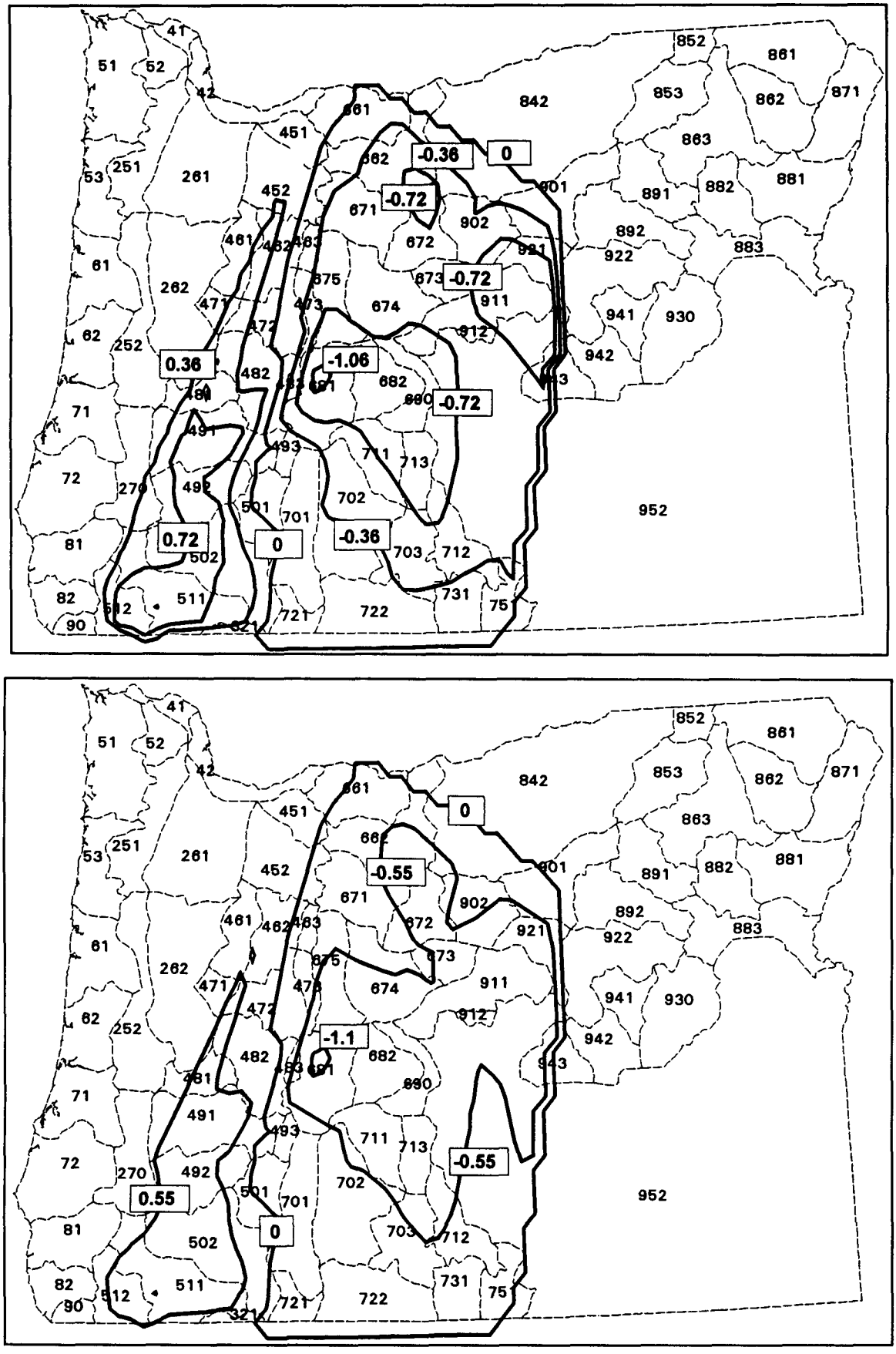


Figure 44: *Maps of relative seed-transfer risk, based on trend surface of predicted scores from first allozyme variate: Total sampling area. Distances between isolines are scaled to represent: Upper: 20% (x=0.36), Lower: 30% (x=0.55) relative transfer risk between isolines*

ZONE	270	491	492	501	502	511	512
270	0.00						
491	0.03	0.00					
492	0.35	0.32	0.00				
501	0.17	0.21	0.50	0.00			
502	0.02	0.01	0.33	0.20	0.00		
511	0.15	0.12	0.21	0.32	0.13	0.00	
512	0.13	0.10	0.23	0.30	0.11	0.02	0.00

Table 41: Relative seed-transfer risk among seed zones, based on canonical scores from the first allozyme variate: Southwest Oregon

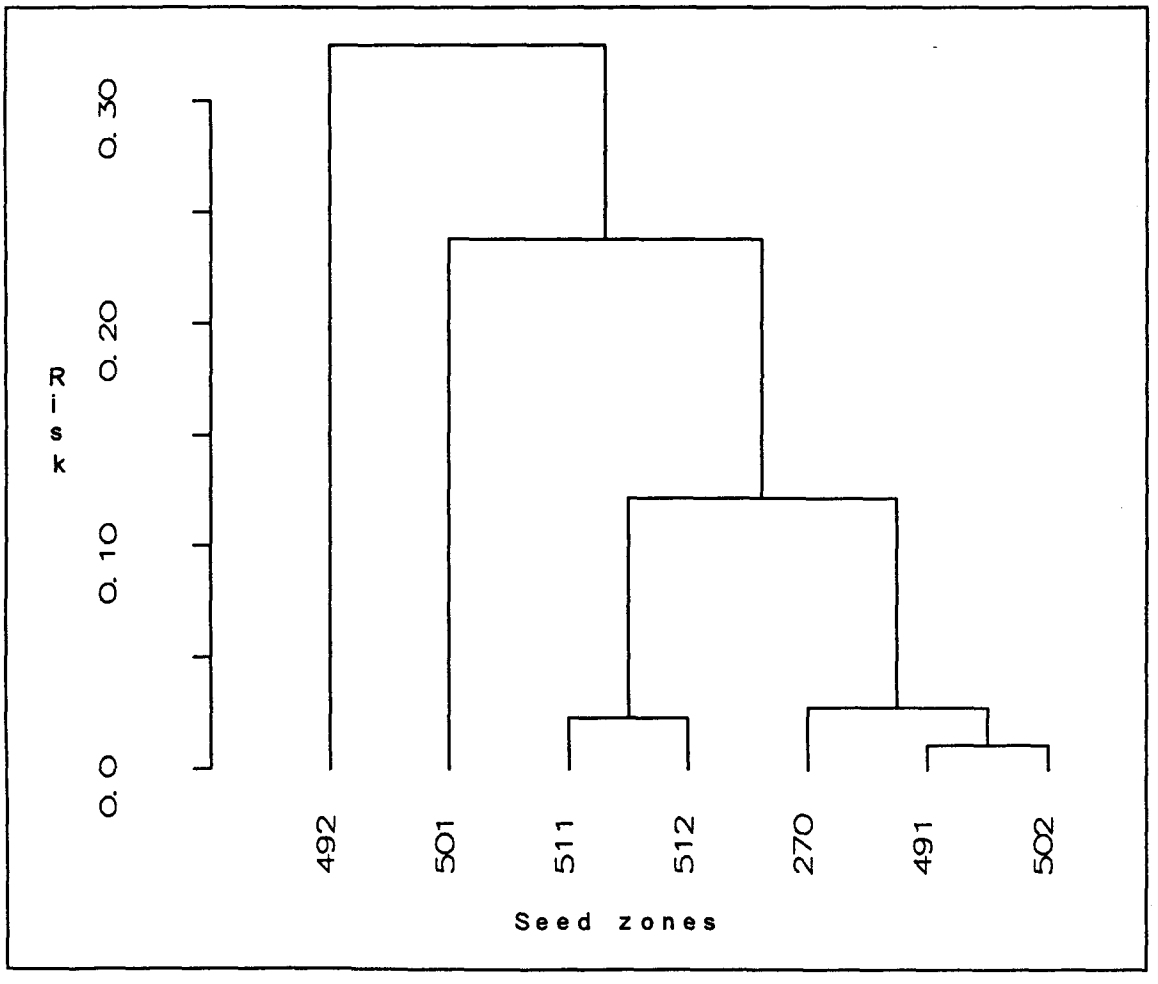


Figure 45: Group formation of similar zones, based on relative seed-transfer risks among zones: Southwest Oregon. Dendrogram resulting from average linkage cluster analysis based on matrix of transfer risks among zones (see Table 41, p. 165)

Steep gradients of changes in transfer risk over very short distances were especially apparent in the western part of the region (zones 512, 502 and 270) and in the northeast (zones 491 and 492). The geographic patterns of the contours seemed to parallel the boundaries of the seed zones to a certain extent, suggesting that the present seed zones basically reflect the adaptive patterns. However, two to three isolines representing 60% to 90% of transfer risk were observed within most of the seed zones, indicating that the present seed zones are most likely too large to effectively guard against insufficient match between source location and planting site.

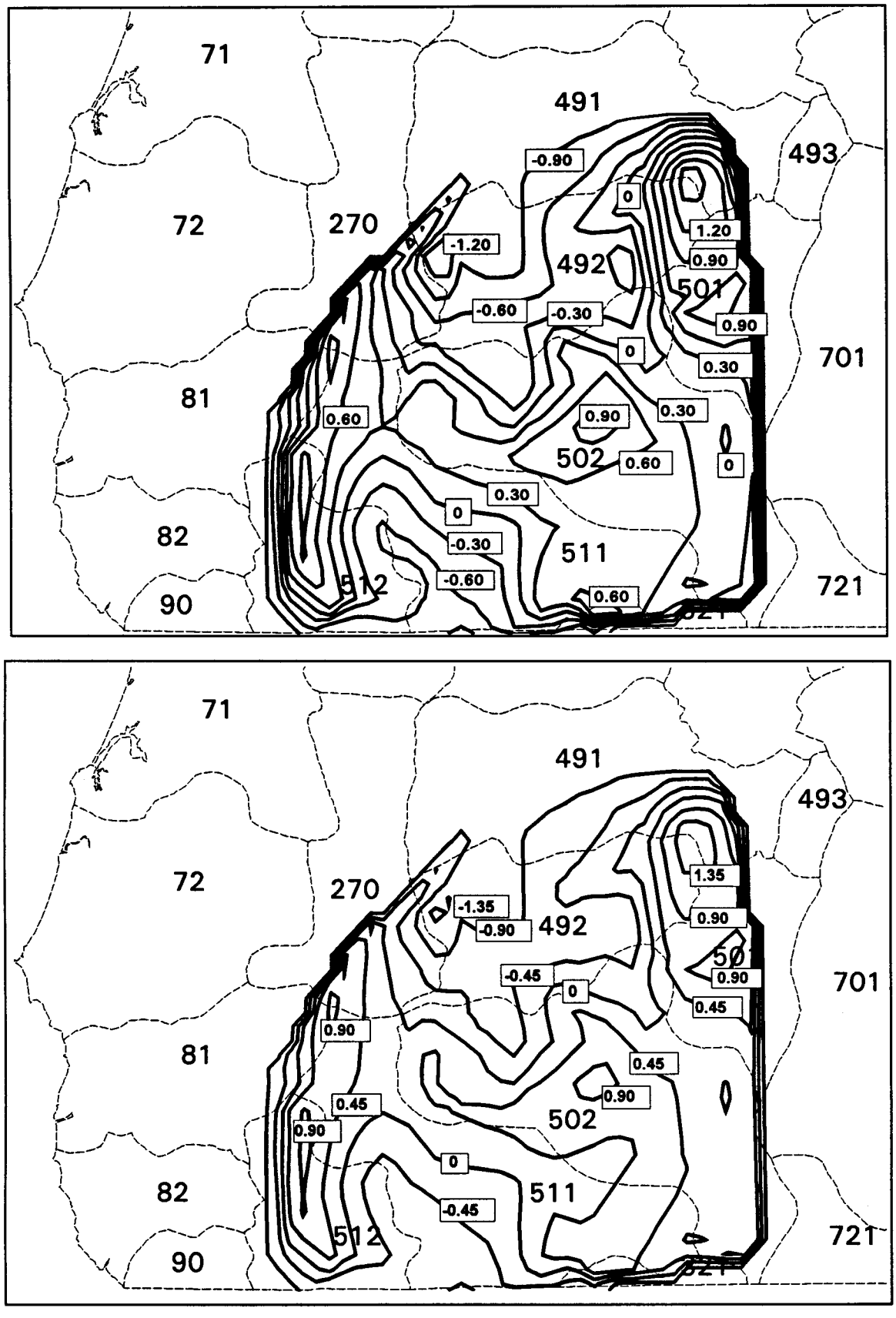


***Figure 46**: Maps of relative seed-transfer risk, based on trend surface of predicted scores from first allozyme variate: Southwest Oregon. Distances between isolines are scaled to represent: Upper: 20% (x=0.30), Lower: 30% (x=0.45) relative transfer risk between isolines*

ZONE	661	662	671	672	673	674	675	681	682	690	701	702	703
661	0.00												
662	0.02	0.00											
671	0.28	0.30	0.00										
672	0.32	0.34	0.04	0.00									
673	0.48	0.50	0.22	0.18	0.00								
674	0.32	0.34	0.04	0.01	0.18	0.00							
675	0.51	0.52	0.25	0.22	0.04	0.21	0.00						
681	0.53	0.54	0.28	0.24	0.06	0.24	0.03	0.00					
682	0.24	0.26	0.05	0.09	0.27	0.09	0.30	0.32	0.00				
690	0.54	0.55	0.29	0.25	0.07	0.25	0.04	0.01	0.34	0.00			
701	0.02	0.04	0.27	0.30	0.46	0.31	0.49	0.51	0.22	0.52	0.00		
702	0.19	0.21	0.10	0.14	0.32	0.15	0.35	0.37	0.05	0.38	0.17	0.00	
703	0.15	0.17	0.14	0.18	0.35	0.18	0.38	0.41	0.09	0.42	0.13	0.04	0.00
711	0.41	0.43	0.14	0.10	0.08	0.09	0.12	0.15	0.19	0.16	0.39	0.24	0.28
712	0.15	0.17	0.13	0.17	0.35	0.18	0.38	0.40	0.09	0.41	0.14	0.03	0.01
721	0.01	0.03	0.28	0.31	0.47	0.32	0.50	0.52	0.23	0.53	0.01	0.18	0.14
722	0.10	0.12	0.19	0.22	0.39	0.23	0.42	0.45	0.14	0.46	0.08	0.09	0.05
731	0.28	0.30	0.00	0.04	0.22	0.05	0.25	0.28	0.05	0.29	0.27	0.10	0.14
751	0.37	0.39	0.09	0.06	0.13	0.05	0.16	0.19	0.14	0.20	0.35	0.20	0.23
842	0.16	0.18	0.13	0.17	0.34	0.17	0.37	0.40	0.08	0.41	0.14	0.03	0.01
911	0.45	0.47	0.19	0.15	0.03	0.14	0.07	0.10	0.23	0.11	0.44	0.28	0.32
912	0.03	0.05	0.26	0.29	0.46	0.30	0.49	0.51	0.21	0.52	0.01	0.16	0.12
921	0.45	0.46	0.18	0.15	0.04	0.14	0.07	0.10	0.23	0.11	0.43	0.28	0.32
943	0.53	0.55	0.28	0.25	0.07	0.24	0.03	0.00	0.33	0.01	0.52	0.38	0.41
952	0.40	0.42	0.13	0.09	0.09	0.08	0.13	0.16	0.17	0.17	0.38	0.23	0.26

ZONE	703	711	712	721	722	731	751	842	911	912	921	943	952
703	0.00												
711	0.28	0.00											
712	0.01	0.27	0.00										
721	0.14	0.40	0.15	0.00									
722	0.05	0.32	0.05	0.09	0.00								
731	0.14	0.14	0.13	0.27	0.18	0.00							
751	0.23	0.04	0.23	0.36	0.28	0.10	0.00						
842	0.01	0.26	0.00	0.15	0.06	0.13	0.22	0.00					
911	0.32	0.05	0.31	0.44	0.36	0.19	0.09	0.31	0.00				
912	0.12	0.38	0.13	0.02	0.07	0.26	0.35	0.13	0.43	0.00			
921	0.32	0.05	0.31	0.44	0.36	0.18	0.09	0.31	0.00	0.42	0.00		
943	0.41	0.15	0.40	0.52	0.45	0.28	0.19	0.40	0.10	0.51	0.10	0.00	
952	0.26	0.01	0.26	0.39	0.31	0.13	0.03	0.25	0.06	0.37	0.06	0.16	0.00

Table 42: *Relative seed-transfer risk among seed zones, based on canonical scores from first allozyme variate: Central Oregon*

Mean relative transfer risks among the seed zones in Central Oregon, estimated from original canonical scores based on the Central Oregon sample only (*Table 37, p. 154*), are presented in *Table 42*. A dendrogram showing the hierarchy of seed zones based on the matrix of relative transfer risk among zones and the geographic illustration of the main groups are furnished in *Figure 47, p. 168*. Average relative transfer risk ranged from 1% to 54% mismatch. The dendrogram revealed 3 major groups of seed zones with an average transfer risk of less than 20% within groups. A clear separation of the central zones from both northern as well a southern zones was found. A transfer of material from northern or southern zones to planting sites in the center of the area resulted in an average mismatch of about 35%. In contrast, transfers of material from the northern to the southern zones showed an average mismatch of only 15%. The central zones were again divided into two groups; seed zones 671, 672, 674 and 682 exhibited an average risk of 20% with the remaining zones in the center.

Maps of contours on the predicted canonical trend surface, representing relative transfer risk of 20% or 30% between two neighboring isolines, are shown in *Figure 48, p. 170*. Compared with

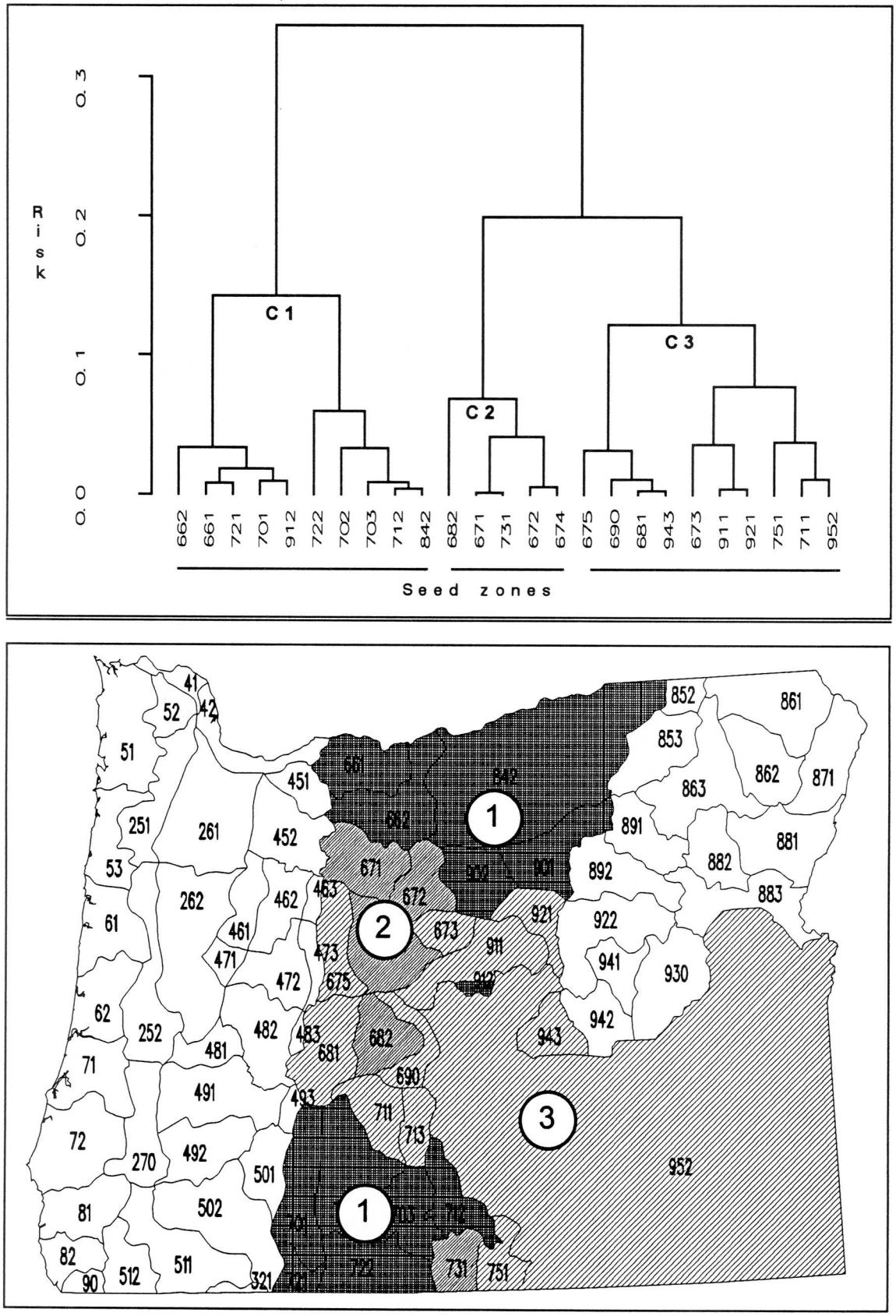


Figure 47: *Group formation of similar zones, based on relative seed-transfer risks among zones: Central Oregon. Upper: Dendrogram resulting from average linkage cluster analysis based on matrix of transfer risks among zones (see Table 42, p. 167). Lower: Mapped groups of seed zones with an average transfer risk of less than 20% within groups.*

Southwest Oregon, contours scaled to 30% of risk among each pair of contours delineated rather large areas which normally consisted of several to many seed zones, indicating that material may be transferred within Central Oregon over rather large geographic distances without surpassing a risk of 30%. Steeper gradients of transfer risk were only found in the highest elevations of the Cascade Range. Occasionally, a risk of more than 30% was found within short geographic distances, for example in seed zones 661, 662, 911, 711 and 722. Contours did not parallel the existing seed zone boundaries.

6.5 Discussion

Application of cluster analysis on matrices of genetic dissimilarity among individuals, assigning each tree to a subset of samples with similar genotypic scores, produced complex geographic patterns of multilocus genotypic variation (*Figure 28, p. 128*). Genetically similar individuals occurred all over the sampling area. No ecotypic distribution of similar individuals was apparent, not even on a very local scale. One exception, however, was found in the southwestern corner of the sampling area. Individuals in this part of southwestern Oregon (seed zone 90) and northwestern California, were characterized by a distinct combination of genotypic scores which did not occur elsewhere. The distinct genetic structure in this small part of the area has already been observed and discussed in *sections 4* and *5*.

It must be emphasized however, that patterns of multilocus genotypic variation, illustrated as the geographic distribution of individual group memberships, are not truly multivariate patterns. Cluster analysis, used for assigning each individual to a subset with a similar genotypic combination, is based on a dissimilarity matrix which is calculated from the original scores of all 71 allozyme variables. All measures of dissimilarity, including Euclidean distance, assume that the variables are un-correlated within the clusters. The covariance structure of the data is thus not reflected in the dissimilarity space between the entities. In other words, intercorrelations among the variables are not accounted for, and the orientations of the differences in multivariate space are not reflected in the resulting cluster solution. Moreover, traditional genetic distance measures (GREGORIUS, 1974; NEI, 1978) weight all alleles equally, even though not all alleles exhibit adaptive patterns as has been demonstrated in *section 5*. Although differences in frequencies are weighted proportionally in the calculation of Euclidean distance, intercorrelations and directions of differences are likewise not reflected in multivariate dissimilarity space. Hence, eventual linkage disequilibria cannot be portrayed by the cluster solution. In addition, variation in multilocus variable space may be continuous and a classification in discrete, discontinuous subsets might only give a crude picture of the real situation. Moreover, deciding on the number of subsets present in the data is a common problem of cluster analysis. Numbers of subsets, however, will influence results of subsequent discriminant analyses. Variation may also be highly nonlinear and the pattern may therefore not be readily detectable in a cluster analysis.

Associations between individual multilocus genotypic structures and habitat conditions, inferred by descriptive discriminant analysis, partitioning the variance in habitat conditions into among and within-group proportions, were statistically significant but rather weak (*Table 25, p. 130*). Estimated associations differed according to the dissimilarity matrices which were used to assign the samples to homogeneous subsets of genotypes. As expected, results based on cluster solutions using Euclidean distance, giving more weight to the highly differentiated alleles, revealed associations 1.5 to 3 times higher than the associations estimated on the basis of the distance measures which do not weight the different alleles. Hence, results involving cluster solutions based on Euclidean distance seem to portray the real situation better than the results based on traditional genetic distance measures. Nonetheless, since cluster analysis is clearly limited in portraying truly multivariate patterns, all estimated associations, even the estimates based on Euclidean distance, most likely underestimate the real multivariate relationships among genetic and environmental variation. Estimates of 7% to 13% may thus be regarded as the lower limits, the real associations most likely being higher than these estimates.

Apart from these limitations regarding the estimated amounts of association, results of discriminant analyses are clearly in line with the conclusions already reached in *sections 4 and 5*.

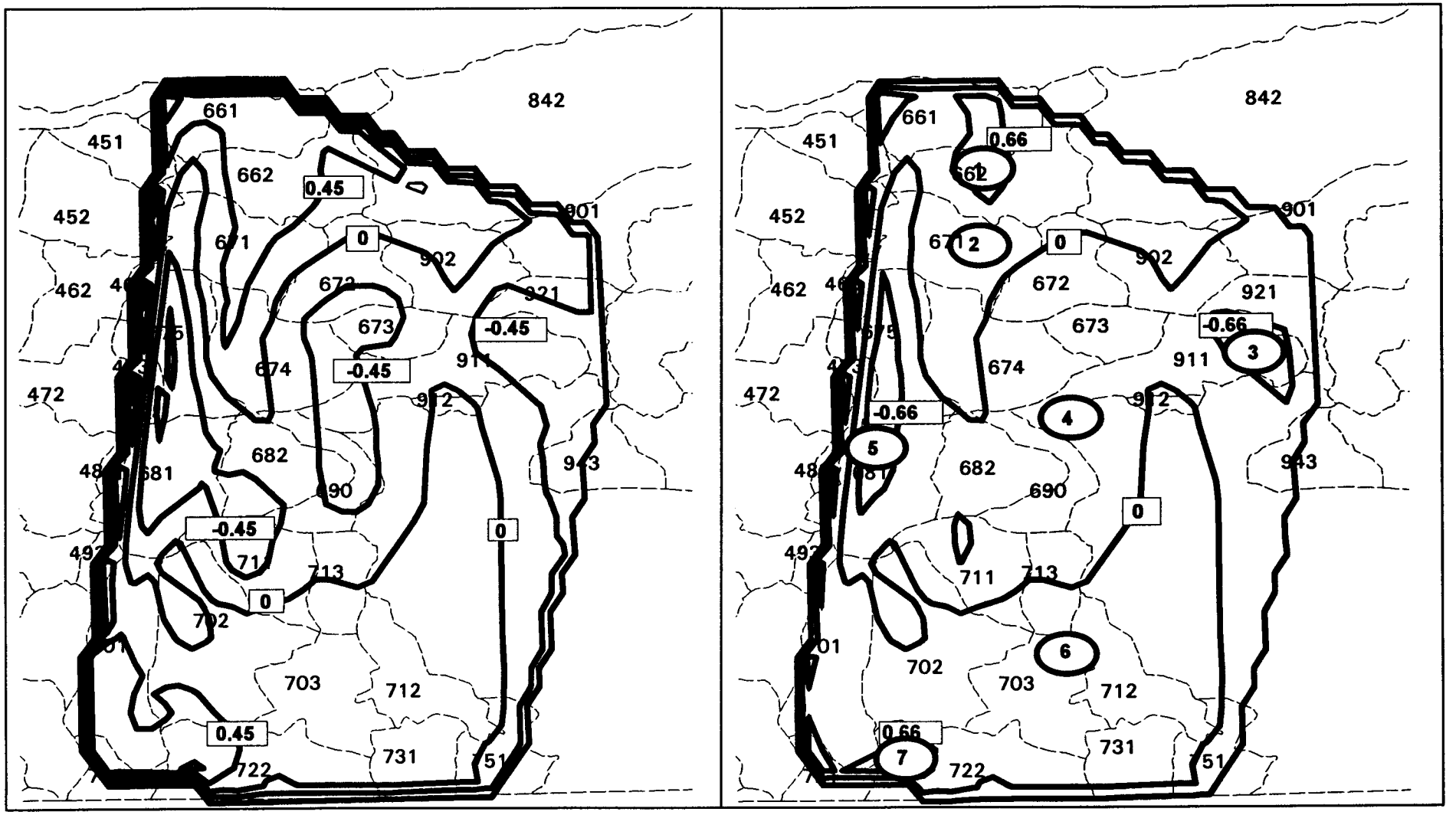


Figure 48: *Maps of relative seed-transfer risk, based on trend surface of predicted scores from first allozyme variate: Central Oregon. Distances between the isolines are scaled to represent: Left: 20% (x=0.45), right: 30% (x=0.66) relative transfer risk between isolines*

Results confirm that the individual genotypic structure is primarily associated with the temperature condition at source locations while no significant associations with moisture conditions of the site seem to exist, except for Southwest Oregon. The prime importance of temperature as selective agent has already been postulated in the analysis of differentiation patterns; it has been observed in matrix comparisons and in the multinomial response models. Interestingly, moisture seems to have a slightly more pronounced effect on genotypic structure in Southwest Oregon than in Central Oregon. Although associations are not significant, the rather high structure coefficient in the cluster solution using Euclidean distance suggests that the high gradients in moisture conditions found within this area may be partly reflected in the genotypic structure. Generally, however, associations between genotypic structure and habitat conditions seem to be weaker within Southwest Oregon than within Central Oregon. It has already been shown in *sections 4 and 5* that differentiation is less pronounced within Southwest Oregon compared to differentiation within Central Oregon. In addition, no associations among single alleles and climate factors were found in the multinomial response models performed on the Southwest Oregon sample. These results are now confirmed by low and not significant proportions of habitat variation which were related to differences among genotypic groups (*Table 25, p. 130*). Lack of significant associations is most likely a consequence of a less extreme selection pressure due to warmer winters and lower temperature extremes in this area.

The distribution of ponderosa pine in the sampling area can be grouped in 8 to 10 different habitats characterized by specific combinations of temperature and moisture conditions (*Figure 29, p. 132*). If these habitats are plotted by location, a fairly distinct geographic pattern becomes apparent (*Figure 30, p. 133*). On an area-wide scale, habitat conditions show an ecotypic pattern of three distinct habitat types within Southwest Oregon while Central Oregon shows a more complex, locally variable pattern in the high elevations of the Cascade Range and more ecotypic distribution in the high desert area to the east.

Describing differences in multilocus genotypic structures relative to these different habitat types revealed associations which were much stronger than those found in the discriminant analyses on groups of genotypes (*Table 27, p. 135*). Canonical discriminant analyses assigned proportions of 20% to 40% of variation in the first canonical variates to differences in habitat conditions. The canonical variates, which explained most of the among-group variance, were dominated by several alleles at many loci, suggesting that many alleles contribute in small proportions and in an additive fashion to an overall adaptive multilocus genotypic frequency distribution.

The geographic distribution of habitat conditions should depict the true multivariate pattern of climate variation in the area. Principal components are orthogonal transformations of the original climate variables. Hence, they are un-correlated with each other and Euclidean distances calculated from climate *Factor 1* and *Factor 2* should reflect the multivariate variance-covariance structure of the original climate data set. Accordingly, results are expected to portray associations among genotypic structure and habitat conditions better than the former analysis. Resulting associations were in fact higher than the results based on subsets of similar genotypes, supporting our speculation that these estimates may indeed be regarded as lower limits.

Nonetheless, the estimated associations of 20% to 40% revealed by discriminant analysis on groups of habitats, must also be interpreted with caution. Comparisons between results of classification and cross-validation provided strong evidence that the discriminant axes were not very robust. Due to insufficient sample size and violations of assumptions, stability of the canonical functions was insufficient in all analyses.

Classification may be used to indirectly infer the robustness and stability of the derived canonical functions and to validate the reliability of results from the analysis. The premise of this validation approach is that unstable results of classification indirectly infer that the canonical functions are also unstable. Unstable canonical functions and classification criteria may result from violations of one or several assumptions or from a sample size too small to obtain accurate and precise estimates of means and dispersions. Stability of classification results may be validated using resampling procedures such as split sample validation or cross-validation procedures. Cross-validation is a procedure where one sample at a time is omitted from the data set before the canonical functions are derived. The omitted sample is then reclassified using the derived

discriminant functions. The procedure is sequentially repeated for each sample and classification accuracy is thus evaluated. Biased classification rates resulting from reclassification of the same samples that were used to derive the canonical functions may thus be overcome. A substantial reduction in the classification accuracy suggests that the estimated means and dispersions are not very reliable and that a larger sample size would be required for more robust estimates and stable classification results. Robust estimates of the functions depend on the homogeneity of group dispersions (*i.e.*, variance-covariance matrices), the multivariate normality of data, low intercorrelations among the variables and a large sample size relative to dimensionality. Violations of assumptions may result in unstable estimates of the canonical functions. In this case, the among-group to within-group variance may be distorted and patterns are at best suggestive and descriptive. However, the impact of these violations on the results remains totally unknown (WILLIAMS, 1981). While there is evidence that certain of the assumptions can be violated moderately, without large distortions and changes in classification results (KLECKA, 1975; HARRIS, 1975; LACHENBRUCH, 1975), equality of covariance matrices is considered to be important for stable estimates of the canonical functions (WILLIAMS, 1981).

In all three analyses, cross-validated results indicated unstable estimates of the canonical functions. Variance-covariance matrices of the genetic data *i.e.* group dispersions were in fact not homogeneous (within-group variance-covariance structure were not the same for all groups). A multivariate chi-square test as proposed by MORRISON (1976) was highly significant for all three analyses. Since quadratic functions, which are commonly used when the within-group variances and covariances are heterogeneous, have limited use in allozyme data sets because some of the loci are fixed in one or more groups and not in others (WESTFALL and CONKLE, 1992), linear functions were used in spite of the unequal variance-covariance matrices. Results of classification and cross-validation may therefore not be very reliable. Robust estimates of canonical functions depend largely on a sufficient number of samples relative to the number of discriminating variables. According to WILLIAMS and TITUS (1988), each group should contain at least three times as many samples as the discriminating variables to insure stable estimates. Hence, our sample to variable ratio was clearly insufficient to allow for robust estimates of the canonical functions and therefore, reported results must be considered as exploratory and descriptive only.

Nevertheless - with the exception of Southwest Oregon - the results obtained were highly in accordance with the conclusions already reached from single locus analyses. Alleles which dominated the first canonical variate of total sampling area were identical with alleles that exhibited high differentiation among the regions. Moreover, alleles with a correlation coefficient higher than 0.20 with the first variate were all characterized by a spatial pattern that significantly deviated from a random spatial arrangement. All alleles with structure coefficients higher than 0.30 on axis 1 had already shown a significant association with temperature in the univariate multinomial response models. Based on mean scores on discriminant axis 1, seed zones could be clearly assigned to their respective regions. Southwestern zones all had negative mean scores on this axis while zones from Central Oregon were all positive. Zone 661 was again characterized by an intermediate mean score, indicating that this zone may indeed be regarded as a transition zone between the two regions, having a genetically distinct structure. The clear separation on axis 1 reflects either a real association of genotypic structures with the different temperature regimes found in the two areas or again is a result of indirect effects linked to the two regions (different evolutionary history). Indirect effects may also be responsible for high structure coefficients of certain alleles which in fact are not associated with habitat conditions but rather show a high correlation with axis 1 due to such indirect effects. The question of indirect effects will be addressed separately in more detail below.

Results regarding associations among genotypic structure and moisture conditions of the site are also in accordance with the former conclusions. Although a slight tendency of associations between the mean scores on axis 2 may be inferred from the plotted scores on both discriminating axis, non significance of axis 2 and inconsistent results with many exceptions clearly indicate that associations are at best only weak.

Discriminant analysis, performed on 10 subsets within Southwest Oregon only, produced a highly significant first variate which assigned 40% of variation in genotypic scores of 32 allozyme variables to group differences (*Table 27, p. 135*). Results from cross-validation how-

ever suggest that the first canonical function is not robust, thus producing spurious results of associations. Classification accuracy was reduced from 40% (higher than by chance) to only 3% to 5% correctly assigned individuals in crossvalidation. Inspection of alleles which exhibited high redundancies with canonical variate 1, also indicates that variate 1 is most likely not robust. None of the alleles which loaded high on this variate had shown either a spatial geographic structure or any association with climate in the multinomial response models. Likewise, plotted canonical scores did not furnish any meaningful patterns. Hence, there is strong evidence that assumptions for discriminant analysis seem to have been violated (due to insufficient samples per subset) and results of discriminant analysis are most likely not only spurious but also substantially overestimate the real associations between genotypic structure and climate conditions.

Concordant results, however, were observed for the analysis on subsets within Central Oregon. Associations with habitat types were estimated as 20%. All alleles having a correlation higher than 0.20 with variate 1 were alleles which had already shown an association with temperature in the multinomial response models. However, plots of mean scores on the first and second canonical axis revealed only weak trends with habitat conditions, even with temperature as the prime factor of influence. Considering the low level of significance and the weak associations that may be inferred from the plot of mean scores by seed zones, the low association of 20% seems to be a rather realistic estimate of the relationship.

Direct multivariate associations between the dependent allozyme data set and an independent climate model may be obtained using canonical trend surface analysis. Although canonical correlation analysis has been successfully applied to ecological problems (GITTINS, 1985), only recently have applications to population genetics been reported (WARTENBERG, 1985; CONKLE and WESTFALL, 1984; WESTFALL and CONKLE, 1992; MERKLE et al., 1988). With the exception of sensitivity to outliers, canonical correlation analysis is less sensitive to errors than principal component analysis, and errors that occur usually reside in the residual variation (WESTFALL and CONKLE, 1992). In addition, canonical correlation is superior to principal component analysis for detecting irregular geographic structures (WARTENBERG, 1985). The technique is truly multivariate and more straightforward than the indirect approach using principal component and regression analysis. It is also superior to the categorization approach based on cluster and discriminant analysis. However canonical trend surface analysis also has some shortcomings, (SMITH,1981, JOHNSON, 1981). A minimal sample size of three times the sum of dependent and independent variables is required for reliable results; if the sample to variable ratio gets smaller, canonical correlation coefficients become inflated and statistical significance may be assured even if the results do not have biological significance. Moreover, robustness depends on normality of the data. Finally, several parameters must be assessed in order to get a complete and accurate picture of the relationships. Interpretation of results may thus be complex. Canonical correlation analysis was described in detail in *section 3.6.4.4.*

With canonical trend surface analysis over the total sampling area, a rather high proportion of the variance (33.5 %) in genotypic scores of 41 alleles was related to a second order model of climate factors (*Table 33, p. 145*). The trend surface, representing the multivariate adaptive genotypic variance, proved to be highly significant with adequate fit and normally distributed and unbiased residuals. Moreover, results were meaningful and in accordance with results from earlier analyses. For example, the estimated amount of association was strikingly in accordance with the association revealed by cluster and discriminant analysis based on subsets of similar habitat conditions. Alleles which contributed most to the trend surface were the same as those which were already found to be associated with climate in single-locus analyses and which loaded high on the first discriminant axis of discriminant analysis on subsets of habitats.

With a number of samples that was more than 9 times larger than the number of variables, the canonical functions and weights are expected to be robust and reliable. Relationships should be truly multivariate. In contrast to the indirect approach with cluster and discriminant analysis, the assumptions for canonical correlation analysis were met and the resulting canonical trend surface should thus describe the adaptive pattern of multilocus frequency distribution as close to reality as possible.

The canonical model clearly demonstrated that the variation in multilocus genotypic frequencies is moderately related to the habitat conditions at source locations. Although individual alleles only occasionally exceeded a correlation of 0.25 with the canonical model, indicating that only small proportions of variation in any one allele are associated with climate conditions, the aggregate pattern was relatively strong. Thus, the loci and their alleles behave much in the way one would expect of quantitative trait loci. On the level of the gene, the trait is broken down into many genes with each gene contributing only very little to the expression of the character (LEWONTIN, 1984). While the small individual contributions may be difficult to detect statistically, the aggregate pattern may be strong and meaningful. The present study thus provides further evidence that allozyme markers behave much in the same way as quantitative genes, each marker contributing only small proportions to an overall pattern of alllozyme variation which is closely associated with climate, suggesting that this allozyme variate is contributing to adaptation. Similar results were reported by WESTFALL and CONKLE (1992), analyzing variation patterns of different species in southwestern Oregon and California.

Not all alleles were equally related with this overall adaptive pattern. Several loci and alleles seemed to be completely unrelated (42% of the 71 allelic markers), others showed weak relationships, while a few alleles were moderately correlated with the pattern. Of the 71 allelic variables, 41 (58%) showed a minimal correlation with the pattern of at least 10% while 22 alleles (31%) revealed a correlation of at least 20% and 8 alleles (11%) of at least 30%. The highest contribution to the pattern was found for the two alleles at locus *Idh* with a structure coefficient of 50% or a proportion of 25% of shared variance with the trend surface. Most of the loci which showed an association with the adaptive pattern code for enzymes with an important physiological role in glycolysis or the Krebs cycle (*Idh*, *Mdh*, *G6p*, *Lap*, *Pgm*). The same enzymes were also reported to pattern in all four species which were investigated by WESTFALL and CONKLE (1992).

With one exception, all alleles with a structure coefficient higher than 0.20 with the trend surface were characterized by a spatial pattern that had deviated from random spatial arrangement in spatial autocorrelation analysis (*section 5*), the exception being locus *Mdh-3*. In contrast, *Lap2-4*, *Got1-3*, *Acp1-4*, *Aco1-3* and *Aco1-4* had shown a spatial pattern but showed no relationship with the adaptive trend surface. In addition, all alleles with a correlation higher than 0.20 with the adaptive trend surface were already significantly associated with temperature in the single locus analyses using multinomial response models. Our study thus provides further evidence that multivariate statistical techniques may be successfully applied to a multilocus data set, revealing in one single analysis all the important single locus patterns and associations, and at the same time completing the picture at the multilocus level, producing information about the aggregate pattern of multilocus variance.

The magnitude of multivariate associations of allozyme genotypes with climate in western Oregon ponderosa pine is similar to that observed for four conifers (also based on canonical trend surface models) in the mixed conifer zone of California's Sierra Nevada (WESTFALL and CONKLE, 1992). The amount of variation described by a geographic model, involving latitude, longitude and elevation, was 28% for White-fir (*Abies concolor Lindl.*), 44% for Sugar pine (*Pinus lambertiana Dougl.*), 40% for ponderosa pine (*Pinus ponderosa Laws.*) and 63% for Douglas-fir (*Pseudotsuga menziesii Franco*). The higher proportion of allozyme variance explained for ponderosa pine in the Sierra Nevada versus western Oregon is most likely due to the larger sample in the Sierra Nevada region, extending over 700 km in the north-south direction (and thus covering a wider variety in habitat conditions) compared to the north-south extension of 420 km in the present study. The White-fir sample, on the other hand, extended only over 325 km in the north-south direction, thus a weaker patterning might be expected. However, patterning of ponderosa pine in Oregon seems to be weaker than the patterns reported for Douglas-fir and Sugar pine which extended over an area similar to that in western Oregon. Several reasons may be envisioned for the weaker association found in Oregon. Adaptive patterns may be weaker due to the shorter period of time since immigration of ponderosa pine into the area, and thus a shorter amount of time for adaptation. Alternatively, habitat conditions may be less variable and selection pressure less stringent than in the Sierra Nevada of California. Especially in southwestern Oregon the habitat conditions are mild and adaptation to temperature is less pronounced (see below). Alternatively, species may differ in the amounts of genetic variation due to a different evolutionary past; associations may thus differ since selection may

only act on the genetic variation which is present. Finally, a difference in methodology may have produced the differing amounts of associations. Instead of a geographic model which was used by WESTFALL and CONKLE (1992), our results are based directly on climate conditions at source locations. Since climate data have been estimated by local regression procedures, a portion of the variability in habitat conditions may have been removed. A similar amount of association of multilocus frequencies with environmental variation was reported by GURIES (1984). In western white pine, *Pinus monticola Dougl.*, the first principal component, representing 37% of total genotypic variance at 12 loci, was highly correlated with latitude and elevation, both variables, of course, being highly redundant with climate conditions. According to YEH et al. (1985) who reported on variation in Lodgpole pine, *Pinus contorta Dougl. ex. Loud.*, in the Yukon and British Columbia, 28% of the variation in the first discriminant function, representing 20 loci, paralleled a strong north-south geographic pattern, this pattern obviously mainly representing variations in temperature.

Multilocus genotypic frequencies were primarily related to temperature (*Table 33, p. 145*). Climate *Factor 1*, representing the temperature regime at source locations, clearly dominated the model (canonical variate 1 representing 97 % of the variance in *Factor 1*). Associations with moisture, on the other hand, were not significant and weak in magnitude. In all former analyses, without exception, temperature has been identified as the major factor influencing genetic structure. Consequently, the results clearly lead to the conclusion that temperature at the source location seems to be the most important environmental factor responsible for the observed adaptive pattern of multilocus frequency distributions found in the area. The important role of temperature may best be seen when comparing the response surfaces of predicted allozyme scores of the first canonical variates (*Figure 37, p.* 148, *Figure 42, p. 157*) with the variation patterns of temperature, expressed as climate *Factor 1*. The two patterns of variation are in fact strikingly similar, not only as regards their similarity in the overall trend but clearly also on a very local scale.

Our results are highly comparable with results that were recently reported by SORENSEN (1994). SORENSEN analyzed several quantitative traits of one- to three-year old seedlings, using the same families which were included in the Central Oregon sample of the present isozyme investigation. Applying principal component and regression techniques, he found a strong association between seedling traits and elevation and concluded that genetic differentiation in seedling vigor was much stronger across temperature gradients than across moisture gradients. These results indicate that allozyme data can provide similar information about adaptive patterns. In *section 7*, patterns of variation revealed by seedling quantitative traits (applying the same statistical procedure as for the allozyme analysis) will be analyzed and compared with the patterns found for the isozyme markers. A detailed comparison of variation patterns from the two sources of information will thus be possible.

In *section 1* we had stated the following 3 hypotheses to be tested:

Geographic variation in multilocus allozyme genotypes is the result of:

1) ***recent adaptation*** *to current environments, caused by natural selection*

2) ***a different evolutionary history,*** *caused by migration from different refugia populations*

3) ***random processes,*** *caused by genetic drift among populations*

The fact that there is genetic differentiation between Southwest and Central Oregon is consistent with all 3 hypotheses. Associations of patterns of multilocus allozyme frequencies with environmental variation within the two regions, is potentially consistent with hypothesis 1 and 2. Complex patterns of allozyme variation associated with complex environmental variation, however, are unlikely under hypothesis 2. This is especially so if patterns are repeated in different places as seems to be the case for ponderosa pine in Oregon. Moreover, the amount of variation in multilocus allozyme frequencies that can be explained by spatial variation alone (*i.e.* which is unrelated to environmental variation) is very small which clearly argues against migration. In addition, close asscoations of allozyme patterns with patterns seen in quantitative traits, as observed in Central Oregon, seem to be strong evidence for hypothesis 1 since patterns of quantitative traits generally strongly reflect adaptation.We thus conclude that the observed pat-

terns of mutlilocus allozyme frequencies are most likely the result of adaptation to current environments, caused by natural selection.

Variation at several enzyme loci appeared to be related to temperature. Highest contributions to the adaptive pattern were found for the loci *Mnr-1*, *Mnr-2*, *Lap-2*, *Pep-3*, *G6p-2*, *Acp-1*, *Gdh-1*, *Idh-1*, *Pgm-1*, *Skd-2*, *Fdp-2*, *Adh-2* and *Mdh-3*. Hence, variation at 13 out of 31 loci seemed to be directly or indirectly related to temperature. Even if the associations were low at the level of the alleles, for the number of loci chosen at random, this is a very high number of observed associations, arguing against the neutrality of these allozyme markers. With few exceptions (LAP, ADH), all these enzyme systems belong to the so-called group I enzymes (GILLEPSIE and LANGLEY, 1974) which are characterized by their function in the primary metabolism (glycolysis, citric acid cycle, pentose phosphate cycle, etc.) and most of which act only on a single physiological substrate. Allelic variation at these loci, which code for enzymes that are most relevant for fitness and adaptation, may be viewed as a result of balancing selection caused by differences in the function and kinetic properties of the variants, leading to differences in fitness under differing environmental conditions (BERGMANN et al., 1990). A huge body of evidence for the existence of differences in the kinetic properties and fitness of allelic variants under specific environmental conditions has been reported in the literature. An overview has been presented in *section 1*. The frequent or the two most frequent alleles are believed to constitute the operating adaptive potential which becomes apparent in the set of environmental conditions to which a population can adapt (or which it can survive) without the necessity to change its genetic composition. The rare variants, on the other hand, can be viewed as the latent genetic potential to be utilized for colonization or adaptation to changed environments (BERGMANN et al., 1990). The large proportion of allelic variants which show an association with the temperature related pattern clearly support this view. Our results strongly argue for the existence of differences in properties and functions among the enzymatic variants, for their differences in fitness under differing environmental conditions and for their selective non-equivalence under natural selection. The non-neutrality of these loci is also strongly supported by the striking congruence of our results with results reported in the literature (see *section 1*). For most of the enzyme systems with correlations of more than 0.20 with the adaptive response surface *i.e.* IDH, PGM, MNR, LAP, G6P, ACP, GDH and MDH, direct or indirect associations with temperature have already been reported in many publications and for a number of different species. No comparative results exist only for the enzyme systems ADH, FDP, SKD and PEP. The latter two enzymes, however, were assayed only very occasionally.

A high correlation among allele frequencies at an *Idh* locus and temperature has been reported for Pitch pine, *Pinus rigida Mill.*, (GURIES and LEDIG, 1981) and beech, *Fagus sylvatica* L., (GÖMÖRY et al., 1992). Indirect evidence for associations with temperature may be inferred from the reported associations with latitude found in different studies on Silver fir, *Abies alba Mill.*, (MOLLER, 1986; BERGMANN et al. 1990; KONNERT, 1992). In the same species, the observed clinal variation in frequencies of the two alleles could be linked to kinetic differences regarding thermostability and catalytic efficiency of the two variants under different temperature regimes (BERGMANN and GREGORIUS, 1993). Similar results have also been reported for allelic variants in trout species by MOON and HOCHACHKA (1971, 1972). In Lodgpole pine, *Pinus contorta Dougl. ex. Loud.*, in the Yukon and British Columbia, *Idh* showed the highest association with the first discriminant axis which paralleled a strong north-south pattern (YEH et al., 1985). A contribution of *Idh* to patterning was also reported for White fir, *Abies concolor Lindl.*, Sugar pine, *Pinus lambertiana Dougl.*, ponderosa pine, *Pinus ponderosa Laws.*, and Douglas-fir, *Pseudotsuga menziesii Franco*, in the Sierra Nevada in California (WESTFALL and CONKLE, 1992).

YEH and O'MALLEY (1980) reported a high correlation of allele frequencies at a *Pgm* locus with latitude in Douglas-fir, *Pseudotsuga menziesii Franco*, which indirectly implies an association with temperature. Likewise, significant differences in allele frequencies at the *Pgm* locus which were observed between mesic, north-facing slopes and xeric, south-facing sites in ponderosa pine, *Pinus ponderosa Laws.*, populations (HAMRICK et al., 1989) suggest an adaptive role of this enzyme with respect to temperature and moisture regime. Variation at *Pgm* was also moderately correlated with the first discriminant axis which paralleled a strong north-south pattern in Lodgpole pine, *Pinus contorta Dougl. ex. Loud.*, in the Yukon and British Columbia

(YEH et al., 1985). Moreover, variation at *Pgm* was also associated with the second discriminant axis which was related to differences in elevation.

Associations of allele frequencies at *Mnr* with temperature were reported for beech, *Fagus sylvatica L.*, by GÖMÖRY et al., (1992). Indirect evidence for the adaptive significance of *Mnr* relative to temperature is also provided by results found for Douglas-fir, *Pseudotsuga menziesii Franco*, in southwestern Oregon as reported by MERKLE et al. (1988). The highest correlation with a second order spatial model including latitude, longitude and elevation was found for a *Dia* locus. Diaphorase is a synonym for MNR, describing the same enzyme system.

Different authors have reported associations of variation at the LAP enzyme with temperature. Direct correlations among allele frequencies and temperature have been reported for Pitch pine, *Pinus rigida Mill.*, (GURIES and LEDIG, 1981) and Norway spruce, *Picea abies Karst.*, (STUTZ, 1990). Indirect evidence comes from associations of allele frequencies at a *Lap* locus with latitude which were described in Norway spruce, *Picea abies Karst.*, (BERGMANN, 1973, 1975b) and Pitch pine, *Pinus rigida Mill.*, (FRYER, 1987). Associations with altitude were found in Norway spruce, *Picea abies Karst.*, in Sweden by LUNDKVIST (1979). The importance of LAP for adaptation to the environment has also been demonstrated by results reported by KIM (1980) and MÜLLER-STARCK and HATTEMER (1989) who described associations between the viability and survival of beech, *Fagus sylvatica L.*, seedlings and genotypic structure at the *Lap* locus.

Frequencies of allelic variants at the *G6p* locus were correlated with temperature in pitch pine, *Pinus rigida Mill.*, (GURIES and LEDIG, 1981). Indirect support for an adaptive role with respect to temperature conditions comes from associations with latitude which were reported for black spruce, *Picea mariana (Mill.) B.S.P.*, in Newfoundland by YEH et al. (1986). Likewise, *G6p* was moderately correlated with the first discriminant axis paralleling a strong north-south pattern in Lodgpole pine, *Pinus contorta Dougl. ex. Loud.*, in the Yukon and British Columbia (YEH et al, 1985). An important role of the G6P enzyme in the adaptation process to temperature may also be concluded from the results published by KONNERT (1991). High correlations of allele frequencies at a *G6p* locus with phenology and growth traits were found in one-year old Norway spruce, *Picea abies Karst.*, seedlings. Moreover, allelic variants at a *G6p* locus seemed to play an important role in the stress tolerance against air pollution (SCHOLZ and BERGMANN, 1984).

Associations between alleles at a *Acp* locus and temperature have been reported by several authors. High correlations between *Acp* allele frequencies and temperature were described for beech, *Fagus sylvatica L.*, by GÖMÖRY et al. (1992). Indirect associations with latitude and altitude were observed repeatedly in Norway spruce, *Picea abies Karst.*, by BERGMANN (1973, 1975a, 1978) and in pitch pine, *Pinus rigida Mill.*, by FRYER (1986). Variation at the *Acp* locus was also found to be related to stress tolerance against SO_2 pollution (MEJNARTOWICZ, 1983).

Gdh was found to be related with altitude in Norway spruce, *Picea abies Karst.*, (LUNDKVIST, 1979). In Lodgpole pine, *Pinus contorta Dougl. ex. Loud.*, in the Yukon and British Columbia, variation at *Gdh* was associated with altitude, loading high on the second discriminant axis which represented variation in altitude (YEH et al., 1985). Significant differences in allele frequencies at the *Gdh* locus which were observed between mesic, north-facing slopes and xeric, south-facing sites in ponderosa pine, *Pinus ponderosa Laws.*, populations (HAMRICK et al., 1989) suggest an adaptive role of this enzyme with respect to temperature and moisture regime.

Several reported results indicate an association of allelic variation at *Mdh* with temperature. A high correlation between variation at locus 2 and temperature was reported for Pitch pine, *Pinus rigida Mill.*, by GURIES and LEDIG (1981). Associations with latitude were reported for Douglas-fir, *Pseudotsuga menziesii Franco*, by YEH and O'MALLEY (1980). Associations with altitude were observed for Black spruce, *Picea mariana (Mill.) B.S.P.*, in Newfoundland by YEH et al. (1986) and for Lodgpole pine, *Pinus contorta Dougl. ex. Loud.*, in the Yukon and British Columbia (YEH et al., 1985). In the latter case, it was also variation at locus 3 which showed a high association.

In conclusion, associations with temperature have been repeatedly described for most of the loci which were associated with the adaptive trend surface. Lack of associations have also been reported for these loci, this has, however, mostly been on the single-locus level, using traditional univariate measures of diversity and differentiation. In our case, all these loci had already shown a relationship with temperature in the single locus analyses. Based on these comparable results, we conclude that these loci *i.e. Idh, Pgm, Mnr, Lap, G6p, Acp, Gdh* and *Mdh* are adaptive or tightly linked to adaptive markers.

No associations with temperature have been reported for *Fdp*, although this enzyme has been investigated in most of the allozyme studies. The rather high association with the adaptive response surface may be due to an indirect, region related effect. Results reported in *section 5* suggested such an indirect effect. According to the results of the multinomial response models, the association of variation at the *Adh* locus is also most likely caused by indirect effects. However, *Adh* has been found to show associations with altitude in Lodgpole pine, *Pinus contorta Dougl. ex. Loud.*, in the Yukon and British Columbia (YEH et al, 1985). Moreover, fitness differences in animals and plant species have been reported for the *Adh* locus (*see section 1*).

The importance of separating direct from indirect effects has been extensively discussed in *section 4.5*. Especially the question regarding the existence of different races (Pacific race, North Plateau race) is of special interest, since several results, mainly based on quantitative traits, were taken as evidence for the existence of different races on the east and west sides of the Cascade Range in Oregon (*see section 2*). Results from single locus analysis in this study also revealed differentiation between the two regions for allele frequencies at several loci. Since CONKLE and CRITCHFIELD (1987) and NIEBLING and CONKLE (1990) have already reported differentiation between the Pacific and the North Plateau races for the same loci, the genetic distance of 0.11 being in the range reported for races (NEI, 1974), the racial theory seems to be supported by the patterns of allele frequency differentiation observed in this study.

Potential past and present forces involved in generating the observed variation patterns could, however, not be separated using univariate procedures. With multivariate techniques, however, the explained variance in the model could be partitioned into different components, representing different possible effects. The proportion of model variance which was related to location only, was interpreted as being most likely due to either evolutionary events in the past (migration) or random processes which are unrelated to natural selection. In contrast, proportions of model variance which were associated with either pure environmental variation or with the shared variation between environment and location (due to spatially structured environments) were taken as being most likely a result of natural selection.

Since only 2.5% of the significantly explained variance in the adaptive response surface of multilocus frequencies was attributable to pure spatial effects, historical events such as a long separate evolution of the base populations or the immigration from two different refugia are extremely unlikely. Ninty seven percent of the variance in the response surface were associated with the climatic conditions at source locations. The response surface thus represents adaptive multilocus genotypic variance. This very high proportion of climate effects leads to the conclusion, that the observed differences in allele frequencies between the two regions are rather the result of natural selection than the consequence of racial differentiation due to a different evolutionary history. Results from multivariate analysis clearly indicate that the observed differences in allele frequencies at several loci are primarily due to the marked environmental differences in climate conditions which exist between the east and the west side of the Cascade Range. Moreover, the associations of multilocus frequencies with climate variation strongly argue for natural selection being the major force involved in generating the patterns of variation of the present populations.

Based on estimates of time of divergence, we had speculated in *section 4.5* that ponderosa pine in the area must have had a long common evolutionary history. In combination with the relatively short history of immigration we had argued that differentiation between the east and the west side of the Cascade Range were most likely the result of the relatively recent separation of the same base population, the differentiation being the result of restricted gene flow in combination with different selection pressures due to the marked environmental differences.

Morphological differences were similarly interpreted and viewed as additional arguments for natural selection being the major force responsible for the observed differentiation between the two areas. Results of multivariate analysis clearly support these speculations. Preliminary results on racial differentiation between the Pacific race and the North Plateau race of ponderosa pine in southern Oregon using RAPD markers did also not indicate a racial differentiation between families from about 100 km west to 150 km east of the presumed boundary between the races (AAGAARD et al., 1993).

Multivariate allozyme patterns appeared to be rather weak, highly complex and mostly nonlinear within Southwest Oregon. With a sample to variable ratio of about 5, results are expected to be stable and reliable. Although 25 % of the variance was accounted for by the climate model (which is about the same as in Central Oregon), weak patterning was indicated by a moderate level of significance ($p = 0.051$) of the first response variate and especially by low cross-loadings of the climate variables in the model (*Table 35, p. 150*). Although the level of significance could have been partly caused by the relatively low sample size, the low loadings of climate variables and former results rather argue for a weak patterning in this area. A weaker patterning in Southwest Oregon compared to Central Oregon, was already apparent in the former analyses. Discriminant analysis on groups of similar genotypes produced no significant proportion of climate variance that was associated with group differences. Results of discriminant analysis on groups of similar habitats were complex, discriminant axis were unstable and discriminating variables were not congruent with former results, suggesting that patterns are weak. In addition, no significant associations of genotypic frequencies with climate were found in the single locus analyses using multinomial response models.

A weak patterning of multilocus genotypic variation in the Klamath National Forest, just south of our sampling area, was reported for Douglas-fir, *Pseudotsuga menziesii Franco,* by WESTFALL and CONKLE (1992). In contrast to the Sierra Nevada populations where two significant canonical vectors described 63% of variance, none of the vectors were significant in the Klamath sample. The amount of variation described by the first canonical variate was only 17%. The strongest pattern in multilocus frequencies was east-west which is comparable with the pattern found in our sample (see below). Similar results were reported by MERKLE et al. (1988) who described statistically significant but weak patterning in Douglas-fir populations from Southwest Oregon. The first two canonical vectors were significant and described 25% of variation in haploid scores of 27 loci. Cross loadings of geographic variables with the first two vectors were weak, however. Highest associations were found with latitude and distance from the ocean, indicating that the pattern shows a northeast-southwest direction. An east-west pattern for seedling quantitative traits of Douglas-fir in this area was reported by CAMPBELL, (1986, 1991). Weak patterning *i.e.* weak associations with habitat variables, in Southwest Oregon were also reported for seedling traits of Sugar pine, *Pinus lambertiana Dougl.*, by CAMPBELL and SUGANO (1987). The regression equation, relating factor scores of the first principal component (representing growth vigor) to habitat variables, explained only 31% of the sums of squares for the first component. This amount of association is small compared to results reported for other species or areas. SORENSEN (1994), for example, reported a multiple R^2 of 0.69 between the scores of the first principal component (representing growth vigor) and a geographic/topographic model for ponderosa pine families from central Oregon. Since patterns of adaptive variation seem to be weak for all the species growing in southwestern Oregon, it may be hypothesized that the populations in southwestern Oregon undergo a less stringent selection pressure because climate conditions are rather mild and lack the extreme values which are, for example, found to the east of the Cascade Range or in the Sierra Nevada in California. Alternatively, the environmental models used to describe associations possibly do not adequately account for the selection regimes in this area.

Both climate factors were only weakly related with the adaptive trend surface in Southwest Oregon. The trend surface was about equally associated with temperature and moisture conditions of the site, non-linear moisture gradients contributing slightly more to the explained variance than temperature. Concordant with the results from single locus analyses and discriminant analysis, the results of canonical trend surface analysis indicate that moisture regime seems to be an important factor influencing adaptation of populations in this region. Climatically, Southwest Oregon is characterized by remarkable gradients in precipitation. Moist conditions prevail in the higher elevations of the northern Siskiyou National Forest, the Kalmiopsis Wilderness

and in the Umpqua National Forest while the central areas are characterized by less annual rainfall and especially by moisture deficits in summer, receiving locally less summer rainfall than sites in the high desert country to the east of the Cascade Range. Temperature, on the other hand, is on average mild and temperature gradients are less pronounced than gradients in precipitation. Hence, it is tempting to interpret the observed pattern of variation as a result of natural selection. Based on our results, temperature generally seems to influence adaptation more than moisture conditions. However, strong moisture gradients in Southwest Oregon in combination with a rather mild and oceanic climate (with less extremes) may lead to a more pronounced effect of moisture as a selective agent than in Central Oregon. Moreover, water availability may be selectively more important in southwestern Oregon because of the distinct geologic substrate found in this area. Large masses of peridotite and dunite, in most places altered to serpentinite, are widely distributed throughout the Klamath Mountains (BALDWIN, 1976). In southwestern Oregon, ponderosa pine is frequently found on infertile soils which are low in contents of various minerals and which locally contain high levels of chromium, magnesium and nickel (WALKER, 1954). It may be speculated that moisture conditions are more important on such sites because nutrient uptake depends on a minimal level of transpiration and water uptake. Adaptation of ponderosa pine to ultramafic soils in this area has been reported by JENKINSON (1974). Observed differences among progenies were related to differences in the capacity to take up calcium from the soil; the better growing progenies having a higher concentration of calcium. Periods of drought may also be more important because heavy metal ions may have a toxic effect in the concentrated soil solution. Directional selection in relation to toxic minerals has been reported for Jack pine, *Pinus banksiana Lamb.*, by XIE and KNOWLES (1992), Bishop pine, *Pinus muricata D. Don.*, by MILLAR (1989) and herbaceous plant species such as *Agrostis tenuis Sibh.*, by ANTONOVICS et al. (1971). Serpentines and ultramafic soils are known to support a unique, tolerant flora (KRUCKEBERG, 1967, 1987). Directional selection and adaptation to poor soil conditions has been hypothesized as a possible cause for observed genetic differences found among populations of Jeffrey pine, *Pinus Jeffreyi Grev. & Balf.*, in the Klamath Mountains of southwestern Oregon (FURNIER and ADAMS, 1986). Likewise, soil moisture has been considered as the most important limiting factor for the growth of ponderosa pine in the Klamath area by DYRNESS and YOUNGBERG (1966).

Associations between seedling quantitative traits and moisture characteristics of sites in the Klamath Mountains have been reported for sugar pine, *Pinus lambertiana Dougl.*, by CAMPBELL and SUGANO (1987). Regression equations indicated that variation in scores on the first principal component, representing growth vigor, was associated with annual precipitation, elevation and distance from the ocean. Interestingly, a similar conclusion may be inferred from the results published by MERKLE et al. (1988) who studied multivariate allozyme patterns of variation of Douglas-fir, *Pseudotsuga menziesii Franco*, in Southwest Oregon breeding units. In the plot of scores on the first two discriminant axis, breeding units with positive scores on the first axis seem to be associated with dry habitat conditions while zones with negative scores seem to be situated in areas with higher moisture availability.

In conclusion, patterns of adaptive variation in Southwest Oregon seem to be weak, highly complex and seem to reflect the topographic and ecological complexity of the region (FRANKLIN and DYRNESS, 1973). The presented adaptive response surface was the result of a climate model which was only moderately significant. Patterns therefore must be regarded as preliminary and descriptive. Nonetheless, the canonical model clearly indicated that a complex combination of environmental conditions seems to be involved in the shaping of genetic variation. Our data provide evidence, that the strong moisture gradients found in southwestern Oregon are reflected in the genetic structure to about the same degree as temperature conditions.

With 23% of total variance, canonical trend surface analysis of the Central Oregon sample revealed a moderate proportion of variance in 33 allele frequencies which was associated with the second order climate model (*Table 37, p. 154*). The adaptive trend surface was highly significant with adequate fit and normally distributed residuals. Results were meaningful and in accordance with the results of the former analyses. For example, the amount of association was only slightly higher than the estimate based on discriminant analysis on groups of similar habitats. Moreover, alleles which contributed most to the trend surface were the same alleles which were already found to be potentially adaptive in the single locus analyses and which were loading high on the first discriminant axis resulting from discriminant analysis on groups of

habitats. With a sample to variable ratio of 5.8, canonical trend surface analysis is expected to produce stable canonical functions and weights.

Not all enzyme systems which were related to the adaptive surface in the overall analysis were also related to the response variate in the Central Oregon sample (*Lap-2*, *G6p-2*, *Acp-1* and *Gdh-1* were related in overall analysis but not in Central Oregon). Two explanations for the lack of associations of these loci are possible. First, the high associations in the overall model may have been caused by indirect effects. Even if such indirect effects are small regarding the overall variance over all loci, as we have shown, they may occur at the level of single genes. Alternatively, variation in temperature within Central Oregon may be too small to reveal associations. It is possible that the associations are discernible only when the variation in habitat conditions and the sample size is large enough. In addition to the enzyme systems which proved to be potentially adaptive in the overall analysis, three more alleles revealed an association with climate: *Mpi1-1*, *Ugp1-1* and *Ugp1-2*. These alleles also showed associations with temperature or moisture in the multinomial response models.

Multilocus frequencies in Central Oregon were related to both temperature and moisture conditions. Temperature, however, was clearly the primary factor of influence, contributing four times as much to the sums of squares than moisture conditions. Associations are mostly linear, but nonlinear terms and the interaction term also contribute small but significant proportions to the model variance. The important role of temperature is best illustrated by the high congruence between the trend surfaces for the predicted allozyme scores and for climate *Factor 1* (*Figure 42, p. 157*).

Based on several quantitative seedling traits from the same families which were also used in our investigation, SORENSEN (1994) concluded that temperature was the main factor of influence on genetic differentiation in Central Oregon ponderosa pine. He hypothesized that shoot elongation potential of ponderosa pine is strongly selected by temperature, but is less selected by moisture characteristics of the seed source. According to SORENSEN, growth potential is plastic when moisture is the selective agent but it is more inherently fixed when temperature is the agent. Seedlings from sites with a short growing season induced by summer drought respond to favorable growing conditions with increased (plastic) growth. Seedlings from sites with a short growing season induced by temperature, however, have much less ability to respond to the same favorable conditions (fixed response). Similar results for ponderosa pine were reported by REHFELDT (1984, 1986a, 1986b). Based on common garden studies in the northern Rocky Mountains, REHFELDT described rather gentle clinal variation across latitude and longitude but relatively steep clinal variation across elevation. Results indicated a close association between elevation of the seed source and various components of shoot elongation. In addition, in a number of tests with ponderosa pine in which water stress was applied in one of the test environments, seedlings responded in a plastic manner to moisture deficits with decreased height and reduced needle length. In no case, however, was there any evidence for a genetic response which could be interpreted as adaptive to different moisture regimes. Even under moisture stress, genetic differences were best explained as adaptation to frost free periods of variable length (REHFELDT, 1986a, 1986b, 1990a, 1993).

The adequacy of the adaptive multilocus response surface which we have presented for Central Oregon may best be evaluated by comparing our pattern of variation with the pattern published by SORENSEN (1994). This excellent and unique opportunity for a comparison of patterns of variation based on both gene markers and quantitative traits is possible since the same families were used for both investigations. SORENSEN subjected 10 seedling traits which were evaluated in a common garden study (*see section 3.4*) to principal component analysis. The first principal component (PCA 1) accounted for 46.6% of location related variance. It was primarily size related; large scores identified locations with tall, slender seedlings with a large overall rate of elongation, but proportionally little elongation early in the season. PCA 1 was primarily correlated with elevation (*-0.542*), with distance from the Cascade crest (*-0.205*) and latitude (*0.144*). Interestingly, the regression model based on 24 geographic terms explained 24% of variance in factor scores of PCA 1; this amount is highly congruent with the 23.3% of explained variance that we have found in our canonical model. Explained variance increased, however, by 55% when topographic terms such as aspect and slope were added to the regression model. This result suggests that the amount of association found in our canonical models

most likely still underestimates true relationships since climate variables were estimated based on longitude, latitude and elevation only, so that estimated climate conditions at seed source do not reflect important local modifications caused by topography.

The pattern of variation on the first principal component was interpreted by SORENSEN as being a result of a strong selection pressure induced by temperature, mainly acting on shoot elongation. The spatial pattern of variation in factor scores of the first PCA is illustrated geographically on page 13 of SORENSEN's publication. Since this plot shows pure factor scores, unadjusted for elevation, it represents the simple geographic variation pattern of scores predicted for each sample location. Hence, it is directly comparable with our contour plot developed from the allozyme response surface. In order to facilitate comparison, the two patterns are presented side by side in *Figure 49, p. 183* .

The two patterns of variation are strikingly congruent. Minor differences may only be observed in the southern part where the zone marked with the + signs extends slightly more to the east and the zone marked with the 0 signs extends more to the south than the areas in our plot. With the exception of these minor differences, the isolines are more or less congruent. The areas of low PCA scores (marked with = and - signs) are identical with zones exhibiting scores of less than -0.45 on the allozyme trend surface. These areas with low seedling vigor are in fact, as SORENSEN speculates, characterized by the lowest temperatures as may be seen in the lower half of *Figure 42, p. 157,* which shows temperature variation in the area. The areas marked with the + signs are identical with the areas showing scores of 0.45 and higher on our response surface. The areas in between, marked with the 0 signs, are congruent with areas exhibiting scores between -0.45 and + 0.45 on the allozyme surface. The zone with the highest scores on PCA 1 (marked with the * sign) in the northwestern corner of the area (Columbia river gorge) also shows the highest scores on our canonical response surface. Although it is not apparent in *Figure 49, p. 183*, scores in this part of the area are above 0.66 (see the plot scaled to 30% transfer risk shown on the right of *Figure 48, p. 170*). A similar zone of high scores above 0.66 on our response surface was also apparent in the southwestern corner of the sampling area. The PCA surface of SORENSEN does not show this zone. As noted by SORENSEN, the Klamath river, like the Columbia river in the north, cuts through the Cascade Range, forming a drainage opening towards the Pacific Ocean. Penetrating maritime environmental influence is believed to be responsible for the locally favorable climate conditions (and high grow vigor) in the Columbia river basin. If this is really the case, a similar zone may be expected also in the southwestern corner of the sampling area. In fact, a maritime influence across the Cascade crest in this area may be seen in *Figure 30, p. 133.*

In conclusion, the multilocus pattern of allozyme frequency distribution in Central Oregon is significant and moderately strong. Patterns are primarily related to temperature at source location. Small contributions can also be attributed to moisture characteristics of the sites. Patterns are meaningful and fully in accordance with the former analyses. They agree surprisingly well with patterns based on seedling quantitative traits. The two response surfaces are nearly identical. With minor differences, they furnish the same information on patterns of genetic variation in Central Oregon, thus providing strong evidence that allozyme markers can be very useful in describing ecologically important patterns of adaptation. Moreover, it appears that canonical trend surface analysis is capable of retrieving the same information in one single analysis, making the analysis easier and more straightforward than the indirect approach based on principal component and regression procedures which was used by SORENSEN.

Relative transfer risk estimates the proportion of plants in a seed source population that are presumed to be at risk from from maladaptation if planted in another location (environment). Risk does not necessarily imply mortality but rather a lack of ability to respond to site conditions as would the local population. Deciding on the level of an acceptable relative transfer-risk for practical applications is a complex task. Based on silvicultural considerations, an acceptable transfer-risk should be a function of tree species, site conditions and management objectives. In fact, different opinions regarding acceptable levels of relative transfer-risk for field guidance have been presented in the literature. For ponderosa pine, SORENSEN (1994) considered a risk of less than 51% as acceptable for practical purposes. This risk level is based on a planting density of 1' 076 seedlings per hectare (3.05 x 3.05 m spacing), 30% loss due to random me-

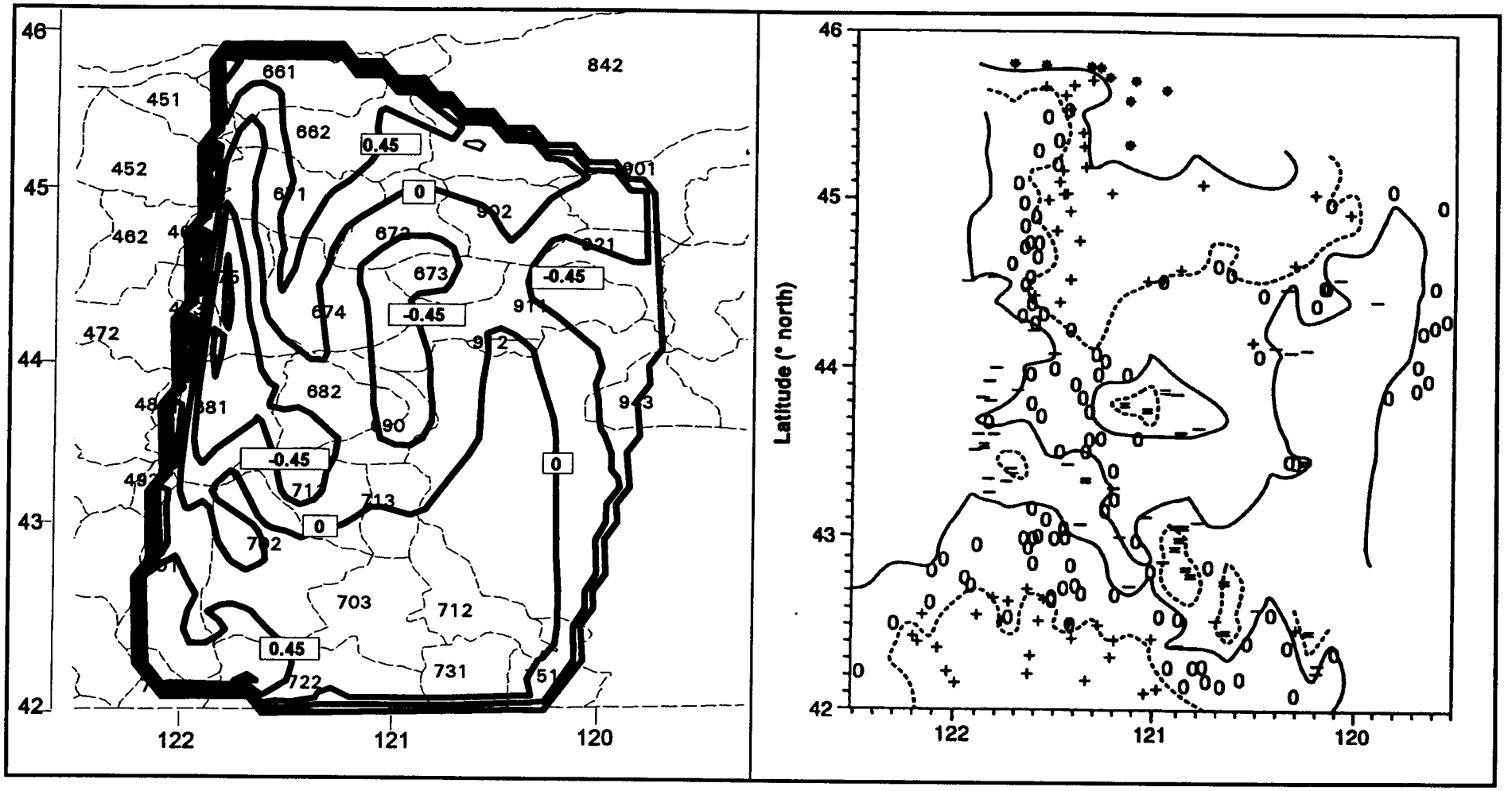


Figure 49: *Patterns of genetic variation, based on allozyme scores and seedling quantitative traits: Central Oregon. Left: Contour plot of adaptive trend surface based on allozyme scores, scaled to represent a relative transfer risk of 20% among isolines. Right: Factor scores of first principal component plotted by latitude and longitude as given by SORENSEN (1994). Factor scores are not adjusted for elevation. PCA 1 is associated primarily with vigor traits. Symbols grade from ▪ (lowest factor scores, least vigor) through -, 0, + to ▪ (highest factor scores)*

chanical factors and a density of 176 trees per hectare as targeted crop trees. CAMPBELL (1986) and CAMPBELL and SUGANO (1987) recommended levels of 25% to 30% as an acceptable risk for the transfer of Douglas-fir, *Pseudotsuga menziesii Franco*, and Sugar pine, *Pinus lambertiana Dougl.*, respectively, in southwestern Oregon. A conservative level of acceptable risk was proposed by MILLAR and WESTFALL (1992). These authors considered a transfer-risk of 12% as upper limit for within-zone transfers, the zones serving as genetic resource management units for gene conservation of White fir, *Abies concolor Lindl.*, on the west slope of the Central Sierra Nevada in California. To illustrate mapped transfer risks based on adaptive trend surfaces, two levels of relative risks (20% and 30%) were used for ponderosa pine in western Oregon. It can be seen from the contour maps, that this difference in 10% in the acceptable risk can have a large impact on zone formation. In areas with highly variable topography, harsh planting sites and steep moisture and temperature gradients, transfer risk may exceed 10% on very short geographic distances.

As a general rule, seed should not be moved across the Cascade Range. Based on genotypic frequencies, 50% of the individuals on average will not be adapted to sites in the other region. In extreme cases, a mismatch between genotypes and planting sites as high as 90% may result from such transfers. However, average values of transfer risk can provide only crude guidelines for seed transfer since transfer risk can be highly variable within some of the zones, but not within others. A better illustration of the continuously changing patterns of transfer risk which is caused by gradual changes in genotypic frequencies in geographic space, is provided by contour maps of transfer risk. Contour maps are helpful in illustrating gradients of transfer risk on the topographic map. Moreover, maps constructed in this way delineate areas of given maximal transfer-risk and thus provide useful information on the adequacy of the current seed zones.

Inferred from sampling the entire study area, seed transfer from central areas in southwestern Oregon to locations east of the Cascade Range has in a transfer-risk of more than 40%. Sources from the higher elevations of the western areas (Kalmiopsis wilderness and Northern Siskiyou National Forest) as well as sources from the eastern part in the Umpqua and Winema National Forests can be transferred to the northern as well as to the southern parts of Central Oregon with a smaller risk of about 30%. Transferring the same sources to eastern areas within Central Oregon, however, will result in a risk higher than 40%. Within southwest Oregon, transfer-risk is generally higher for transfers in the east-west direction than in the north-south direction. Highest risks (> 20%) occur when sources are moved either from the eastern or the western parts of this zone to central areas or vice versa, whereas risks remain below 20% when sources are used within the central area, or when sources from the part of the zone most eastern are used on the most western sites and vice versa. In Central Oregon, highest risks (>30%) occur when either sources from high elevations of the Cascade Range, either from the southernmost (zones 701, 721, 722, 702, 703, 712, 731 and 751) or from the northernmost areas (zones 661, 662 and 842) are transferred to sites in Crook, Jefferson, Wheeler and Lake county or vice versa. As we have shown, overall transfer risks are generally oriented in the same direction as thermal gradients. Greatest transfer-risk thus exist between areas of greatest difference in temperature regime.

Contour maps derived from trend surfaces of adaptive genotypic variation within each region, focus more on local variation of ecological conditions than on the overall thermal gradients, and show more complex and subtle patterns of relative transfer-risk. Transfer risks at the regional scale can differ from risks evaluated over the total sampling area since estimates are relative, depending on the variance of the sample. Estimates derived from total sampling area primarily focus on transfers between the two regions; the within-region risks therefore are smaller in the overall analysis compared to the estimates which are relative within each of the two regions. Because sources should not be moved across the Cascade Range, but only used within the respective region, seed transfer quidelines should be based on maps derived from within-region estimates.

High variability and steep gradients in risk were especially found within Southwest Oregon. The general west-east pattern was slightly modified into a northeast-southwest pattern in the central area. However, since the adaptive pattern was weak and the model only moderately significant, the patterns of transfer-risk presented are only preliminary. Directions of gradients in transfer-

risk may be portrayed adequately, levels of transfer risk, however, may not be reliable. Nonetheless, two observations are of special interest. Current seed zones within Southwest Oregon seem to reflect the general ecological and genetic differences better than seed zones in Central Oregon. Zone boundaries generally parallel the contours quite well, indicating that the important ecological and adaptive gradients are reflected in the zonation. Obviously, delimitation of ecologically distinct areas was easier in this topographic highly variable area than in the Cascade Range and the Ochoco Mountains. However, based on our preliminary results, the current seed zones seem rather large to guarantee an acceptable transfer-risk within the zones. All existing seed zones except zone 492 seem to be characterized by a rather high within-zone variability in transfer risk. The development of adequate seed zones and of seed transfer guidelines seems to be a highly complex task in this area. On the other hand, the selection of suitable planting material is a very important management decision. Seed zones in this region may not be an adequate tool for guiding seed transfer. Models which allow the selection of suitable seed sources for any given site, based on direct estimates of transfer risks between the two sites, seem the best solution for minimizing transfer risk (CAMPBELL, 1986; REHFELDT, 1990b).

In contrast to Southwest Oregon, seed zones in Central Oregon seem to be conservative for ponderosa pine. Based on an acceptable level of transfer risk of 30%, many of the current seed zones could be joined into considerably larger units. The same conclusion was also reached by SORENSEN (1994) on the basis of seedling traits. Based on multilocus genotypic frequencies, sources from the southern area can be moved to northern sites and vice versa with a relative transfer-risk smaller than 30%. Sources from the central area, however, seem to be distinct and should be used only within this area. Both areas, however, could encompass several to many of the current seed zones. Based on contour maps, a total of only about 7 seed zones is required to guarantee a maximal transfer risk of 30% (*Figure 48, p. 170*). A comparison of the existing seed zone boundaries with the isolines of the adaptive response surface suggests that the important environmental and genetic patterns of variation are not well reflected in the present zonation.

6.6 Summary

Multilocus genotypic frequencies are moderately associated with habitat conditions. Based on different multivariate analyses, climatic conditions at source location significantly explain about 33% of total variance of 41 allozyme variables. Within the regions, the proportion of explained variance is about 23%. Overall, 13 out of 31 loci show an association with climatic conditions. These results clearly argue for the adaptivness of certain alleles or enzyme systems. In accordance with the former results, especially enzyme systems involved in important physiological pathways such as the loci *Mnr-1*, *Mnr-2*, *Lap-2*, *Pep-3*, *G6p-2*, *Acp-1*, *Gdh-1*, *Idh-1*, *Pgm-1*, *Skd-2* and *Mdh-3* seem to be adaptive. Not all alleles contribute to the adaptive pattern of variance. Of the markers, 42% were not related with the pattern. Amounts of associations differ among the alleles. Of the alleles, 16% showed a correlation of more than 10%, 31% of more than 20% and 11% of more than 30% with the adaptive response surface. Single alleles thus behave much in the way one would expect of quantitative trait loci; each gene contributing only small amounts to the adaptive multilocus pattern of variation. While these small individual contributions may be difficult to detect with single locus analyses, the aggregate patterns resulting from multivariate analyses can be strong and meaningful.

All results generally agree with results reported in the literature and are congruent with the results reported in *section 4* and *section 5*.
Close associations of patterns of multilocus allozyme frequencies with environmental variation within each of the two regions are potentially consistent with migration (hypothesis 2) and adaptation (hypothesis 1). Complex patterns of allozyme variation associated with complex environmental variation, however, are unlikely under hypothesis 2. Since patterns are repeated in different places and since they closely resemble patterns seen in quantitative traits published for Central Oregon by SORENSEN (1994), adaptation to current environments caused by natural selection seems the most likely cause for the observed patterns of multilocus allozyme frequencies.

Multilocus genotypic frequencies were primarily related to temperature. All results clearly lead to the conclusion that temperature at source locations seems to be the most important environmental factor responsible for the observed adaptive pattern of multilocus frequency distribution. A striking similarity between the adaptive response surface and variation patterns of temperature conditions can be observed. Moisture characteristics of the site, on the other hand, have only a very minor overall effect on genotypic variation except in Southwest Oregon.

The adaptive patterns differ between the two regions. In Southwest Oregon, patterning is weak and both temperature as well as moisture conditions of the site have about an equal influence on multilocus frequencies. Strong moisture gradients combined with relatively mild climate may lead to a more important influence of moisture characteristics of the site in this area. The distinct geologic substrate and the soil conditions may also play an important role. Patterns in Southwest Oregon seem to be highly complex. They seem to reflect the topographic and ecological complexity of this area.

The adaptive pattern in Central Oregon is primarily related to temperature. However, moisture conditions and interactions contribute small but significant proportions to the explained variance. The adaptive pattern of multilocus frequencies is nearly identical with patterns based on seedling quantitative traits published by SORENSEN (1994). This high congruence of the two patterns provides strong evidence that allozyme markers can be very useful in describing ecologically important patterns of adaptation.

Based on multivariate analyses, differentiation among the two regions is rather the result of natural selection in two contrasting environments than the consequence of evolutionary events. The existence of two races (Pacific race, North Plateau race) as a consequence of a different evolutionary past, as suggested by several authors, is highly unlikely. Based on results of canonical and partial canonical analyses, only 2.5% of the significantly explained variance in the adaptive response surface is due to pure spatial effects. Pure spatial effects are reflected in patterns of variation which are independent of environmental variation. Such spatial effects should predominate if historical events were responsible for the present patterns of variation. Since 97.5% of variation in multilocus frequency distribution are caused by environmental effects and only 2.5% by pure spatial effects, historical events such as a long and separate evolution of different base populations or the immigration from two different refugia are highly unlikely. Results from multivariate analyses clearly demonstrate that differentiation between the two areas is primarily and nearly exclusively due to adaptation to different habitat conditions which exist in the two areas.

Contour maps of relative seed transfer-risk, calculated from the adaptive response surfaces, can be used to delineate areas on the adaptive surface with a transfer-risk smaller than a given value. Derived contour maps indicate that seed should not be transferred across the Cascade Range. Within Southwest Oregon, seed generally should not be moved in the east-west direction. Steep gradients in transfer-risk are observable in this area. The current seed zone boundaries seem to reflect the important ecological and genetic gradients rather well. However, zones seem to be too large to guarantee an acceptable transfer-risk. Seed transfer guidelines should be based on models of transfer risk and not on seed zones, since the formation of such zones seems highly complex.

The present seed zones in Central Oregon do not seem to reflect the important adaptive patterns. In contrast to Southwest Oregon, zonation seems to be rather conservative. Zones could be considerably larger, especially in the southern and the central part. To guarantee a maximum relative transfer-risk of 30%, only about 7 zones would be required.

7. Multivariate patterns of variation based on seedling quantitative traits and associations with climate conditions

7.1 Patterns of adaptive variation and relative seed transfer risks, inferred by canonical trend surface analyses

7.1.1 Southwest Oregon

Seedling traits of the same families which were analyzed for their enzyme genotypes were subjected to principal component analysis. Common garden procedures, traits and data transformations for the seedling quantitative data were described in detail in *section 3.4*. The 7 traits measured in the two different environments (cold and warm treatment) were treated as separate variables. Two additional variables, focusing on the different expression of genetic variance in the different environments, were included. Since second-year height was missing for the cold treatment, 15 traits were available. Original traits were subjected to principal component analysis, since growth and vegetative-cycle traits often form a co-adapted, highly correlated, multivariate system. Principal component analysis transforms this multivariate system into a reduced set of few orthogonal components which ideally focus on such co-adapted trait combinations. Moreover, the original data set was transformed in order to reduce the number of variables in canonical correlation analysis and to facilitate interpretation of results.

Results of principal component analysis, based on the correlation matrix of 15 seedling traits, are furnished in *Table 43, p. 188*. The first 4 principal components retained 83% of the original variance in the quantitative data. Final communalities were all above 70%, indicating that the 4-factor solution retained high amounts of variance of all original seedling traits. Since final communalities for some traits were reduced to values below 50% in the 3-factor solution, all 4 components were used in subsequent trend surface analysis.

PC-1, accounting for 34% of the total variance, was related to growth potential. High scores on this component thus identify families with tall seedlings and a good elongation potential. The second component reflected germination and emergence. High scores on *PC-2* identify families with an early development *i.e.* with high germination rates and early emergence. Early development had only a minor influence on first-year height. Early emerging families had slightly taller seedlings than late emerging families. This influence was no longer apparent after the second growing season and inherent growth potential was already fully expressed. Early developing families showed a weak tendency to also set their buds earlier. *PC-3* reflected plasticity of growth vigor. High scores for this component identify families which show a plastic reaction to the bed treatments, while low scores identify families which have much less ability to respond to changing environments (fixed reaction). Component 4 was related to growth timing. High scores on *PC-4* identify families which grow until late in the season. Late budset was associated to a certain extent with late emergence and with slow early development.

Results of canonical trend surface analysis, relating the 4 principal components to a second order climate model, are presented in *Table 44, p. 189*. Canonical correlation analysis produced three significant vectors ($p = 0.0001$, $p = 0.0026$, $p = 0.026$). Forty two percent of the patterned variation in the quantitative traits could be significantly explained by the climate model. The first response variate, accounting for 75% of total trace, was primarily associated with growth potential (*PC-1*) and growth timing (*PC-4*). A weak association was also evident with plasticity of growth potential (*PC-3*). High positive scores on the first response variate thus mainly represent a high growth potential and continuing growth until late in the season. Variation in the first response variate was primarily and positively associated with temperature at source location. Thirty two percent of the linear term and 24% of the quadratic term of climate *Factor 1* were redundant with variation on the first response variate. A minor relationship was manifest for the quadratic term of climate *Factor 2* with 2.85% redundant variation.

Thus, under the climatic conditions at the test site in Corvallis, families from mild and warm habitats showed a high inherent growth potential. Since these families set bud late in the season, this high growth vigor obviously is the result of a longer, inherently fixed growing period.

SEEDLING TRAITS	SOLUTION WITH 4 PRINCIPAL COMPONENTS					SOLUTION WITH 3 PRINCIPAL COMPONENTS			
	FACTOR PATTERN				Final	FACTOR PATTERN			Final
	PC-1	PC-2	PC-3	PC-4	Communality	PC-1	PC-2	PC-3	Communality
CEMERG	-0.104	-0.779	-0.021	0.299	0.707	-0.087	0.824	-0.007	0.686
WEMERG	-0.123	-0.688	0.087	0.467	0.714	-0.122	0.831	0.007	0.711
CGERM	0.094	0.859	0.131	0.019	0.736	0.053	-0.731	0.073	0.541
WGERM	0.109	0.871	0.013	-0.108	0.783	0.078	-0.807	-0.029	0.658
CBDUSET	-0.165	-0.147	-0.177	0.856	0.813	-0.204	0.561	-0.271	0.429
WBUDSET	-0.119	-0.289	-0.035	0.879	0.872	-0.157	0.695	-0.123	0.523
CFHT1	0.761	0.260	0.229	-0.130	0.717	0.750	-0.295	0.234	0.706
WFHT1	0.637	0.257	-0.053	-0.196	0.808	0.836	-0.327	-0.036	0.807
WFHT2	0.899	0.114	-0.059	0.000	0.825	0.893	-0.101	-0.057	0.810
CFHT3	0.850	0.028	0.411	-0.165	0.923	0.847	-0.111	0.439	0.923
WFHT3	0.903	0.000	-0.365	-0.029	0.950	0.910	-0.022	-0.348	0.950
CDIA3	0.824	0.046	0.445	-0.107	0.891	0.817	-0.097	0.460	0.888
WDIA3	0.848	-0.003	-0.419	-0.100	0.906	0.861	-0.056	-0.394	0.900
CFHT3-WFHT3	-0.007	0.034	0.937	-0.168	0.909	-0.018	-0.110	0.947	0.908
CDIA3-WDIA3	-0.004	0.054	0.954	-0.011	0.915	-0.027	-0.047	0.943	0.893
Eigenvalue	5.115	2.836	2.588	1.960		5.113	3.597	2.627	
Proportion	34.10	18.91	17.25	13.07		34.09	23.98	17.51	
Cumulative	34.10	53.01	70.26	83.33		34.09	58.07	75.58	

Varimax rotated factor solutions based on correlation matrix of family mean standard deviates of 15 seedling traits [standard normal deviates are used to remove differences among the 3 replications]

Abvreviations of variables: C: cold treatment - bed without cover in winter
W: warm treatment - bed covered with plastic tent in winter

EMERG: Emergence of seedlings after sowing [low values mean early, high values mean late emergence]
GERM: Germinaten rate of seeds based on 12 seeds per family plot [%]
BUDSET: Date of first year budset [low values mean early, high values mean late budset]
FHT1: Final height at end of first year [values for 2 replications only]
FHT2: Final height at end of second year [missing for uncovered treatment]
FHT3: Final height at end of third year
DIA3: Final diameter at end of third year
CFHT3-WFHT3: Difference in final third year height between cold and warm treatment
CDIA3-WDIA3: Difference in final third year diamater between cold and warm treatment

Final communality: Proportion of a variables variance accounted for by the retained principal components

The principal components represent:

FACTOR 1: Growth vigor
FACTOR 2: Early development [Germination and emergence]
FACTOR 3: Plasticity of growth potential
FACTOR 4: Growth timing

Table 43: *Principal component analysis of seedling traits: Southwest Oregon. Varimax rotated factor solutions based on family mean standard normal deviates of 9 seedling traits from 3 replications and two environments (treatments). The different treatments are included as separate variables (Prefix C and W). Sample size: N= 182 families*

In contrast, families from colder environments set bud earlier and consequently produced smaller seedlings due to the shorter growing season.

The overall model very significantly ($p = 0.0000$) and adequately described the first response variate since lack-of-fit was non significant at $p = 0.1708$ (*Table 45, p. 190*). Associations between the first canonical vector and the climate model were linear and were clearly dominated

SEEDLING TRAITS PC - FACTORS	SEEDLING TRAIT DATA SET								
	CANONICAL VARIATE 1			CANONICAL VARIATE 2			CANONICAL VARIATE 3		
	Struc.	Cross.	Red. %	Struc.	Cross.	Red. %	Struc.	Cross.	Red. %
PC - 1	0.8214	0.4879	23.81	0.5021	0.1474	2.17	0.1673	0.0461	0.22
PC - 2	-0.0541	-0.0321	0.10	-0.1661	-0.0488	0.24	0.9711	0.2678	7.17
PC - 3	-0.2706	-0.1607	2.58	0.7158	0.2102	4.42	-0.0004	-0.0001	0.00
PC - 4	0.4992	0.2965	8.79	-0.4560	-0.1339	1.79	-0.1702	-0.0469	0.22
CLIMATE FACTORS									
Factor 1		0.5670	32.15		0.0822	0.68		-0.0194	0.04
Factor 2		-0.0611	0.37		0.1088	1.18		-0.2068	4.28
Factor 1 square		0.4861	23.63		0.1437	2.06		0.0144	0.02
Factor 2 square		0.1685	2.84		0.1614	2.60		-0.1050	1.10
F1 X F2		0.0617	0.38		0.0241	0.06		-0.2188	4.80
Eigenvalue	0.5452			0.0943			0.0823		
% trace (pt)	75.27			13.02			11.36		
unb. can. corr.	0.5716			0.2936			0.2757		
r square	0.3267			0.0862			0.0760		
Significance	0.0001			0.0026			0.0259		
H-L-trace (Tr)	0.7243								
Model variance	42.01								
Sign. variance	41.86								

Dependent data set:
4 Principal components representing 15 seedling traits:

Dependent data set:
Full second order (quadratic) model with climate Factor 1and Factor 2 [PCA]

CLIMATE DATA SET					
CAN. VARIATE 1		CAN. VARIATE 2		CAN. VARIATE 3	
Struc.	% var	Struc.	% var	Struc.	% var.
0.9445	89.21	0.2800	7.84	-0.0705	0.50
-0.1029	1.06	0.3706	13.73	-0.7499	56.24
0.8183	66.96	0.4894	23.95	0.0522	0.27
0.2836	8.04	0.5496	30.21	-0.3809	14.51
0.1038	1.08	0.0822	0.68	-0.7933	62.93

Struc: Structure coefficients
Cross: Crossloadings
Red %: Redundancy
% var: Percent of variance explained
H-L-trace: Hotelling-Lawley-trace

Table 44: *Canonical correlation analysis: Seedling traits and climate variables: Southwest Oregon. [Sample size: N= 182 families]*

CANONICAL VARIATE 1: TRAITS					
Terms	Degrees of freedom	Type I Sum of Squares	R-Square	F-Ratio	Prob > F
Linear	2	60.604	0.335	45.530	0.0000
Quadratic	2	2.010	0.011	1.510	0.2230
Crossproduct	1	1.249	0.007	1.877	0.1724
Total model	5	63.864	0.353	19.191	0.0000
Variables	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
PCA-Factor1	3	52.936	17.6455	26.513	0.0000
PCA-Factor2	3	2.843	0.9478	1.424	0.2370
Residual	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
Lack of Fit	149	103.443	0.6942	1.369	0.1708
Pure Error	27	13.693	0.5071		
Total Error	176	117.136	0.6655		
CANONICAL VARIATE 2: TRAITS					
Terms	Degrees of freedom	Type I Sum of Squares	R-Square	F-Ratio	Prob > F
Linear	2	3.086	0.017	1.620	0.1966
Quadratic	2	6.038	0.033	3.213	0.0426
Crossproduct	1	6.478	0.036	6.894	0.0094
Total model	5	15.603	0.086	3.321	0.0068
Variables	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
PCA-Factor1	3	10.760	3.5868	3.817	0.0110
PCA-Factor2	3	7.923	2.6409	2.810	0.0410
Residual	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
Lack of Fit	149	149.018	1.0000	1.649	0.0643
Pure Error	27	16.378	0.6066		
Total Error	176	165.397	0.9397		
CANONICAL VARIATE 3: TRAITS					
Terms	Degrees of freedom	Type I Sum of Squares	R-Square	F-Ratio	Prob > F
Linear	2	7.740	0.043	4.073	0.0187
Quadratic	2	5.853	0.032	3.080	0.0484
Crossproduct	1	0.170	0.001	0.18	0.6721
Total model	5	13.764	0.076	2.897	0.0154
Variables	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
PCA-Factor1	3	6.016	2.0000	2.111	0.1006
PCA-Factor2	3	12.159	4.0530	4.266	0.0062
Residual	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
Lack of Fit	149	157.048	1.0540	2.794	0.0013
Pure Error	27	10.187	0.3770		
Total Error	176	167.230	0.9500		

Table 45: *Evaluation of the fitted trend-surface: Relative importance of linear, quadratic and cross-product terms and of independent variables in the canonical model, test of lack of fit of fitted surface: Southwest Oregon*

by temperature which contributed 95% to the total sums of squares. Moisture characteristics, on the other hand, did not show any significant contribution ($p = 0.2370$) to the explained variance of the first trait variate.

The second canonical variate, representing 13% of the total trace, was primarily associated with the plasticity of growth potential (*PC-3*), but also with inherent growth potential (*PC-1*) and

growth timing (*PC-4*). High scores on the second variate thus represent families which were most sensitive (plastic) to test environment. Height growth of these families was negatively impacted by the warm temperature regime. These families also tended to have a higher inherent growth potential and an earlier cessation of growth. The second trait variate was positively and about equally associated with both quadratic terms of *Factor 1* and *Factor 2*. The model was significant ($p = 0.007$) and more or less adequately described the data. Relationships were second order, only weak and highly complex since the interaction term contributed most to the model sums of squares. In comparison with the major pattern revealed in variate 1, the pattern reflected in variate 2 was clearly less important. In tendency, however, families from moist and warm habitats produced greater seedlings, were more negatively impacted by the warm test environment and had an earlier budset than families from drier and cooler locations.

The third trait variate accounted for 11% of total trace. This significant ($p = 0.026$) variate was primarily associated with early development (*PC-2*). High scores on variate 3 thus represent families with a high germination rate and early emergence. Variate 3 was negatively associated with the water balance of the site (*Factor 2*). Both linear and quadratic terms of *Factor 2* significantly contributed to the model sums of squares. The model was significant at $p = 0.0154$; it did not, however, adequately describe the pattern since lack-of-fit was highly significant ($p = 0.0013$). Relationships were weak and comparatively insignificant. The pattern, revealed in variate 3, reflected a weak tendency of families from drier habitats to show a higher germination rate and a faster emergence after sowing than families from moist sites.

The plot of canonical scores on the first and third response vectors, illustrated as mean scores of the seed zones, separated the zones on vector 1 according to the temperature regime (*Figure 50*). In general, zones with mild climate had positive scores on axis 1 while zones with harsher climate (lower *Factor 1*) had negative scores. Since both trait variates 2 and 3 reflected complex and only weak patterns, no clear relationships were apparent in the plot of seed zone mean scores on both these axes (plot of variate 1 versus variate 2 is not shown).

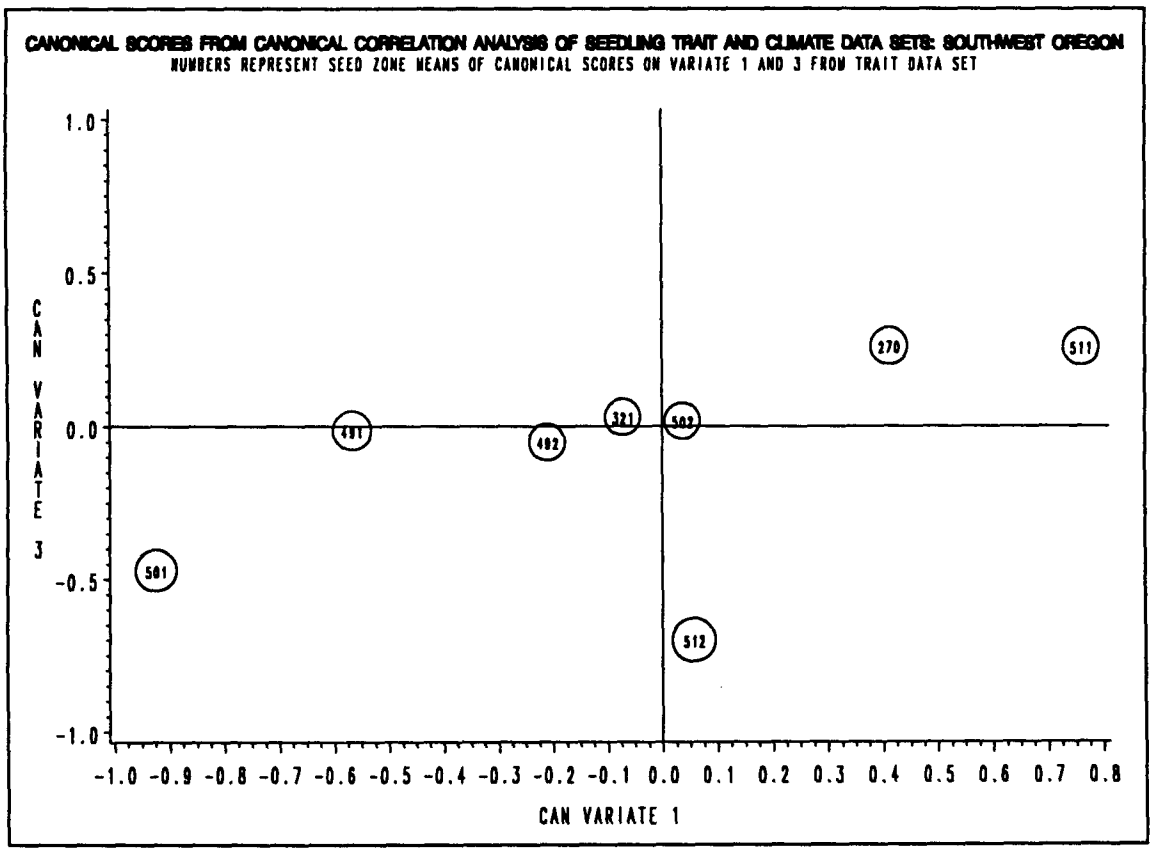


Figure 50: Plot of canonical scores from canonical correlation analysis of seedling traits and climate data set: Southwest Oregon. Trait variate 1 and 3

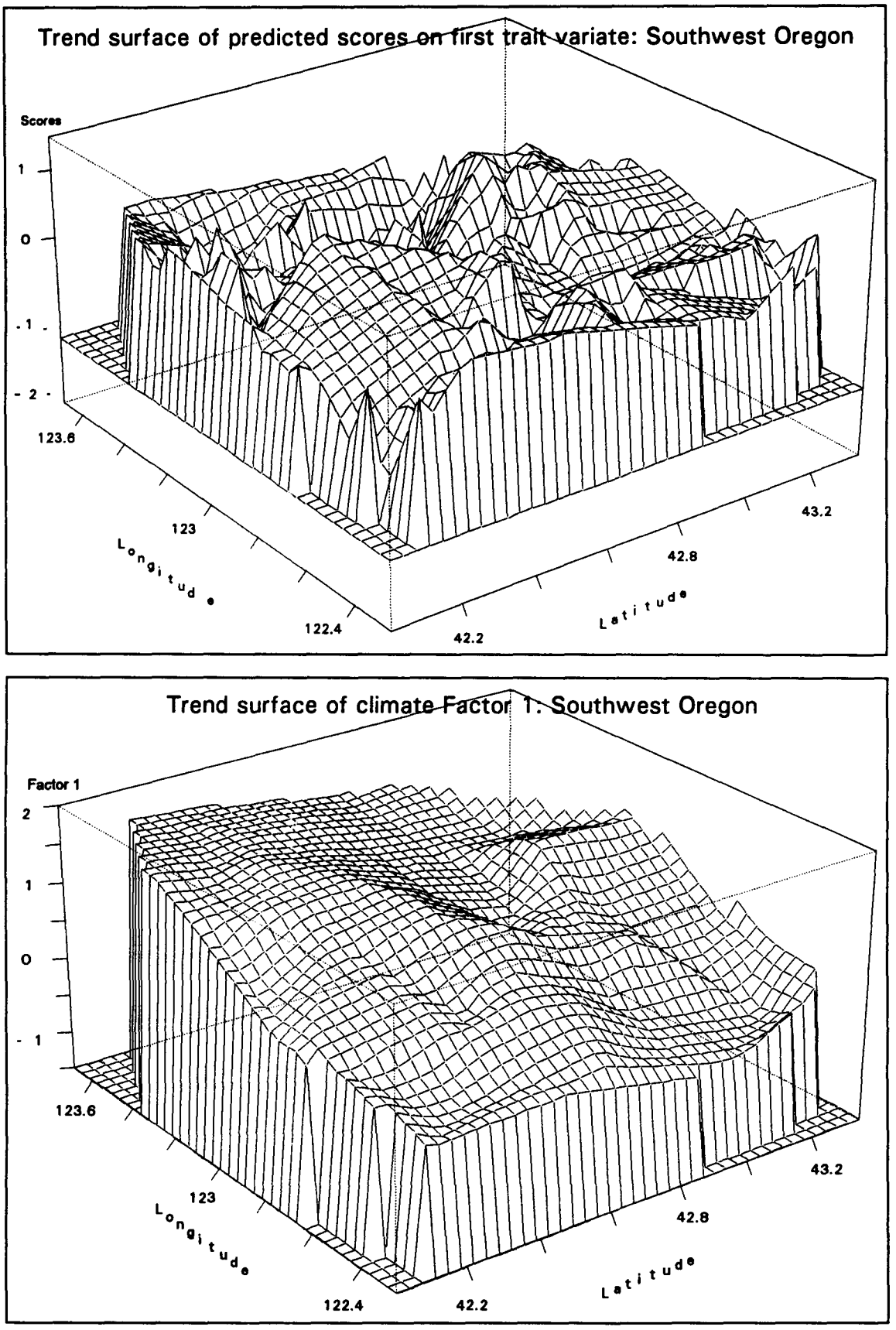


***Figure 51**: Upper: Trend surface of predicted scores on first trait variate for Southwest Oregon. Based on canonical correlation analysis of seedling traits and climate variables as reported in Table 44, p. 189. Lower: Trend surface of actual values of climate Factor 1 (temperature)*

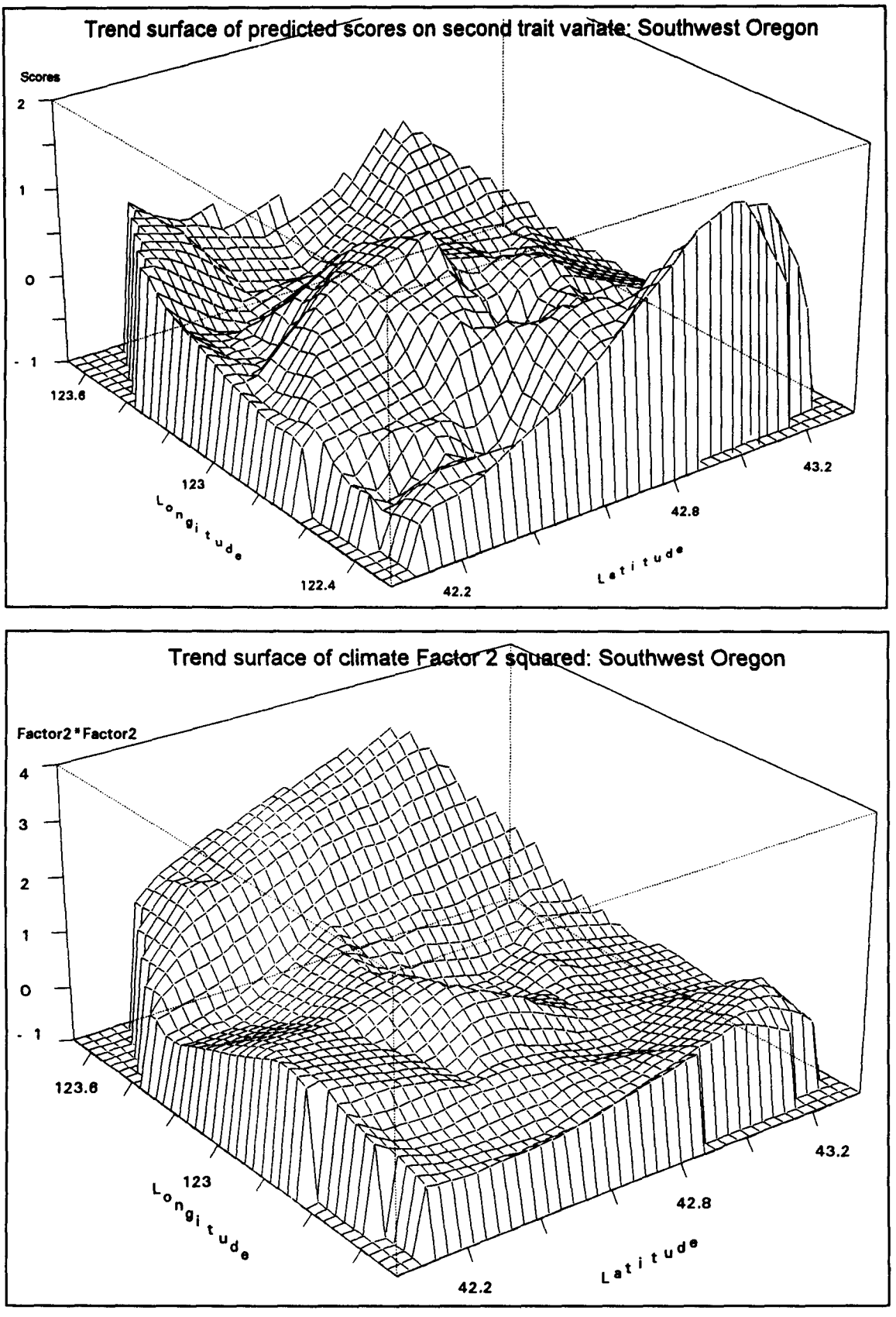


Figure 52: *Upper: Trend surface of predicted scores on second trait variate for Southwest Oregon. Based on canonical correlation analysis of seedling traits and climate variables as reported in Table 44, p. 189. Lower: Trend surface of actual values of climate Factor 2 squared (moisture characteristics of the site)*

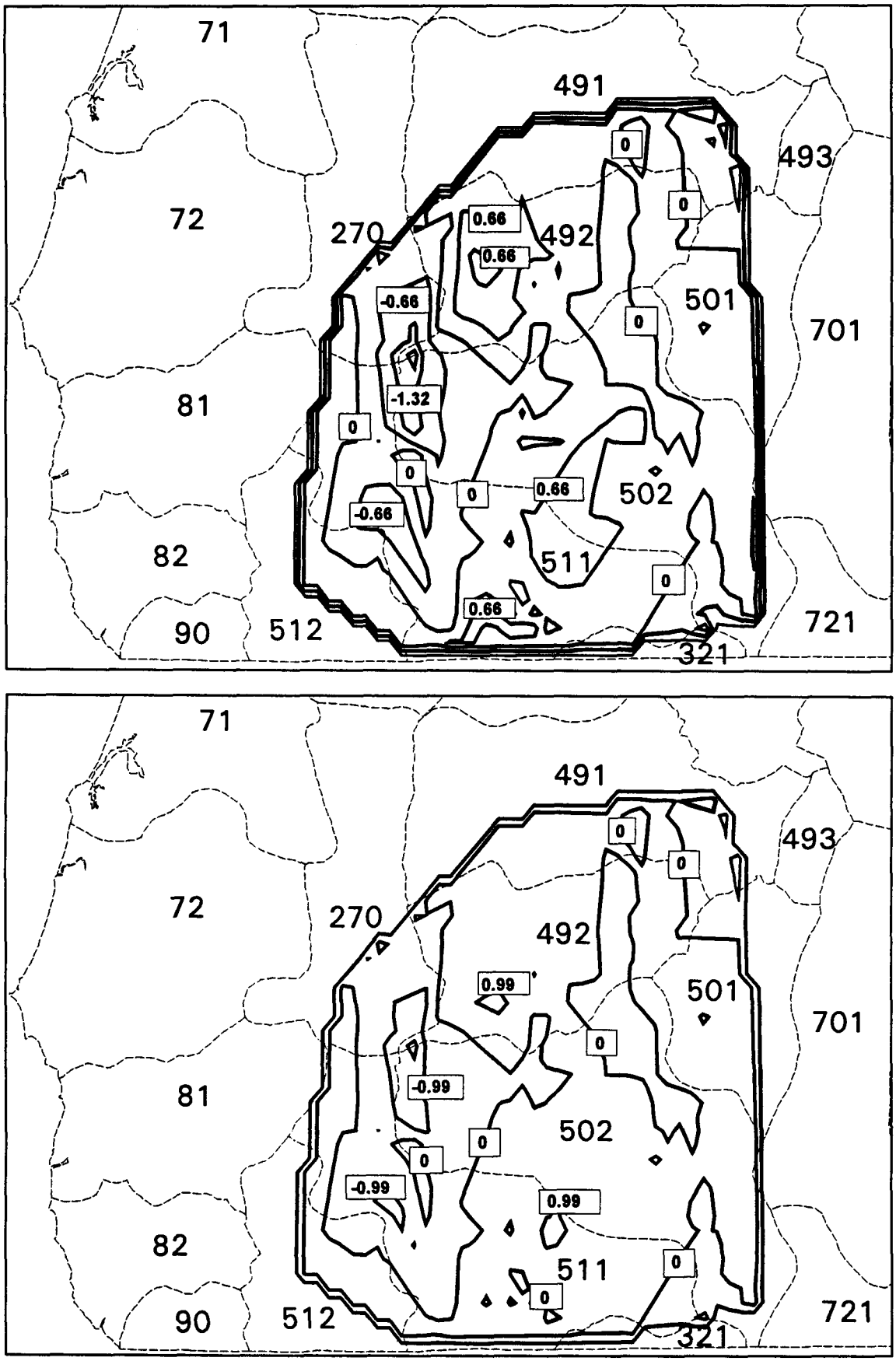


Figure 53: *Maps of relative transfer risk, based on trend surface of predicted scores from first trait variate: Southwest Oregon. Distances between isolines are scaled to represent: Upper: 20% (x=0.66, lower: 30% (x=0.99) relative transfer risk between two isolines*

The trend surface of response variate 1 is illustrated in the upper half of *Figure 51, p. 192.* A rather complex and locally organized pattern of variation was manifest. A weak overall tendency of a west-east oriented pattern is evident. The pattern paralleled the variation in temperature which is characterized by a clinal decrease in the west-east direction (*Figure 51, p. 192, lower half).* Variation, however, was rather patchy, paralleling the clinal variation in temperature only to a certain extent. Discrepancies between scores on the response surface and local temperatures were especially noticeable around longitude 123.2 to 123.4. In this area the scores on variate 1 were clearly lower, *i.e.,* families from these locations had a smaller growth vigor, than expected based on local temperatures.

The trend surface of response variate 2 which was related to both moisture characteristics and temperature, is shown in the upper half of *Figure 52, p. 193.* A highly complex pattern was manifest. The pattern clearly paralleled variation in the second order term of climate *Factor 2* (*Figure 52, p. 193, lower half*) suggesting that this pattern is primarily a reflection of adaptation to moisture characteristics of the site which partly interacts with temperature as indicated in *Table 45, p. 190.* The pattern reflected in variate 2 is only weak, however, accounting for 5 % of the original variation in seedling traits only. The response surface of variate 3 is not shown due to the even minor significance of the pattern.

Since variation of scores on the first response surface is rather complex, exhibiting large changes over very short geographic distances, estimates of mean transfer risks for the seed zones would not be very informative. The complex pattern of transfer risk can best be illustrated by means of contour plots. Such contour maps, illustrating 20% and 30% relative risk among two neighboring contours, are presented in *Figure 53, p. 194.* A relative transfer risk of 20% is equal to a difference of 0.66 units in predicted scores on the first trait variate, a risk of 30%, a difference of 0.99 units. The highly complex, genetic variation pattern illustrated in *Figure 51, p. 192* translates into a highly complex zonation of estimated seed transfer risks. Transfer risks of 30% or even higher are predicted over very short geographic distances. Especially steep gradients of risk are found in the coastal mountains (*i.e.* seed zone 270, west end of zone 511 and zone 512). Areas encompassing the same levels of transfer risk occurred repeatedly in different locations in the sampling area. Risks above 30% occurred over distances of only few kilometers, while transfers over larger distances, for example from the southwestern to the northeastern part of the sampling area, have risks less than 30%. Contours did not follow existing seed zone boundaries; potential transfers could exceed 30 % risk in all zones.

7.1.2 Central Oregon

Results of principal component analysis for Central Oregon, which included 14 seedling traits, are presented in *Table 46, p. 196.* The varimax rotated factor solution with four principal components extracted 92% of total variance contained in the original data set. Final communalities were all high, indicating that the four orthogonal transformations retained between 85% and 96% of the original variance of the single trait variables. The first component (*PC-1*), accounting for 49% of total variance, was related to growth vigor. High scores on this component thus identify families with a high growth potential that is more or less independent of early development. *PC-2* reflected the plasticity of growth potential. High scores on this component thus identify families which are most sensitive (plastic) to test environment. Component 3 was related to budset. High scores on this axis identify families that grow until late in the season. Finally, *PC-4* reflected early development. High scores on this component identify families with a late emergence after sowing.

Results of canonical trend surface analysis, correlating the four principal components to a second order model of climate variables, are furnished in *Table 47, p. 197.* Canonical correlation produced two significant ($p = 0.0001$ and $p = 0.0173$) canonical variates. 50% of the patterned variation in the quantitative traits could be significantly explained by the second order climate model. The first canonical variate, representing 89% of total trace, was associated primarily with growth potential (*PC-1*) and growth timing (*PC-3*). High scores on the first trait variate thus represent a high inherent growth potential combined with a late cessation of growth in autumn.

Variation in the first response variate was primarily and positively associated with temperature. The first canonical vector of the climate model explained 91% of the variance in the linear term

SEEDLING TRAITS	SOLUTION WITH 4 PRINCIPAL COMPONENTS					SOLUTION WITH 3 PRINCIPAL COMPONEN			
	FACTOR PATTERN				Final	FACTOR PATTERN			Final
	PC -1	PC - 2	PC - 3	PC - 4	Communality	PC - 1	PC - 2-	PC - 3	Communality
CEMERG	-0.133	-0.038	0.116	0.936	0.907	-0.243	0.037	0.710	0.564
WEMERG	-0.195	0.063	0.149	0.913	0.898	-0.295	0.133	0.720	0.623
CBDUSET	0.067	-0.110	0.941	0.116	0.915	0.166	-0.176	0.779	0.664
WBUDSET	0.104	0.025	0.939	0.140	0.914	0.200	-0.039	0.794	0.672
CFHT1	0.905	0.129	0.030	-0.123	0.851	0.912	0.122	-0.062	0.850
WFHT1	0.886	-0.239	0.028	-0.171	0.873	0.894	-0.248	-0.096	0.870
CFHT2	0.956	0.204	0.019	-0.065	0.961	0.954	0.203	-0.031	0.953
WFHT2	0.936	-0.227	0.094	-0.098	0.946	0.942	-0.234	0.002	0.941
CFHT3	0.942	0.255	0.028	-0.083	0.960	0.944	0.252	-0.037	0.956
WFHT3	0.936	-0.269	0.076	-0.087	0.961	0.937	-0.274	-0.004	0.953
CDIA3	0.913	0.280	0.074	-0.096	0.927	0.924	0.271	-0.011	0.926
WDIA3	0.903	-0.348	0.071	-0.101	0.953	0.905	-0.353	-0.017	0.945
CFHT3-WFHT3	0.041	0.946	-0.085	0.004	0.905	0.043	0.947	-0.059	0.902
CDIA3-WDIA3	-0.064	0.941	0.000	0.014	0.890	-0.052	0.935	0.010	0.876
Eigenvalue	6.885	2.305	1.840	1.834		7.089	2.331	2.279	
Proportion	49.18	16.46	13.14	13.10		50.64	16.65	16.28	
Cummulative	49.18	65.64	78.78	91.80		50.64	67.29	83.57	

Varimax rotated factor solutions based on correlation matrix of family mean standard deviates of 14 seedling traits [standard normal deviates are used to remove differences among the 2 replications and 2 sowing years]

Abvreviations of variables: C: cold treatment - bed without cover in winter
W: warm treatment - bed covered with plastic tent in winter

EMERG: Emergence of seedlings after sowing [low values mean early, high values mean late emergence]
BUDSET: Date of first year budset [low values mean early, high values mean late budset]
FHT1: Final height at end of first year
FHT2: Final height at end of second year
FHT3: Final height at end of third year
DIA3: Final diamater at end of third year
CFHT3-WFHT3: Difference in final third year height between cold and warm treatment
CDIA3-WDIA3: Difference in final third year diamater between cold and warm treatment

Final communality: Proportion of a variables variance accounted for by the retained principal components

The principal components represent:

FACTOR 1: Growth vigor
FACTOR 2: Plasticity of growth potential
FACTOR 3: Growth timing
FACTOR 4: Early development [Germination and emergence]

***Table 46**: Principal component analysis of seedling traits: Central Oregon. Varimax rotated factor solutions based on family mean standard normal deviates of 8 seedling traits from 2 replications, 2 sowing years and two different environments (treatments). The different treatments are included as separate variables (Prefix C and W). Sample size: N= 216 families*

and 46% in the quadratic term of *Factor 1*. The first response variate shared 45% or 22% of variation with the linear or the quadratic term of *Factor 1*.

Hence, under the climate conditions at the test site in Corvallis, families from warm and mild environments had a high inherent growth potential and grew later in the season, while families from locations with cold and harsh climate set their buds earlier and produced smaller seedlings. The relationship between growth vigor and temperature appeared to be inherently (gene-

SEEDLING TRAITS PC - FACTORS	SEEDLING TRAIT DATA SET					
	CANONICAL VARIATE 1			CANONICAL VARIATE 2		
	Struc.	Cross.	Red. %	Struc.	Cross.	Red. %
PC - 1	0.9052	0.6330	40.06	0.0954	0.0260	0.07
PC - 2	-0.0212	-0.0149	0.02	-0.5392	-0.1470	2.16
PC - 3	0.4060	0.2839	8.06	-0.4546	-0.1240	1.54
PC - 4	0.1235	0.0863	0.74	0.7025	0.1916	3.67
Eigenvalue	0.9595			0.0797		
% trace (pt)	88.87			7.40		
unb. can. corr.	0.6882			0.2229		
r square	0.4736			0.0497		
Significance	0.0001			0.0173		
H-L-trace (Tr)	1.0796					
Model variance	51.91					
Sign. variance	49.98					

Dependent data set:
4 Principal components representing 14 seedling traits

Dependent data set:
Full second order model with climate Factor 1 and 2

CLIMATE PCA FACTORS	CLIMATE DATA SET			
	VARIATE 1		VARIATE 2	
	Struc	% var	Struc	% var
Factor 1	0.9541	91.00	-0.2715	7.37
Factor 2	0.1519	2.31	0.9859	97.20
Factor 1 square	-0.6750	45.56	0.3812	14.53
Factor 2 square	0.0917	0.85	0.3746	14.03
F1 X F2	0.0638	0.41	-0.7991	63.86

Struc: Structure coefficient
Cross: Crossloadings
Red %: Percent redundancy
% var: Percent variance explained
H-L-trace: Hotelling-Lawley-trace

Table 47: *Canonical correlation analysis: Seedling traits and climate variables. Central Oregon. [Sample size: N=216 families]*

tically) fixed because *PC-2*, representing the difference in the expression of growth potential in the two different environments, was completely unrelated to the first response variate.

The canonical model was highly significant ($p = 0.0000$) and adequately described the surface since lack-of-fit was non significant ($p = 0.0802$). Results furnished in *Table 48, p. 198* indicate that relationships were primarily linear. Temperature clearly dominated the model, contributing 91.5% to the total sum of squares. Moisture characteristics of the site contributed only a small (7.5%), but highly significant ($p = 0.0006$) proportion to the model sum of squares.

The second canonical response variate, representing 7.4% of total trace, was positively associated with early development (*PC-4*), and negatively related to plasticity growth potential (*PC-2*) and growth timing (*PC-3*). Associations with the climate model were linear, focusing primarily on moisture conditions of the site. As indicated by the redundancies, relationships were weak, however. Moreover, lack-of-fit was significant at $p = 0.0265$, indicating that the climate model did not adequately describe the pattern (*Table 48, p. 198*). In comparison with the pattern revealed in variate 1, the second variate was clearly insignificant, reflecting a rather weak, moisture-induced relationship. At the test site in Corvallis, families from dry sites showed a tendency to emerge earlier, to set bud later and to respond in a more plastic manner to the warmer climate during winter than families from locations with a more favorable water regime.

The plot of canonical scores on the first two response vectors, illustrated as seed zone mean scores, is shown in *Figure 54, p. 199.* With few exceptions, seed zones were aligned on canonical variate 1 according to their mean thermal conditions. Zones with colder conditions were separated from zones with warmer climate. Notable exceptions were the seed zones 675, 673, 952 and 751. Separation on axis 2, however, was non conclusive.

The trend surface of predicted scores on the first trait variate is shown in the upper half of *Figure 55, p. 200.* The general pattern of adaptive variation resembled a double cline running in

CANONICAL VARIATE 1: TRAITS					
Terms	Degrees of freedom	Type I Sum of Squares	R-Square	F-Ratio	Prob > F
Linear	2	103.457	0.481	98.851	0.0000
Quadratic	2	0.773	0.004	0.739	0.4790
Crossproduct	1	0.876	0.004	1.676	0.1969
Total model	5	105.107	0.489	40.171	0.0000
Variables	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
PCA-Factor1	3	102.512	34.1708	65.299	0.0000
PCA-Factor2	3	9.478	3.1594	6.038	0.0006
Residual	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
Lack of Fit	162	90.883	0.561	1.4165	0.0802
Pure Error	48	19.010	0.396		
Total Error	210	109.893	0.5232		
CANONICAL VARIATE 2: TRAITS					
Terms	Degrees of freedom	Type I Sum of Squares	R-Square	F-Ratio	Prob > F
Linear	2	15.830	0.074	8.353	0.0003
Quadratic	2	0.088	0.000	0.046	0.9548
Crossproduct	1	0.006	0.000	0.006	0.9369
Total model	5	15.930	0.074	3.361	0.0061
Variables	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
PCA-Factor1	3	0.456	0.1517	0.16	0.9230
PCA-Factor2	3	13.564	4.521	4.77	0.0031
Residual	Degrees of freedom	Sum of Squares	Mean-Square	F-Ratio	Prob > F
Lack of Fit	162	168.291	1.038	1.62	0.0265
Pure Error	48	30.778	0.6412		
Total Error	210	199.069	0.9479		

Table 48: *Evaluation of the fitted trend-surface: Relative importance of linear, quadratic and cross-product terms and of independent variables in the canonical model, test for lack of fit of fitted surface: Central Oregon*

opposite direction from north to south, resulting in minimum values in the middle section of the Central Oregon sampling area. The trend surface of predicted scores on the first trait variate, representing variation in growth vigor, strongly resembled the variation pattern of climate *Factor 1*, emphasizing the strong association of growth vigor with temperature at source locations. With few exceptions, the surface of predicted scores paralleled the variation of temperature in great detail (*Figure 55, p. 200, lower half*). Lower growth vigor than expected, based on temperatures, were apparent for families from the southeastern corner of the sampling area (zones 751, 952 and 712). In contrast, higher inherent growth potentials than expected were observed for families of seed zones 675 and 673.

Since variate 2 accounted for only 3.5 % of original variation in seedling traits, no trend surface is shown for this very minor pattern of variation.

Estimated relative seed transfer risks among all seed zones, based on seed zone mean scores on the first response variate, are furnished in *Table 49, p. 201*. Zones with few families were joined for the analysis as follows: Zones 713 with 711, 902 with 672 and 901 with 921. Transfer risks varied in a wide range between 1% and 71%. Risks were especially high when transferring families from the Columbia River (zones 42 and 661) to central - eastern Oregon or vice versa.

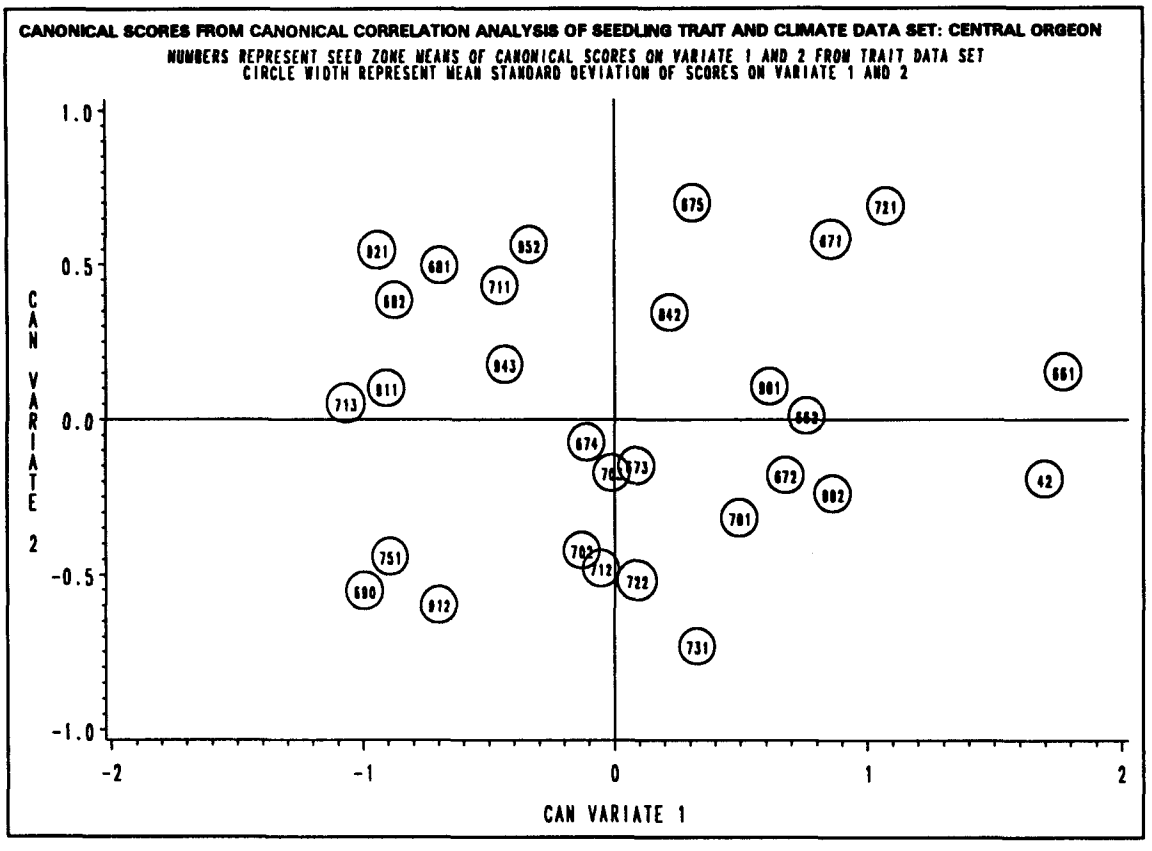


***Figure 54**: Plot of canonical scores from canonical correlation analysis of seedling traits and climate data set: Central Oregon. Trait variate 1 and variate 2*

Average linkage cluster analysis on the matrix of estimated transfer risks among zones, as given in *Table 49, p. 201*, produced the hierarchical grouping illustrated in *Figure 56, p. 202.* Four main clusters, separated by an average transfer risk of more than 20%, were manifest. Seed zone 661 was again very distinct from the rest of the area, indicating that the milder climate in this area (maritime influence) is reflected in the response of families growing in this seed zone. Moving sources from the Columbia River to plantation sites on the east side of the Cascade Range resulted in a high average transfer risk of more than 50%. Risks were as high as 70% when families were moved to sites in central eastern Oregon. Also short distance transfers to plantation sites in the adjoining zones 662 and 671 resulted in a risk of about 30%. A relatively high average transfer risk of about 30% was estimated for transfers from central zones to northern and southern sites and vice versa. An additional, quite distinct area was discernible in the northwest, comprising the seed zones 662, 671, 672 and probably 902 (the few individuals of zone 902 were joined with zone 672 for the analysis, they may, however, be more similar to families of zones 842, 901, 921 or 673).

Estimated relative seed transfer risks are also illustrated in the maps presented in *Figure 57, p. 203*, showing isolines which are scaled to represent 20% and 30% of relative transfer risk between two neighboring isolines. The scaled contours delineated 7 major zones within which sources can be transferred with an estimated risk below 20% (*Figure 57, p. 203, left hand side*).

Isolines roughly paralleled the directions of the seed zone boundaries. In the northern half of the area they ran in a southwest-northeast direction while in the southern part they followed a northwest-southeast direction, thus delineating a rather large area in the center of eastern Oregon, that encompassed seed zones 911, 912, 690, 682, 711, 713 and 952. As a result of the adaptation to the cold and harsh climate, the sources from this area showed a low growth potential. The lowest growth vigor was manifest in a small zone around Silver Lake (east side of zone 713).

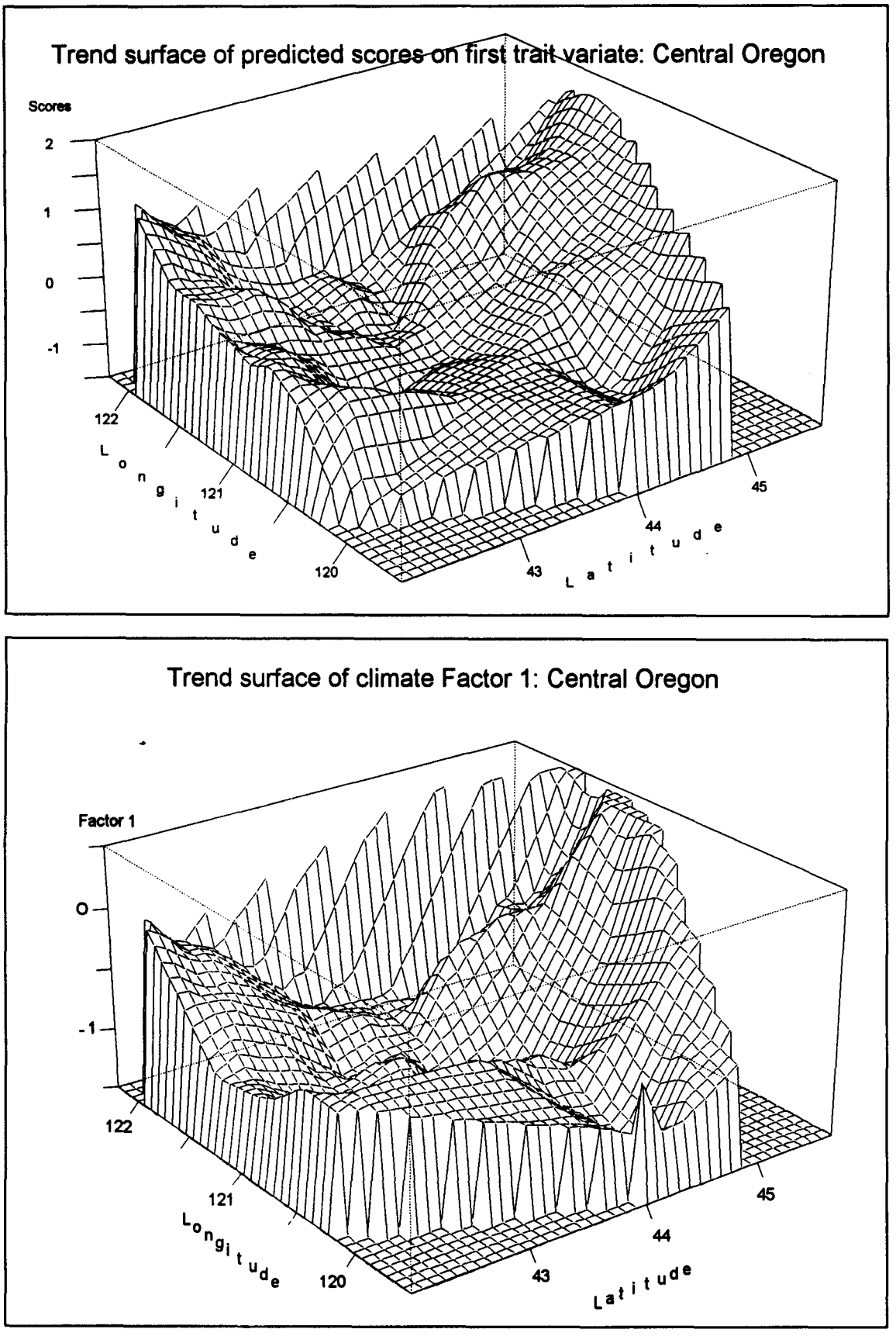


Figure 55: *Upper: Trend surface of predicted scores on first trait variate for Central Oregon. Based on canonical correlation analysis of seedling traits and climate variables as reported in Table 47. Lower: Trend surface of actual values of climate Factor 1*

ZONE	661	662	671	672	673	674	675	681	682	690	701	702	703
661	0.00												
662	0.30	0.00											
671	0.27	0.03	0.00										
672	0.30	0.01	0.04	0.00									
673	0.48	0.20	0.23	0.20	0.00								
674	0.52	0.26	0.29	0.25	0.06	0.00							
675	0.42	0.14	0.16	0.13	0.07	0.13	0.00						
681	0.65	0.42	0.44	0.41	0.23	0.18	0.30	0.00					
682	0.68	0.47	0.49	0.46	0.29	0.23	0.35	0.05	0.00				
690	0.71	0.50	0.52	0.49	0.32	0.26	0.38	0.09	0.04	0.00			
701	0.37	0.08	0.11	0.07	0.12	0.18	0.06	0.35	0.40	0.43	0.00		
702	0.53	0.26	0.29	0.26	0.07	0.01	0.13	0.17	0.22	0.26	0.19	0.00	
703	0.50	0.23	0.26	0.22	0.03	0.03	0.10	0.21	0.26	0.29	0.15	0.04	0.00
711	0.64	0.41	0.43	0.40	0.22	0.16	0.28	0.02	0.07	0.11	0.33	0.15	0.19
712	0.51	0.24	0.27	0.24	0.04	0.02	0.11	0.19	0.24	0.28	0.17	0.02	0.01
721	0.21	0.09	0.07	0.10	0.29	0.35	0.23	0.50	0.54	0.57	0.17	0.35	0.32
722	0.48	0.20	0.23	0.20	0.00	0.06	0.07	0.23	0.29	0.32	0.12	0.07	0.03
731	0.42	0.13	0.16	0.12	0.07	0.13	0.01	0.30	0.35	0.39	0.05	0.14	0.10
751	0.69	0.47	0.49	0.47	0.29	0.23	0.35	0.06	0.01	0.03	0.40	0.23	0.26
842	0.42	0.14	0.17	0.13	0.06	0.12	0.00	0.29	0.34	0.38	0.06	0.13	0.09
911	0.69	0.47	0.50	0.47	0.30	0.24	0.36	0.06	0.01	0.03	0.41	0.23	0.27
912	0.65	0.42	0.45	0.42	0.24	0.18	0.30	0.00	0.05	0.09	0.35	0.17	0.21
921	0.70	0.48	0.51	0.48	0.30	0.25	0.37	0.07	0.02	0.02	0.42	0.24	0.28
943	0.60	0.35	0.38	0.35	0.16	0.10	0.22	0.08	0.13	0.17	0.28	0.09	0.13
952	0.58	0.32	0.35	0.32	0.13	0.07	0.19	0.11	0.16	0.20	0.25	0.06	0.10

ZONE	703	711	712	721	722	731	751	842	911	912	921	943	952
703	0.00												
711	0.19	0.00											
712	0.01	0.18	0.00										
721	0.32	0.49	0.33	0.00									
722	0.03	0.22	0.04	0.29	0.00								
731	0.10	0.29	0.12	0.22	0.07	0.00							
751	0.26	0.08	0.25	0.55	0.29	0.36	0.00						
842	0.09	0.28	0.11	0.23	0.06	0.01	0.35	0.00					
911	0.27	0.08	0.25	0.55	0.30	0.36	0.00	0.35	0.00				
912	0.21	0.02	0.19	0.50	0.24	0.30	0.06	0.30	0.06	0.00			
921	0.28	0.09	0.26	0.56	0.30	0.37	0.01	0.36	0.01	0.07	0.00		
943	0.13	0.06	0.12	0.43	0.16	0.23	0.14	0.22	0.14	0.08	0.15	0.00	
952	0.10	0.09	0.09	0.41	0.13	0.20	0.17	0.19	0.17	0.11	0.18	0.03	0.00

Table 49: *Relative seed-transfer risks among seed zones, based on canonical scores from first trait variate: Central Oregon*

Five major zones were delineated by the isolines scaled to represent 30% of estimated transfer risk among two neighboring contours (*Figure 57, p. 203, right hand side*). The distinct zone around Silver Lake, exhibiting the lowest scores on the first response variate, was enlarged in the northwest direction, encompassing now the three zones 690, 682 and 713. A small area in zone 911 also showed sources with a similar adaptive response.

7.2 Discussion

Much of the collective variation in seedling traits could be associated with climate variables of source locations. As much as 50 % of the patterned variation in growth and phenology traits could be explained. Moreover, patterns were reasonable, given the assumption that they reflect adaptation to the environment. These two observation are strong evidence that much of the variation in seedling traits reflect adaptation to local climate. Adaptive variance of single traits tended to be intercorrelated, forming co-adapted trait combinations. Temperature at source locations was the most important factor shaping the major patterns of adaptive variance. The moisture characteristics of the site, on the other hand, had only a minor influence.

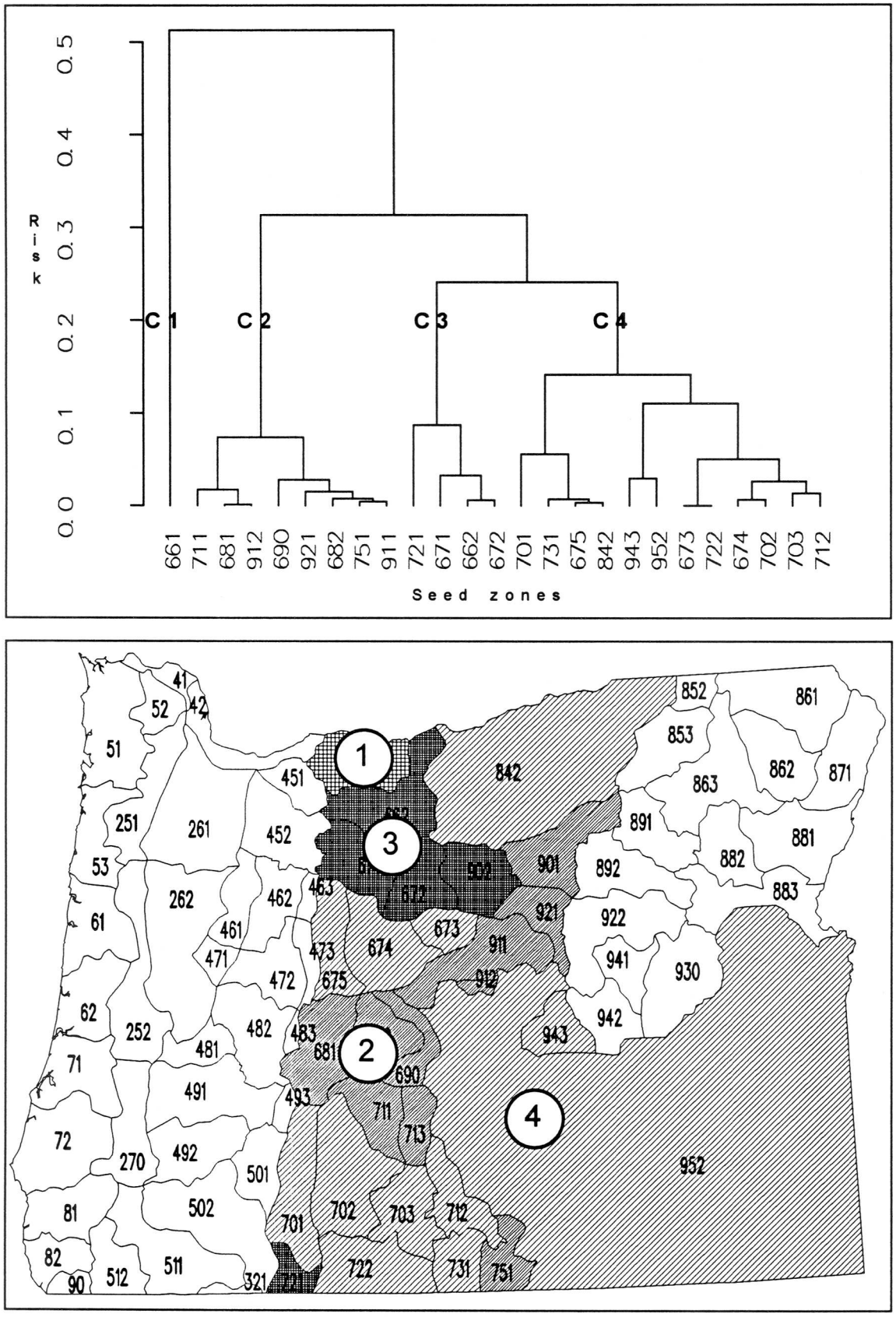


Figure 56: *Group formation of similar zones, based on transfer risks among zones: Central Oregon. Dendrogram resulting from average linkage cluster analysis based on matrix of transfer risks among zones (see Table 49). Lower: Mapped groups of seed zones with an average transfer risk of less than 20% within groups*

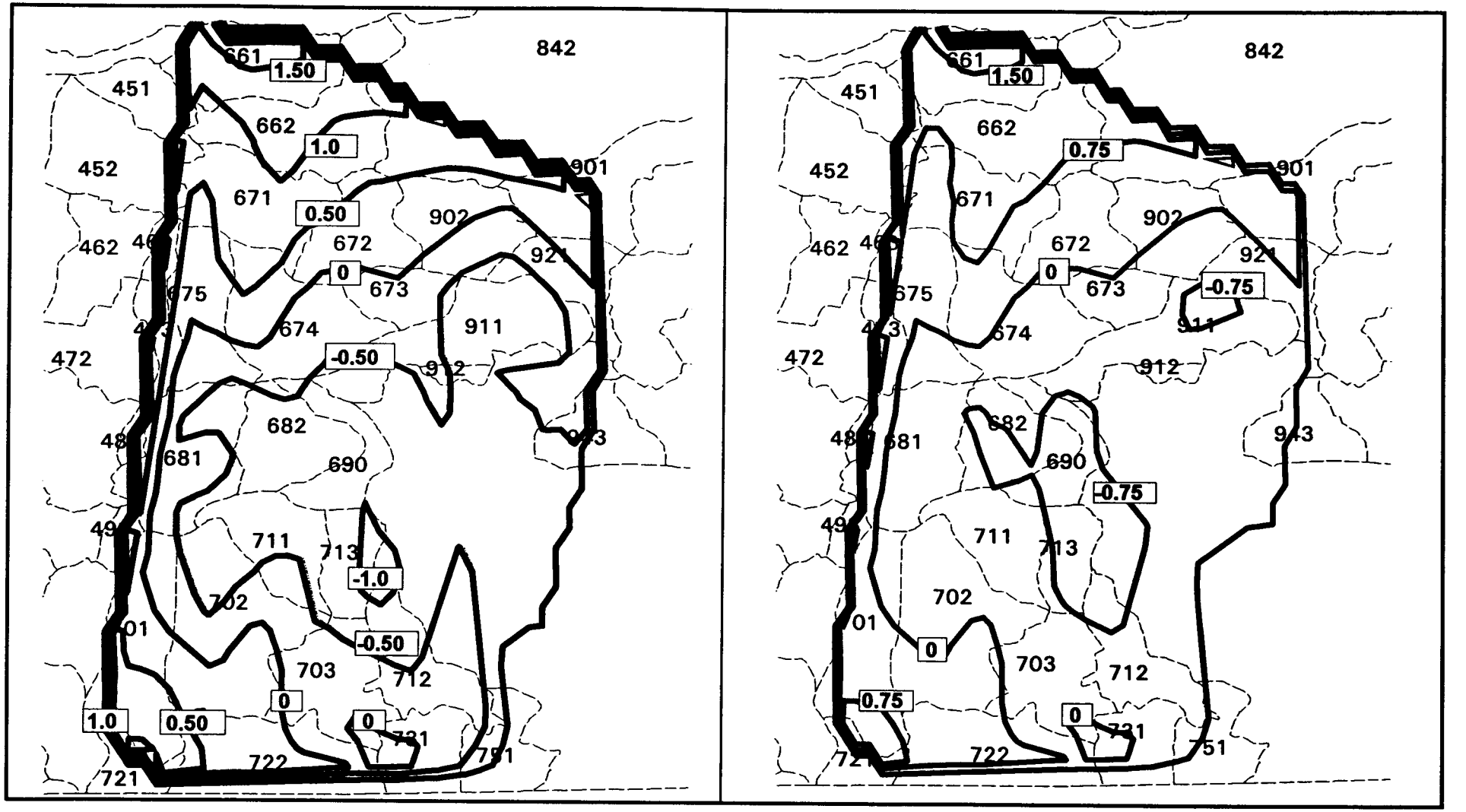


Figure 57: Maps of relative transfer risk, based on trend surface of predicted scores on first trait variate: Central Oregon. Distances between the isolines are scaled to represent: Left: 20% (x=0.50), right: 30% (x=0.75) of relative transfer risk between isolines

The major pattern of adaptive variance reflected primarily relationships of growth vigor with thermal conditions at source locations. About 24 % of variance in growth vigor of southwestern Oregon sources were redundant with variation in temperature. In Central Oregon, growth potential and temperature shared about 40 % of their variance.

Simple climatic or elevational clines in growth vigor have been found for many different tree species. Decreasing height growth with increase in seed source elevation is a common finding in forest tree genetic studies conducted on low-elevation sites (WRIGHT, 1976; HERMANN and LAVENDER, 1986; ROW and CHING, 1973; REHFELDT, 1974). Adaptive patterns in growth vigor of Ponderosa pine have been described by several authors, for example by CALLAHAM and LIDDICOET (1961), SQUILLACE and SILEN (1962), ECHOLS and CONKLE (1971), CONKLE (1973), MADSEN and BLAKE (1977), MITTON et al. (1977), REHFELDT (1980, 1984, 1986a, 1986b, 1990a, 1993) and SORENSEN (1994).

SQUILLACE and SILEN (1962) reported results of 30-year old provenance trials including 18 seed sources of ponderosa pine planted on 6 different test locations in Oregon, Washington and Idaho. Correlations between tree height and altitude of the seed source were, across the different test sites, between $R = 0.45$ and $R = 0.78$. Correlations with annual temperature ranged from $R = 0.26$ to $R = 0.75$ and for spring temperature in April and May from $R = 0.37$ to $R = 0.82$. Interestingly, correlations of the same order of magnitude were also found for precipitation. However, this was most likely caused by an indirect effect since temperature and moisture gradients are often highly redundant in the studied area.

ECHOLS and CONKLE (1971) described a rank correlation of $R = 0.86$ between 16-year height and parent tree elevation in a provenance trial of ponderosa pine in the Sierra Nevada in California.

A clinal association between second year height, measured in the common garden, and temperature at seed source location was reported for ponderosa pine sources from the northern Rocky Mountains (eastern Washington, northern Idaho and western Montana) by MADSEN and BLAKE (1977). A number of characters were strongly related to environmental parameters at seed source. The highest correlation ($R = 0.602$) was found between height growth and elevation. Second year height and temperature showed a correlation of $R = -0.261$, while no correlation was found with moisture ($R = 0.001$). In multiple regression analyses, the strongest relationship was found with 2-year height growth, in which elevation, latitude and longitude explained 44% of the variation.

In a family evaluation test of 37 ponderosa pine populations from southern Idaho, REHFELDT (1980) reported that 51% of the variance in 16-year height could be explained by a multiple regression model including seed source elevation, latitude and longitude.

A strong relationship between growth potential of 4-year old seedlings, assessed in a common garden, and elevation was reported for ponderosa pine sources in the upper Colorado river basin by REHFELDT (1990a). The best model described genetic variation in growth vigor (represented as a principal component reflecting several growth traits expressed in different environments) along elevational and geographic clines, explaining 74% of the variance. For growth potential alone, the coefficient of determination in the multiple regression model was $R^2 = 0.67$.

A weaker association between growth vigor and elevation of seed source, however, was found for ponderosa pine in the northern Rocky Mountains where elevation accounted for only 32% of the variance in growth vigor. This result contrasts with Douglas-fir (*Pseudotsuga menziesii var. glauca Franco*) and lodgpole pine (*Pinus contorta Dougl. ex. Loud.*) in the same region, where elevation accounted for 72 % and 66 % of seedling growth vigor, respectively. When the geographic variables latitude and longitude were added to the ponderosa pine multiple regression, the model explained 61% of the variance in growth vigor (REHFELDT, 1984).

In central Idaho ponderosa pine, 56% of the variation in third year height, measured in a common garden test, could be explained by a multiple regression model including 13 independent geographic variables and elevation (REHFELDT, 1986a).

In the middle Columbia River system (Montana, Idaho and northern Washington), REHFELDT (1986b) found a coefficient of determination of $R^2 = 0.52$ for growth vigor of ponderosa pine when elevation combined with geographic origin of the seed were independent variables.

Adaptive variance of growth vigor in Ponderosa pine is thus well documented. All the reported results showed patterns of variation which were clearly related to independent variables that serve as surrogates for the complex environmental gradients that influence natural selection. Elevational clines were the most prominent.

Our results clearly reflect the same major pattern of adaptive variance in growth potential. Growth vigor appears to be the result of adaptation to temperature constraints at source location. Apparently, growth potential is strongly selected by temperature. Since elongation potential seems to be genetically "fixed" if temperature is the selective agent, as hypothesized by SORENSEN (1994) and supported by our results, selection most likely acts on phenology traits, adapting the sources to the length of the growing season at seed source location. The term "fixed" is used by SORENSEN in the sense that seedlings from sites with a short growing season because of low temperature constraints have less ability to respond to favorable growing conditions (fixed reaction). In contrast, families from locations where growing season is short because of moisture deficits will respond to favorable growing conditions with increased growth (plastic reaction). Families from colder habitats have a genetically fixed vegetation cycle; they set their buds early regardless of the temperature conditions of the common garden and thus are unable to respond to a longer growing season on more favorable lowland sites. In contrast, families from warmer habitats are adapted to a long growing season, grow until late in the season and thus produce larger seedlings. Adaptive variance in our sampling area was the same order of magnitude as that which has been reported for many other forest tree species. Likewise, our results are highly concordant with the results reported for Ponderosa pine in the adjoining areas in Washington, Idaho and Montana, as indicated.

In Southwest Oregon, the association of growth vigor with temperature at seed source location was rather weak (24 % redundant variance) compared to Central Oregon as well as to results found in other areas. This comparatively weak adaptive variance may be due to the rather restricted sampling area, spanning only 155 km in north-south and 145 km in east-west directions. Even if Southwest Oregon has a rather extreme environmental heterogeneity (FRANKLIN and DYRNESS, 1973) caused by the rugged topography, climate heterogeneity is smaller in the sampled area than in Central Oregon (*see section 3.5.2*). Temperatures are rather mild, especially in winter, and lack the extremes found east of the Cascades. Consequently, it may be hypothesized that selection pressure caused by temperature constraints is comparatively weaker in this area. As has been discussed already in *section 6.5*, rather weak patterns of variation in quantitative traits as well as in allozyme markers have been reported for several other species in this area. Moreover, factors other than temperature were found to be involved in adaptation processes. Especially moisture conditions of the site and, possibly, soil properties seem to be of selective significance (FURNIER and ADAMS, 1986; DYRNESS and YOUNGBERG, 1966; JENKINSON, 1974; CAMPBELL and SUGANO, 1987).

Although the major pattern reflected in the first response variate in Southwest Oregon was clearly dominated by temperature, and moisture did not show any significant influence overall, a locally modifying interaction of moisture seems to exist on very dry sites. In the plot of the first trend surface (*Figure 51, p. 192*), a smaller growth potential than expected based on local temperatures was apparent in a north-south extending area around longitude 123.2 to 123.4. Since it is exactly in this part of the sampling area that precipitation is at a minimum, a modifying influence of moisture seems likely.

Even though second and third response variates accounted for only small amounts of original variation in seedling traits of southwestern families, an interpretation may still be interesting since both patterns seem to reflect an adaptive response to moisture conditions. Complex second order relationships involving temperature and moisture conditions were reflected in both second and third response variates. Since temperature and moisture gradients parallel each other in Southwest Oregon (AGER and STETTLER, 1983; WELLS, 1964b), complex interactions and indirect effects are expected. The second response variate related growth potential, plasticity of growth potential and growth timing to a complex model of temperature,

moisture and their interaction. Families from warm and moist habitats produced large seedlings, their growth potential, however, was plastic *i.e.* differed between the two test environments. The second response variate was significantly related to both temperature and moisture conditions of the site. Since growth potential appears to be inherently fixed if temperature is the selective agent, we hypothesize that the different expression of growth potential in the two test environments primarily reflects a moisture related adaptive response. In the warm test environment, families from moister sites showed a tendency to set their buds earlier and consequently to produce smaller seedlings compared to the cold test environment while the expression of these traits did not differ between the two environments for families from drier sites. Since the warm test environment was achieved by covering the nursery beds with tents during winter, we may anticipate not only a change in the temperature regime in this treatment but most likely - as an indirect effect - also a substantial change in the water balance because winter and early spring rain was precluded from entering the soil in this treatment. Precipitation during winter and spring, however, make up the most important part of total precipitation in Southwest Oregon and are thus expected to be of great importance for the water balance of the sites in spring and early summer. We therefore believe that the second response variate portrays primarily an adaptive response to the moisture conditions of the site during early shoot elongation and not a temperature related response.

REHFELDT (1986a, 1986b, 1990a, 1993) has conducted a number of common garden studies with Ponderosa pine from the Rocky Mountains in which moisture stress was applied in one of the test environments. Although seedlings responded in a plastic manner to the water stress with reduced growth, in no case was there evidence of a genetic response that could be interpreted as adaptation to the moisture conditions at source location. Several factors could have contributed to the disparity between our results and those reported by REHFELDT. We have already argued in *section 6.5* that moisture most likely plays a more important role as a selective agent in Southwest Oregon compared to the Rocky Mountains because moisture gradients are extremely steep over short geographic distances and soil conditions are very distinct due to the geologic substrate. On the ultramafic soils, a reduced water uptake of the plants may lead to an insufficient uptake of important minerals or may even lead to toxic effects in the concentrated soil solution. We have already cited much supporting evidence for this assumption in *section 6.5*. In addition, a moisture-related adaptive response has not only been reported for other species in this area (FURNIER and ADAMS, 1986; DYRNESS and YOUNGBERG, 1966; CAMPELL and SUGANO, 1987), it was also reflected in the allozyme patterns as we have demonstrated in *section 6*.

A weak adaptive response to the moisture conditions of the site was also reflected in the third response variate. In contrast to the second pattern which reflects an adaptive response to the moisture conditions during early shoot development, the third pattern seems to portray an adaptive response to the early occurrence of summer droughts since families from drier sites showed a tendency to have a higher germination rate and a faster emergence compared to families from moist habitats. Although the model did not adequately describe the pattern of variation, since lack-of-fit was significant and the relationships were only weak, the pattern was meaningful and clearly in line with expectations and reported results. Rapid and uniform germination is considered adaptive on sites where early summer droughts occur and where growth has to be completed before the onset of the unfavorable growing conditions (CAMPBELL and RITLAND, 1982; WEBER, 1988). On such sites, a fast and vigorous early development may be essential for survival of seedlings.

In conclusion, patterns of adaptive variation in Southwest Oregon were rather complex. Growth potential and growth timing were clearly but comparatively weakly associated with temperature at source location. A rather small-scale, ecotypically organized pattern was evident. This ecotypic pattern of adaptive variance seems to be a result of rather complex interactions of both temperature and precipitation. Steep moisture gradients over short distances seem to be partly reflected in adaptive response of the sources. Even if temperature is clearly the most important selective agent, about 25% of the total trace (second and third variates) appeared to be related primarily to moisture conditions of the site.

Growth vigor and growth timing of sources in Central Oregon were strongly related to temperature. Of the variation in growth potential, 40% were redundant with climate variate 1 which was

primarily respresenting temperature (*Table 47, p. 197*). Growth potential was genetically fixed since plasticity of growth was unrelated to the first response variate. Our results thus clearly support SORENSEN' s two hypothesis that elongation potential is strongly selected by temperature and that growth potential is genetically fixed if temperature is the selective agent. SORENSEN, who analyzed exactly the same data which we used for our investigation, was able to explain 69% of the variation in factor scores of the first principal component (which primarily reflected growth potential) by a complex model of location variables. This comparatively high association may be due partly to overfitting of the model, since 35 location terms were included. Partly, however, it is certainly an indication for the important influence of topographic variables such as aspect and slope. Although slope and aspect explained only 8% of the model sums of squares, their interaction with other location variables contributed 47% to the total sum of squares. Hence, both aspect and slope obviously play an important, modifying role for local climate conditions. Because our estimates of climate conditions at source location included neither slope nor aspect (due to a lack of information), the temperature-related adaptive variance in growth potential is most likely higher in reality than the 40 % that was revealed in our model.

Although we have not used exactly the same seedling traits, have applied a different statistical technique, and have used climate instead of location variables as independent terms in the model, the major pattern of adaptive variance in growth vigor is nearly identical to the pattern of growth vigor (represented as the first principal component) published by SORENSEN (1994). Highest growth vigor, represented as " * " symbols in the plot of SORENSEN (*Figure 58, p. 208, right hand side*) correspond to our highest canonical scores between 1.0 and 1.5. The " + " values correspond to canonical scores between 0 and 1.0, the " 0 " values correspond to moderately negative canonical scores between 0 and -0.5 while the " - " and the " = " symbols match our moderately to high negative scores below -0.5.

Although elevation seemed to dominate geographic differentiation, there were three notable exceptions to the general elevational pattern in SORENSEN's investigation. All three of them were also observable in our contour map (*Figure 58, p. 208, left hand side*). In the extreme south, SORENSEN observed sources from high-elevation locations which produced relatively vigorous seedlings. As suggested by him, this comparatively high growth potential seems to be a result of adaptation to the milder, more oceanic climate in this area (a consequence of the Klamath River cutting through the Cascade Range) since canonical scores on our first response surface paralleled temperature variation in great detail. The second anomalous area, situated between 43 ° 10' N and 43 ° 50' N (north part of seed zone 702), with a growth vigor lower than expected based on elevation, had also low canonical scores on our response surface, clearly paralleling the harsh temperature conditions in this area. Finally, in the southeastern corner of the sampling area, in the Warner Mountains in Oregon, SORENSEN found a large within-zone and within-elevation transfer-risk in his newly formed zone 7. Based on our response surface (*Figure 55, p. 200*), families in this area (zone 751) clearly had a lower growth potential than expected based on prevailing temperatures. Our results thus support SORENSEN's interpretation that possibly some Washoe pine (*Pinus washoensis* Mason and Stockwell) families were mistakenly sampled instead of Ponderosa pine because Washoe pine has a lower growth potential than Ponderosa pine (WELLS, 1964a).

While SORENSEN used multiple regression procedures in order to relate trait combinations to location variables, our response surface and contour map directly illustrate the most important pattern of adaptive variance *i.e.* the temperature-related variation in growth vigor and growth timing. With multiple regression techniques, the effects of location variables are difficult to separate due to intercorrelations of the independent variables. Moreover, the illustration of adaptive patterns is less straightforward since associations have to be mapped separately for each dependent variable. In contrast, response surface analysis seems to provide an excellent statistical technique which may be utilized to directly and simply asses the major patterns of adaptive variance and to easily illustrate these patterns in geographic space.

Similar to Southwest Oregon, a weak, partially moisture-related, adaptive response was also apparent in Central Oregon. Compared to Southwest Oregon where 25% of the total trace was primarily related to moisture conditions of the site (*Table 44, p. 189*), the association was much weaker in Central Oregon, where 7.4 % of the total trace, represented in the second variate,

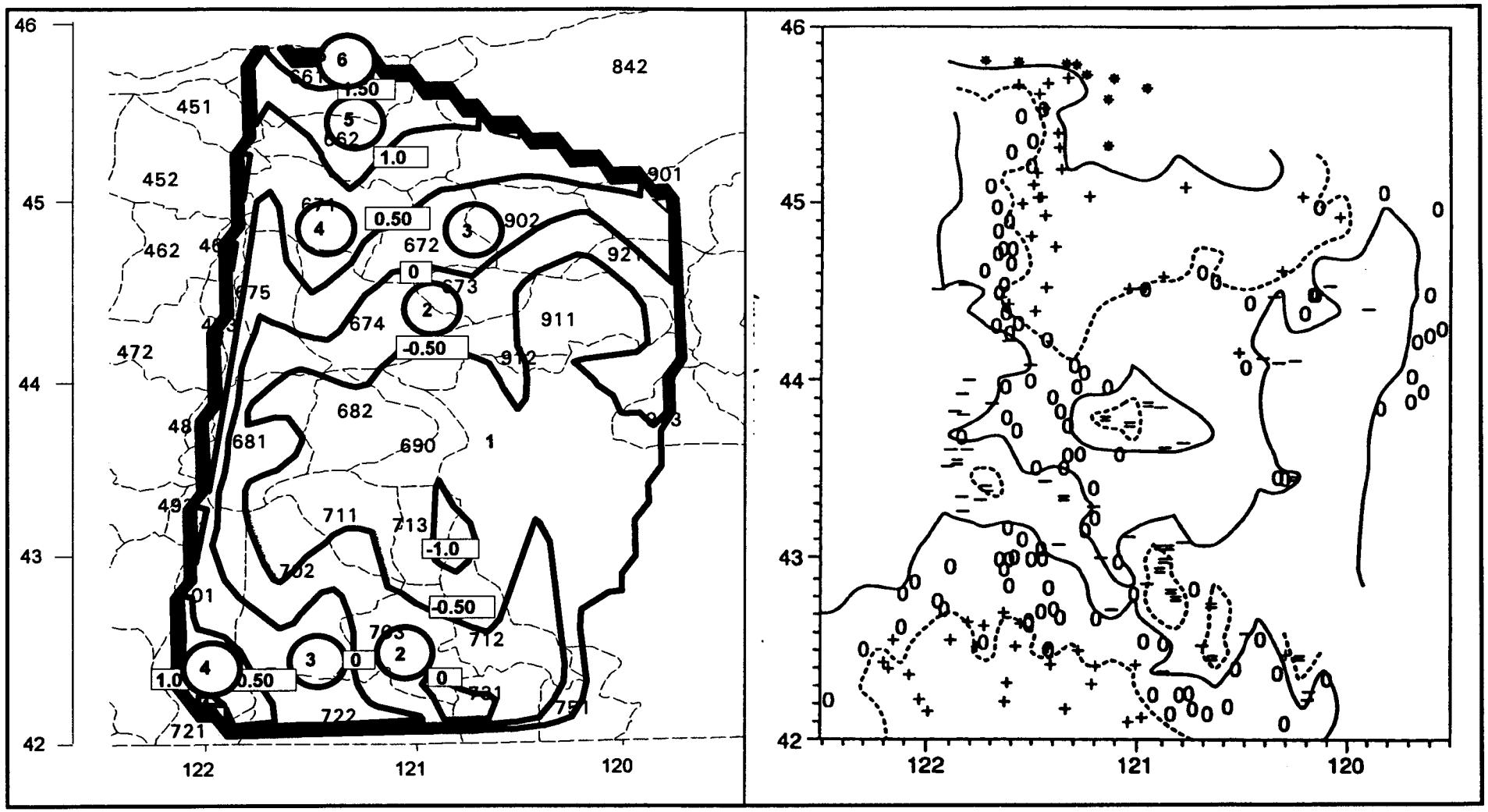


***Figure 58**: Comparison of adaptive patterns of variance in growth vigor for Central Oregon. Left: Contour plot of adaptive trend surface of canonical scores, scaled to represent a relative transfer risk of 20% among isolines. Right: Factor scores of first principal component, representing growth vigor, plotted by latitude and longitude as given by SORENSEN (1994)*

was related to moisture conditions (*Table 47, p. 197*). Although the second pattern was dominated by the water balance of the site, the revealed relationships were rather complex, however. As in Southwest Oregon, sources from dry habitats tended to emerge earlier than sources from moist sites. The other relationships portrayed in the second response variate are not easily explained; most likely some of the effects are temperature-related, others are moisture-related as the redundancies imply. The later budset of sources from dry sites reflected in the second variate is most likely the result of an indirect effect through temperature since the drier sites are also warmer in temperature. In contrast, the more plastic reaction of these sources is most likely rather an effect of moisture conditions.

Seed transfers over relatively limited distances within Southwest Oregon bear higher risks than transfers over the same distance within Central Oregon. Due to a rather ecotypic and rather complex pattern of adaptive variance, short distance transfers within Southwest Oregon may result in considerable mismatch between source locations and planting sites. In general, moving seed within Soutwest Oregon in a north-south direction is expected to cause less mismatch than moving it from west to east or vice versa. However, adaptive patterns were rather weak, indicating factors other than temperature and precipitation are involved in shaping the pattern. Possibly, geologic substrate and soil properties may play an additional role. Since we have no means of deciding on the role of other important factors which might explain the remaining proportions of adaptive variance, our contour maps must be regarded as preliminary. Our results suggest, however, that the current seed zones in this area most likely do not adequately portray the major patterns of adaptive variance.

The Columbia River basin seems to possess highly distinct sources of ponderosa pine. Seed from this area should be used only locally and be moved neither to the west- nor to the east. We agree with SORENSEN (1994) that seed zones or breeding blocks of ponderosa pine in Central Oregon could be considerably larger in longitude and latitude than the current zones. Based on our results, we would however propose a zonation different from the one published by SORENSEN. We are aware of the fact that our pattern of adaptive variation is only an approximation because climate conditions were estimated based on longitude, latitude and elevation only, disregarding local modifications caused by aspect and slope. On the other hand, our estimated transfer-risks are based on a combined adaptive response, combining growth vigor as well as growth timing in one major pattern. It may be seen from our contour plot that 9 zones are maximally needed to guarantee a risk smaller than 30%. For practical reasons, the 6 major zones indicated in *Figure 58, p. 208, left hand side* would be sufficient. Although the existing zones seem to be too conservative in number and size, they in general reflect the major environmental gradients which are of adaptive importance rather well. Several existing seed zone boundaries parallel the contours which delineate zones with a transfer risk below 30% (*Figure 58, p. 208*). The general orientation of the seed zone boundaries in a northeast-southwest direction in the northern part or in a northwestern-southeastern direction in the southern part of the sampling area clearly match the general geographic pattern of the major adaptive variance.

SORENSEN has proposed 9 zones based on estimates of cumulative transfer risk associated with the first 3 principal components (by summing the risks for each PCA) and on convenience for subdividing the area. The discrepancy between SORENSEN' s proposed breeding blocks and our contour map is difficult to interpret since his delineation procedure is not described in detail in the publication.

7.3 Summary

Much of the collective variation in seedling traits could be associated with climate variables of source locations. As much as 50 % of the patterned variation in growth and phenology traits could be explained. Moreover, patterns were reasonable, given the assumption that they reflect adaptation to the environment. These two observation are strong evidence that much of the variation in seedling traits reflect adaptation to local climate. Adaptive variance of single traits tended to be intercorrelated, forming co-adapted trait combinations. Temperature at source locations was the most important factor shaping the major patterns of adaptive variance. The moisture characteristics of the site, on the other hand, had only a minor influence.

The major pattern of adaptive variance reflected primarily relationships of growth vigor with

thermal conditions at source locations. About 24 % of variance in growth vigor of southwestern Oregon sources were redundant with variation in temperature. In Central Oregon, growth potential and temperature shared about 40 % of their variance.

Growth vigor appears to be the result of adaptation to temperature constraints at source location. Apparently, growth potential is strongly selected by temperature. Since elongation potential seems to be genetically "fixed" if temperature is the selective agent, selection most likely acts on phenology traits, adapting the sources to the length of the growing season at seed source location. Families from colder habitats have a genetically fixed vegetation cycle; they set their buds early regardless of the temperature conditions and thus are unable to respond to a longer growing season on more favorable lowland sites. In contrast, families from warmer habitats are adapted to a long growing season, grow until late in the season and thus produce larger seedlings. Adaptive variance in our sampling area was the same order of magnitude as that which has been reported for many other forest tree species. Likewise, our results are highly concordant with the results reported for Ponderosa pine in the adjoining areas in Washington, Idaho and Montana.

Patterns of adaptive variation in Southwest Oregon were rather complex. Growth potential and growth timing were clearly but comparatively weakly associated with temperature at source location. A rather small-scale, ecotypically organized pattern was evident. This ecotypic pattern of adaptive variance seems to be a result of rather complex interactions of both temperature and precipitation. Steep moisture gradients over short distances seem to be partly reflected in adaptive response of the sources. Even if temperature is clearly the most important selective agent, about 25% of the total trace appeared to be related primarily to moisture conditions of the site.

Growth vigor and growth timing of the sources in Central Oregon were strongly related to temperature. 40 % of the variation in growth potential were associated with temperature at source location. Growth potential was genetically fixed. Our results clearly support the hypothesis that elongation potential is strongly selected by temperature as has been hypothesized by SORENSEN (1994) who analyzed the same data using a different methodology. Moreover, the major pattern of adaptive variance, which reflected growth vigor and growth timing strongly resembled the pattern of variation in growth vigor published by SORENSEN (1994). Our result demonstrate that response surface analysis, using canonical correlation procedures, provides an excellent statistical technique which may be utilized to simply asses the major patterns of adaptive variance and to easily illustrate these patterns in geographic space.

Seed transfers over relatively limited distances within Southwest Oregon bear higher risks than transfers over the same distance within Central Oregon. Due to a rather ecotypic and rather complex pattern of adaptive variance, short distance transfers within Southwest Oregon may result in considerable mismatch between source locations and planting sites. In general, moving seed within Soutwest Oregon in a north-south direction is expected to cause less mismatch than moving it from west to east or vice versa. Our results suggest that the current seed zones in this area most likely do not adequately portray the major patterns of adaptive variance.

The Columbia River basin seems to possess highly distinct sources of ponderosa pine. Seed from this area should be used only locally and be moved neither to the west- nor to the east. Seed zones for Ponderosa pine in Central Oregon could be considerably larger in longitude and latitude than the current zones. Based on our map of relative seed transfer risk, 9 zones are maximally needed to guarantee a risk smaller than 30%. For practical reasons, 6 major zones appear to be sufficient. Although the existing zones seem to be too conservative in number and size, they in general seem to reflect the major environmental gradients which are of adaptive importance rather well.

8. Utility of allozymes versus metric traits for describing adaptive patterns of genetic variation

Zone designation for breeding programs or seed transfers are made on the basis of either geoclimatic and ecological data or on the evaluation of metric traits from short or long term genetic tests (CAMPBELL, 1984; REHFELDT, 1986a). The utilization of geoclimatic and ecological data assumes that local populations are optimally adapted and perform best at their original sites. Only under this assumption, do source populations, which define the seed zone, and the planting zone, defined by the planting environments, occupy the same geographic location. When non-local sources are optimal (NAMKOONG, 1969; NAMKOONG et al., 1988), the planting locations and the important traits determine the appropriate zones in the source populations. Finding appropriate sources requires testing of different sources over planting sites of interest. Moreover, the optimum source may depend on the trait or traits of interest. Difficulties arise from low inheritance of traits which reduce the precision in estimates of geographic patterns and from genotype-environment interactions that complicate the establishment of seed-transfer guidelines. Alternatively, adaptive variance may be determined by means of seedling common garden tests. In this case, associations between certain traits of interest and environmental conditions at seed source locations are investigated and patterns of adaptive variance are mapped in terms of ecological conditons or geographical location. However, at least 3 years are needed to collect the necessary data (and perhaps many more if traits are not adequately expressed in juvenile material). In contrast, allozyme data can be assessed easily and within short time periods, multilocus genotypes can be determined directly and are not subject to genotype-environments interactions. Allozyme patterns would thus be very helpful in providing easily available data on natural population genetic structures. However, to be of practical use, *i.e.* to allow zone designation, allozyme patterns would have to contribute information similar to that of metric traits.

Both seedling common garden and allozyme studies have the same limitations, however. Although patterns of genetic variation on the landscape can be mapped, the meaning of this variation for long-term survival and productivity remains speculative. It can only be assumed that the greater the difference between the source and potential planting environment, the greater is the risk in seed transfer. Moreover, it has to be assumed that local populations are adapted and perform best which is not necessarily the case (NAMKOONG, 1969; NAMKOONG et al., 1988).

Associations between allozyme genotypes and ecological factors have not been consistently established for forest trees (HAMRICK and GODT, 1990; HAMRICK et al. 1992). Moreover, the precision of methods used for detecting geographic patterns of variation in single loci is low and patterns of allozyme variation have not been consistently correlated with those from morphological traits (HAMRICK and GODT, 1990). Besides methodological problems (*e.g*, low statistical power of methods used, no consistent data sets), patterns in some morphological and physiological traits are not necessarily correlated with each other (DICKINSON et al. 1988; REHFELDT, 1986a). Nevertheless, for defining breeding or seed zones, metric traits have been used in most cases since metrical traits have a more readily distinguishable adaptive role.

A comparison between patterns of variation from allozymes and metric traits may thus provide additional information about the utility of geographic patterns of allozymes for developing breeding zones or seed-transfer guidelines. Since the ability to detect patterns is fundamentally a matter of the analytical methods used, we focus our comparison on multivariate patterns only. Because the same parent trees in the allozyme analysis were assessed in the nursery tests, our data allow a matched comparison of both patterns of variation.

The trend-surface models described different proportions of genetic variance in the response variates of the metric traits and the allozyme data (*Table 50, p. 212*). As expected, metric traits showed a higher association with climate conditions than did the allozymes. As we have outlined in detail in *section 1.4*, allozymes and quantitative data have fundamentally different properties and associations uncovered by each data set in this way are expected to differ even when multivariate techniques are used. Metric traits are the expressed sums of many different genes interacting in a specific environment. Even if the geographic variation in a metric trait is strongly associated with source environment of parent trees, this pattern will be much more dif-

	ALLOZYME DATA	SEEDLING TRAITS
SOUTHWEST OREGON		
SIGNIFICANT VECTORS	1	3
R^2_1	0.24	0.33
R^2_{Model}	0.24	0.42
CENTRAL OREGON		
SIGNIFICANT VECTORS	1	2
R^2_1	0.23	0.47
R^2_{Model}	0.23	0.50

***Table 50**: Proportion of variation described by the trend surface models using allozymes or seedling traits as dependent data sets. R^2_1: Proportion of variation described by the trend surface model in the first canonical variate; R^2_{model}: Total proportion of variation significantly described by the model*

ficult to detect at the level of the individual genes influencing this trait. Since many loci may contribute to the expression of a metric trait, it would be very unlikely that all of them would be included in a random sample of individual genes. Allozymes, however, are gene markers. They generally contain an "average" sample of structural genes affecting many different characters. Consequently, allozymes are expected to exhibit weaker associations with habitat conditions than metric traits do, even when multivariate analysis is used. Likewise, due to these fundamental different properties of the two data sets (allozymes vs. metric traits), the detection of relationships between individual allozyme markers and metric traits is rather unlikely.

Viewed from this perspective, the associations between multilocus allozyme markers and climate revealed in this study are striking. Thus, in spite of the fundamental limitations of individual allozyme markers, multivariate analysis of allozymes at many loci may reveal adaptive patterns of genetic variation on the landscape.

The levels of association between multilocus allozyme variates and climate revealed in this study (23% and 24% in the two regions) are very similar to the levels reported for ponderosa pine in the Sierra Nevada Mountains of California (WESTFALL and CONKLE, 1992). Using a second order model with latitude, longitude and elevation, these authors could explain 25 % multilocus allozyme variance in the first vector and 40 % over the first four significant vectors combined. Our results contrast with the findings in the Sierra Nevada in two respects. Although we have analyzed about the same number of loci, only one vector significantly described adaptive variance in allozyme scores within each region. Interestingly, however, the same loci associated with geographic variables in the Sierra Nevada sample were associated with climate variables in our study. Secondly, while the proportion of family variation in seedling quantitative traits accounted for by geographic variables was similar in the WESTFALL and CONKLE (1992) study to the proportion of allozyme data accounted for (26 % in the first vector, 36 % in the first three vectors), our data clearly revealed much stronger relationships for the metric traits. We see three possible explanations for these contrasting results. The overall higher association which was reflected in four significant vectors in the Sierra Nevada sample may be due to a higher environmental heterogeneity caused by a larger sampling area extending primarily in the north-south direction. Secondly, compared to results in the literature, the proportion of genetic variance (36 %) in metric traits, explained in the Sierra Nevada study appears rather low. This relatively low association may be due to the inclusion of a high number of morphological traits. Our results, however, are based on growth and phenology traits which are commonly strongly associated with habitat conditions of the source location. As a third explanation, genetic variation in metric traits may be differently associated with climate than with geographic variables, since processes other than natural selection may be reflected in the latter.

Although quantitative traits appear to reflect much higher proportions of adaptive variance than allozymes, patterns of variation in multilocus allozyme traits clearly approximate the patterns for seedling traits. For ease of comparison, both patterns of variation are shown side by side for for Southwest Oregon in *Figure 59, p. 213* and for Central Oregon in *Figure 60, p. 215.*

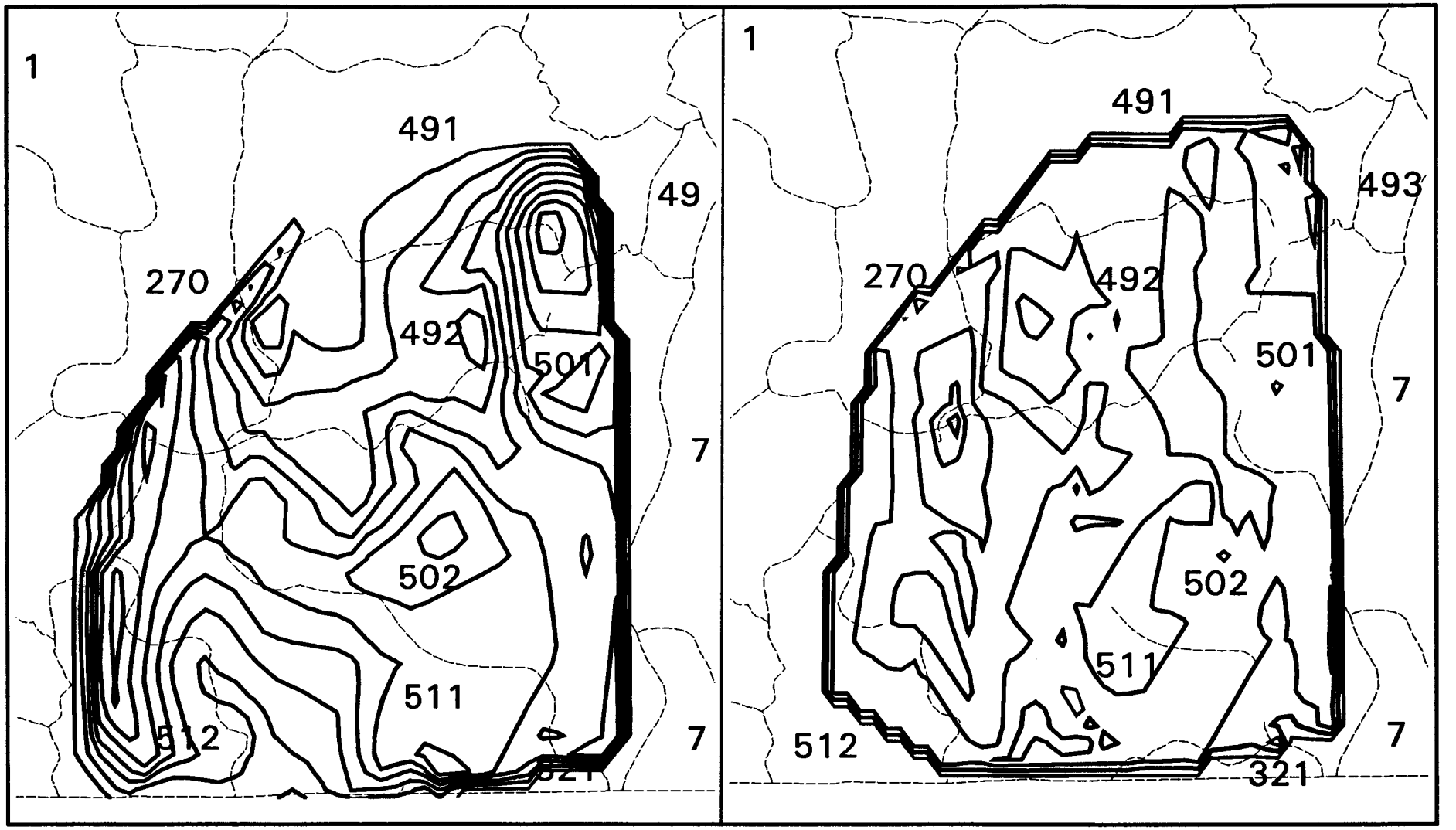


***Figure 59**: Comparison among patterns of adaptive variance based on allozyme data and seedling quantitative traits: Southwest Oregon. Contour plots of predicted scores on the first response variate for the allozyme vector (left) and the quantitative trait vector (right), scaled to represent a seed-transfer risk of 20% among two neighboring contours*

Inferred from the allozyme pattern, changes in multilocus frequencies in Southwest Oregon are very rapid over short geographic distances, especially in the western part and in the north-eastern corner of the sampling area (*Figure 59, p. 213*). A complex pattern of adaptive variance is also evident in the contour map derived from the seedling traits. Gradients, however, are clearly less pronounced than those in the allozyme pattern. Visually, the patterns do not seem to differ considerably. With a value of $R = 0.142$, the correlation among the predicted scores of the first allozyme vector and those of the first metric trait vector is rather low, however, indicating that the two trend surfaces differ considerably. This comparatively low congruence of the patterns in Southwest Oregon may be due to several reasons. First of all, while overall trends may be similar (as the contour maps suggest) for both traits, the two surfaces may still differ locally due to the complex and small scale pattern of variation observed in this area. Secondly, both sets of traits (allozymes and metric traits) may be responding to the same selection pressures, but error of estimation (for individual points on the surface) is rather large, leading to the low congruence of the two surfaces. Another possible explanation is that the two sets of traits may be influenced to different degrees by the effects of different evolutionary forces. Traits may vary in the degree of response to selection pressure, as well as to other factors. Finally, both sets of traits may reflect adaptation to local environment, but they are responding to different selection pressures (even different metric traits may individually show quite different patterns). The fact that both sets of traits are associated with climate variables suggest that both may be responding to selection pressure, but that they either vary in the degree of their response or that they are affected to different degrees by different forces.

As we have discussed in detail in *sections 6 and 7*, adaptive patterns in Southwest Oregon differed in several respects from Central Oregon and from the total sampling area. Patterns in Southwest Oregon were comparatively weak for both allozymes and metric traits. Moreover, both patterns were associated with temperature and moisture conditions. As a result of steep moisture gradients in combination with distinct soil properties, populations in Southwest Oregon appear to be partly adapted to the moisture conditions of the site. About half of the variation in allozyme scores and about 25 % of the variation in seedling traits were associated with moisture conditions. Considering these differences, several reasons for the low congruence of the two patterns are conceivable. The degree of congruence may be a function of the strength and the quality of the respective patterns. Adaptive variance of allozymes and quantitative traits may resemble each other only if patterns are strong enough or if they are dominated by the same environmental factor governing adaptation. Complex and interacting selective environments may result in differing patterns because allozymes and traits may be affected differently or to a different degree by various selective agents or their varying interactions. Since allozymes and traits have different properties and may reflect different parts of the genome, it is likely that an allozyme sample will contain a certain number of loci which are unrelated to the studied quantitative traits but which are associated with other characters that are affected by natural selection. Under such circumstances, patterns will differ because they provide different information. Hence, based on our results, a conceivable hypothesis would be that the sampled allozyme loci (especially those enzymes associated with glycolysis and the Krebs cycle such as IDH, PGM, ACO and MDH) reflect similar patterns of adaptive variance as growth and phenology traits when temperature is the major selective agent, whereas patterns are less congruent when other factors such as moisture or complex interactions of different factors are involved. Allozyme scores and growth traits appear to be associated with temperature to about the same degree whereas associations with moisture seem to differ. In Southwest Oregon, allozymes showed a higher association with moisture than did the metric traits. Allozymes thus appear to focus more on the moisture conditions of the site than do the studied traits. This may be a likely explanation for the low congruence of the two patterns in Southwest Oregon with respect to small scale, local variation. The general trends of the two surfaces, however, seem much more congruent. In fact, a visual inspection of the two contour maps reveals that both patterns portray the same major ecological gradients. Both vectors are mainly oriented in a west-east direction. Both patterns indicate that coastal populations differ from populations in the eastern part and that they are again different from populations in the center of the sampled area. Even if the contours differ in number, form and shape, the size, location and form of a distinct zone in the center appears very similar in both maps. This central zone coincides with dry habitat conditions. Both patterns thus clearly reflect a certain importance of moisture as a selective force. Moisture gradients, however, are represented in more detail in the adaptive response of the allozymes than in the response of the selected traits.

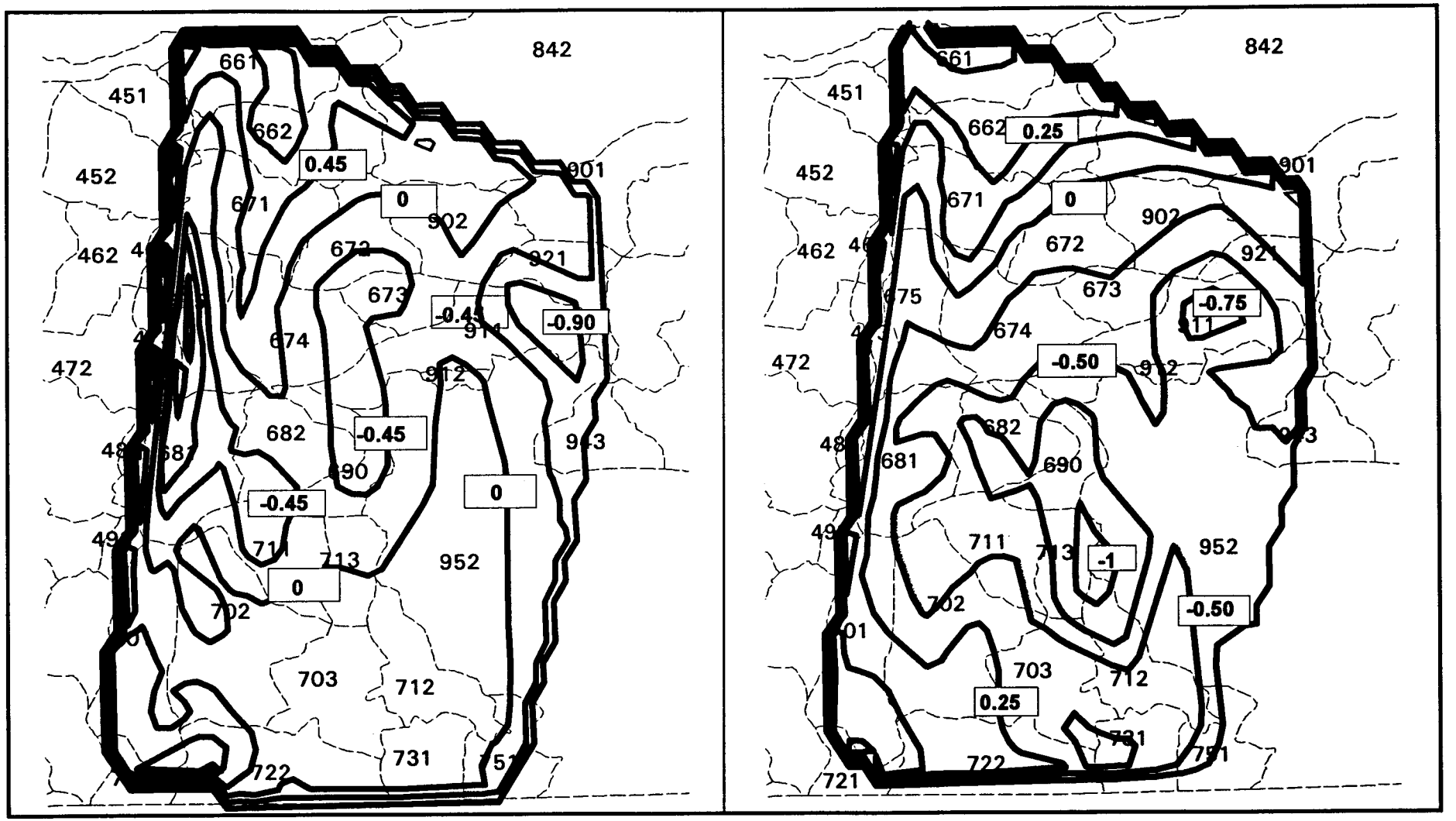


***Figure 60**: Comparison among patterns of adaptive variance based on allozyme data and seedling quantitative traits: Central Oregon. Contour plots of predicted scores on the first response variate for the allozyme vector (left) and the trait vector (right), scaled to represent a seed-transfer risk of 10% among two neighboring contours*

With a correlation coefficient of $R = 0.66$ between the predicted scores of the first allozyme variate and the first metric trait variate, the two patterns in Central Oregon resemble each other rather well. The contour lines (*Figure 60, p. 215)* are very similar except in the southeastern corner of the sampled area. Both patterns show the same northeast-southwest or southwest-northwest orientation of the contours. Both clearly delineate a central zone which appears to differ from the northern and the southern parts of the area. However, form and location of this central zone differ somewhat for the two patterns. While both show a clear minimum of scores in the area around seed zones 911 and 921, the second area with low scores extends slightly more to the north in the allozyme pattern (zones 672, 673, 911, 690) than in the trait pattern (zones 713, 690, 682). We do not have an explanation for this difference. It must be emphasized, however, that ponderosa pine occurs only in small and partly isolated stands in the eastern parts of the sampling area and sampling density was therefore rather small. For example, the large area of seed zone 952 is represented by 8 trees, and zones 712, 912 and 690 by 8, 4 and 9 individuals only. Consequently, random effects may be responsible for these minor differences. Nevertheless, allozymes and metric traits clearly portray the same major ecological gradients of adaptive importance.

In the studied case, patterns of multivariate allozyme variation on the landscape thus bear a resemblance to those of metric traits. Since patterns based on quantitative traits make sense from an adaptation point of view (which is commonly agreed on), allozyme patterns seem to reflect adaptation to a certain extent. This conclusion is in perfect agreement with conclusions already reached in the former sections.

For Southwest Oregon, both data sets clearly indicate that adaptive patterns are highly complex and rather ecotypically organized, that seed should not be transferred over large distances and that transfers in an east-west direction cause more mismatch than transfers in a north-south direction. Even if contour intervals differ for the two patterns, both contour maps demonstrate that the current seed zones do not seem to sufficiently reflect the adaptive genetic variance of ponderosa pine in this area and that zonation should be more conservative. However, both patterns also indicate that seed transfer should rather be based on models which describe genetic variation in relation to environmental variation than on a system of discrete seed zones. Moreover, both data sets clearly reveal an important influence of moisture conditions that is reflected in the genetic structure of ponderosa pine in this area.

Nevertheless, the two trend surface models which may be used to predict the risks of seed transfers in Southwest Oregon for any given source location and planting site, will furnish partly divergent results because allozymes seem to focus more on moisture conditions than do the metric traits. Some of the allozymes are most likely associated with characters other than growth and phenology. These characters (such as water use efficiency, stomata regulation, stress avoidance, mineral uptake or others) may be important for survival and fitness under the geologicaly distinct, dry habitat conditions of this area. Allozymes could thus contribute important additional information which is not reflected in the measured traits. Our results, however, do not allow any conclusions regarding the potential utility nor the ecological or practical significance of such additional information. Only long term field tests under varying habitat conditions could provide an answer to this question.

In Central Oregon, both patterns clearly indicate that seed may be transferred from northern to southern planting sites and vice versa with an acceptable risk but should not be used on sites in the high desert country of central Oregon. Both patterns show that the existing seed zones are very conservative for ponderosa pine and that they could be consolidated into only 5 to 8 zones. The high congruence of both patterns of adaptive variance would permit the basing of the estimates of seed-transfer risk for any given source location and planting site on allozyme data instead of on metric traits. Even if estimates of transfer risk based on allozyme markers would differ in some parts of Central Oregon, the differences would be too small to be of practical importance. If we adopt an acceptable risk of 30%, decisions would be practically identical.

In summary, in the presented case, allozyme markers were thus able to contribute useful information which was highly compatible with that of seedling quantitative traits such as growth characteristics and phenology traits. Consequently, multilocus allozyme patterns could be directly used for practical purposes such as development of seed transfer guidelines or seed-zone designation.

This conclusion can, however, not be generalized. The question of whether multivariate patterns of allozyme variation are useful for describing adaptive variation in other species and in different environmental conditions cannot be answered. This is especially so because only a few studies have applied multivariate statisical techniques to allozymes. With the exception of results reported by WESTFALL and CONKLE (1992) and MILLAR and WESTFALL (1992), whose results are perfectly supported by our findings, little is known about the potential utility of multivariate allozyme patterns for practical purposes. Moreover, our as well as the cited studies investigated allozyme variation in regions with rather pronounced environmental gradients. As suggested by our results from Southwest Oregon, the adaptive significance of allozymes may, however, vary as a function of the environmental variation found in the studied area and the complexity and interaction of the selective agents. Any generalization about the potential and utility of multilocus allozyme variation would thus be merely speculative. However, based on our results, we perfectly agree with the view of MILLAR and WESTFALL (1992), that certain statistical methods may allow allozyme diversity to be partitioned such that the portion of allozyme diversity that correlates with adaptive variation is revealed. We also agree with WESTFALL and CONKLE (1992) that significant, regular geographic patterns exist in allozyme loci in forest trees and that information on allozyme patterns can augment or even replace that of metric traits in certain situations.

Allozymes may, however, not only be useful traits regarding their possible adaptive portion. For conservation purposes, for example, all genetic information should be used together in order to assess patterns of variation important for conservation (MILLAR and WESTFALL, 1992). Allozyme variation thus may provide useful and important additional information. Since we do not know what genes are or will be adaptive in future environments, all traits, including allozymes, should be weighted equally, independent of their role for adaptation. Multilocus allozyme variation and other traits may, however, give discordant patterns since the two traits may be reflecting different evolutionary forces. In this case, the species in question may be divided into the smallest reasonable units suggested by cummulative analysis of individual patterns as has been suggested and demonstrated for bishop pine (*Pinus muricata*) by MILLAR and WESTFALL (1992).

9. Conclusions

Variation of allele frequencies in the studied area is clearly non-random for many alleles at several loci. Multilocus analyses reveal statistically significant geographic patterns of allozyme variation which tend to follow major climatic patterns, especially temperature gradients. Multilocus frequency distributions in the study area seem to reflect important adaptive processes. The observed differentiation in allele frequencies between the coastal and interior regions does not appear to be a result of historical events such as immigration from two different progenitor lines. It seems rather to be the result of natural selection in differing environments.

Allozyme variation seems to have a certain adaptive significance. However, the adaptive significance varies among the different alleles and gene loci. Many alleles at several loci are unrelated to habitat conditions. Multilocus analyses may furnish a clearer picture of adaptive patterns of variation and may help to understand important factors or processes which are involved in shaping the observed patterns. Multilocus trend surfaces seem especially useful in describing and illustrating such adaptive patterns of variation in geographic space.

Patterns of allozyme variation clearly parallel patterns in metric traits. The congruence of the two patterns provides further evidence that multilocus allozyme patterns may reflect adaptation to a certain extent.Thus, we fully agree with WESTFALL and CONKLE (1992) that allozyme data can contribute useful information to the understanding of adaptive patterns of genetic variation and its practical consequences for breeding zone formation, seed-transfer risk estimates or gene conservation measures.

10. Summary

The study describes and interprets patterns of genetic variation in ponderosa pine (*Pinus ponderosa Dougl. Ex. Laws.*) from two contrasting regions of the state of Oregon, USA, namely the east slope of the Cascade Range with parts of the high desert (Warner and Ochoco mountains) and the Klamath and Siskiou mountains in Southwest Oregon. Isozyme data as well as seedling quantitative traits of the same families were analyzed primarily by multivariate statistical techniques. Trend surface analysis was used to describe relationships of allozymes and metric traits with climate conditions at source locations. The resulting multilocus allozyme surfaces were compared with those of metric traits and the utility of allozymes for describing patterns of adaptive variance was investigated. Resulting patterns of adaptive variance were also utilized to estimate relative seed-transfer risks and to map zones of certain transfer risks in geographic space. Resulting contour maps of transfer risks were used to assess the adequacy of the existing seed zones to guide seed transfer for ponderosa pine.

The two regions were described and compared with respect to their single-locus and multilocus genotypic structures. Measures of diversity and differentiation were analyzed with the traditional univariate procedures. Single alleles were also tested for spatial structures by spatial autocorrelation analysis, and associations with climate variables were analyzed with multinomial response models and Mantel tests of matrix associations. In addition, allozyme variation was examined in multilocus sets, applying canonical discriminant and correlation analyses to genotypic scores. Results of single-locus analyses were compared to results of multivariate analyses in order to evaluate the utility of multivariate techniques applied to isozyme data.

Seed of one or two individuals was sampled on a grid at each of 217 locations in Southwest Oregon and 227 locations in Central Oregon. Families were tested in two separate common garden tests. Growth and phenology traits were recorded over a period of 3 years. Leftover seed of the same families was used for electrophoresis. Megagametophytes from six seeds sampled from each of 488 parent trees were analyzed electrophoretically for 31 allozyme loci. Environmental data were estimated for each tree location by the use of local regression models. Long term climate data of 195 weather stations in Oregon served as input for model-building. Climate variables for tree locations were predicted from these models based on their latitude, longitude and elevation.

Single-locus analyses revealed that ponderosa pine in Oregon maintains a high genetic diversity which is comparable to values reported for geographically adjoining parts of the range. Diversity estimates were in the upper range of values found for conifer species. Moreover, genetic variation seems well distributed over the area; much of total diversity is found within local areas (or seed zones) as indicated by low F_{ST} - values and by estimates of average diversity within seed zones. Minor and major polymorphisms were found in about equal proportions. Rare alleles with frequencies smaller than 5 % were, however, found at most loci (24). The high genetic diversity may be perceived as a high potential of adaptability to the extreme and variable habitat conditions which prevail in the area. Although Southwest Oregon had a slightly higher mean number of alleles per locus, a higher proportion of polymorphic loci, and showed 5 rare alleles not found in Central Oregon, differences in genetic diversity between the two regions were small and non-significant. Moreover, observed differences between the two regions most likely reflect the unequal sample sizes (leading to large sampling error) than real differences.

Allelic frequencies were significantly heterogeneous between the two regions when all 31 loci are considered. The two regions differed on average by a proportion of 6.1% unshared alleles. With an average of 0.007 for NEI's unbiased genetic distance, differentiation between the regions, however, was rather small, being in the range of distances commonly found between populations. Differentiation was no higher among the two regions than among populations within regions (within-region estimates are however upwardly biased due to small sample sizes). The small average differentiation between the two regions argues against the existence of different races (Pacific race, North Plateau race) in the areas east and west of the Cascade Range which has been postulated by different authors based on morphological differences and monoterpene composition. Methodological limitations of composite measures of differentiation,

as well as fundamental difficulties involved in decisions on the taxonomic status of taxa, do not allow a final conclusion, however.

In contrast, allele frequency differences at several loci which exist between the two regions would argue in favor of possible racial differentiation. Allele frequency differences may, however, not only be the result of a separate evolution in the past; they may also be a result of natural selection in different environments or a consequence of random genetic processes.

Patterns of differentiation, emerging from cluster analysis on genetic distance by NEI or GREGORIUS, suggest that genetic and environmental variation parallel each other. Such associations between genetic and environmental variation are expected if natural selection plays a major role in shaping genetic variation. Results from univariate genetic analyses are non conclusive, however. Based on single-locus analyses, natural selection is a plausible but not the only possible explanation for the observed patterns of differentiation.

Several alleles at many loci showed spatial variation patterns which clearly deviate from random spatial arrangement. Significant heterogeneities of allele frequencies, significant spatial patterns and observed associations between genotypic frequencies and climate for the same alleles suggest that adaptation to the environment is the most likely cause for the observed variation patterns. Although other causes can not be ruled out, patterns are most likely the result of differential selection for most of the alleles which exhibit a spatial pattern. Associations of allele frequencies with temperature could be confirmed by results from the various analyses using different analytical tools. Temperature seems to be an important selective agent responsible for adaptation. Genetic structure, on the other hand, seems unrelated to water balance of the site. Lack of adaptation of ponderosa pine to hydrological conditions or the inability of the marker genes to portray adaptation to water regime may be plausible explanations for these findings. Based on spatial structures and associations with climate, no final conclusion about the adaptivness of alleles can be drawn, however, since climate effects cannot be separated from other effects. Nested multinomial response models suggest that some of the associations among genotypic frequencies and temperature are most likely caused by indirect effects. Such indirect effects may be due to a different evolutionary past of the populations in the two regions. Since climate differs between the two regions, other region related differences may be mimic an association with climate conditions. According to the results from nested response models, *Mnr-2*, *Pep-3*, *Mpi-1*, *Idh-1*, *Pgm-1*, *Aco-1*, *Skd-2*, *Ugp-1*, *Ugp-2* and *Mdh-3* are most likely associated with temperature at source location. Associations of *Mnr-1*, *Acp-1*, *Adh-2*, *Lap-2* and *Fdp-2* with climate conditions, however, are most likely caused by other than climate effects. Such effects are indirectly related to climate conditions and thus mimic associations with temperature conditions, although temperature and genotypic frequencies are in fact not associated.

Multilocus genotypic frequencies were moderately associated with habitat conditions. Based on different multivariate analyses, climatic conditions at source location significantly explained about 33% of total variance of 41 allozyme variables. Within the regions, the proportion of explained variance was about 23%. Overall, 13 out of 31 loci showed an association with climatic conditions. These results clearly argue for the adaptivness of certain alleles or enzyme systems. In accordance with the former results, especially enzyme systems involved in important physiological pathways such as the loci *Mnr-1*, *Mnr-2*, *Lap-2*, *Pep-3*, *G6p-2*, *Acp-1*, *Gdh-1*, *Idh-1*, *Pgm-1*, *Skd-2* and *Mdh-3* seem to be adaptive. Not all alleles contribute to the adaptive pattern of variance. Of the markers, 42% were not related with the pattern. Amounts of associations differ among the alleles. Of the alleles, 16% showed a correlation of more than 10%, 31% of more than 20% and 11% of more than 30% with the adaptive response surface. Single alleles thus behave much in the way one would expect of quantitative trait loci; each gene contributing only small amounts to the adaptive multilocus pattern of variation. While these small individual contributions may be difficult to detect with single locus analyses, the aggregate patterns resulting from multivariate analyses can be strong and meaningful.

Close associations of patterns of multilocus allozyme frequencies with environmental variation within each of the two regions are potentially consistent with migration from different refugia populations and with adaptation. Complex patterns of allozyme variation associated with complex environmental variation, however, are unlikely for migration as a major evolutionary force.

Since patterns are repeated in different places and since they closely resemble patterns seen in quantitative traits published for Central Oregon by SORENSEN (1994), adaptation to current environments caused by natural selection seems the most likely cause for the observed patterns of multilocus allozyme frequencies.

Multilocus genotypic frequencies were primarily related to temperature. All results clearly lead to the conclusion that temperature at source locations is the most important environmental factor responsible for the observed adaptive pattern of multilocus allozyme frequency distribution. A striking similarity between the adaptive response surface and variation patterns of temperature conditions was observed. Moisture characteristics of the site, on the other hand, had only a very minor overall effect on genotypic variation, except in Southwest Oregon.

The adaptive patterns differed between the two regions. In Southwest Oregon, patterning was weak, with temperature and moisture conditions of the site having about equal influence on multilocus frequencies. Strong moisture gradients combined with relatively mild climate may lead to a more important influence of moisture characteristics in this area. Distinct geologic substrate and soil conditions may also play an important role. Patterns in Southwest Oregon seem to be highly complex. They seem to reflect the topographic and ecological complexity of this area.

Adaptive patterns of multilocus allozyme frequencies in Central Oregon were primarily related to temperature. However, moisture conditions and interactions contributed small but significant proportions to the explained variance. The adaptive pattern of multilocus frequencies was nearly identical with patterns based on seedling quantitative traits published by SORENSEN (1994). This high congruence of the two patterns provides strong evidence that allozyme markers can be very useful in describing ecological patterns of adaptation.

Based on multivariate analyses, differentiation between the two regions appears to be the result of natural selection in two contrasting environments rather than the consequence of a different evolutionary past. The existence of two races (Pacific race, North Plateau race) as a consequence of a different evolutionary past, as suggested by several authors, is highly unlikely. Based on results of canonical and partial canonical analyses, only 2.5% of the significantly explained variance in the adaptive response surface is due to pure spatial effects. Pure spatial effects are reflected in patterns of variation which are independent of environmental variation. Such spatial effects should predominate if historical events were responsible for the present patterns of variation. Since 97.5% of variation in multilocus frequency distribution are caused by environmental effects and only 2.5% by pure spatial effects, historical events such as a long and separate evolution of different base populations or the immigration from two different refugia are highly unlikely. Results from multivariate analyses clearly demonstrate that differentiation between the two areas is primarily and nearly exclusively due to adaptation to different habitat conditions which exist in the two areas.

Contour maps of relative seed transfer-risk, calculated from the adaptive response surfaces, can be used to delineate areas on the adaptive surface with a transfer-risk smaller than a given value. Derived contour maps indicate that seed should not be transferred across the Cascade Range. Within Southwest Oregon, seed generally should not be moved in the east-west direction. Steep gradients in transfer-risk are observable in this area. The current seed zone boundaries seem to reflect the important ecological and genetic gradients rather well. However, zones seem to be too large to guarantee an acceptable transfer-risk. Seed transfer guidelines should be based on models of transfer risk and not on seed zones, since the formation of such zones seems highly complex.

Current seed zones in Central Oregon do not seem to reflect the important adaptive patterns. In contrast to Southwest Oregon, zonation seems to be rather conservative. Zones could be considerably larger, especially in the southern and the central part of this region. To guarantee a maximum relative transfer-risk of 30%, only about 7 zones are required.

Much of the collective variation in seedling traits appears to be associated with climate variables of source locations. As much as 50 % of the patterned variation in growth and phenology traits could be explained. Moreover, patterns were reasonable, given the assumption that they

reflect adaptation to the environment. These two observation are strong evidence that much of the variation in seedling traits reflect adaptation to local climate. Adaptive variance of single traits tended to be intercorrelated, forming co-adapted trait combinations. Temperature at source locations was the most important factor shaping the major patterns of adaptive variance. The moisture characteristics of the site, on the other hand, had only a minor influence.

The major patterns of adaptive variance reflected primarily relationships of growth vigor with thermal conditions at source locations. About 24 % of variance in growth vigor of southwestern Oregon sources were redundant with variation in temperature. In Central Oregon, growth potential and temperature shared about 40 % of their variance.

Growth vigor appears to be the result of adaptation to temperature constraints at source location. Apparently, growth potential is strongly selected by temperature. Since elongation potential seems to be inherently "fixed" if temperature is the selective agent, selection most likely acts on phenology traits, adapting the sources to the length of the growing season at seed source location. Families from colder habitats have a genetically fixed vegetation cycle; they set their buds early regardless of test environments and thus are unable to respond to a longer growing season on more favorable lowland sites. In contrast, families from warmer habitats are adapted to a long growing season, grow until late in the season and thus produce larger seedlings. Adaptive variance in our sampling area was the same order of magnitude as that which has been reported for many other forest tree species. Likewise, our results are highly concordant with the results reported for Ponderosa pine in the adjoining areas in Washington, Idaho and Montana.

Patterns of adaptive variation of seedling traits in Southwest Oregon were rather complex. Growth potential and growth timing were clearly but comparatively weakly associated with temperature at source location. A rather small-scale, ecotypically organized pattern was evident. This ecotypic pattern of adaptive variance seems to be a result of rather complex interactions of both temperature and precipitation. Steep moisture gradients over short distances seem to be partly reflected in adaptive response of the sources. Even if temperature is clearly the most important selective agent, about 25% of the total trace appeared to be related primarily to moisture conditions of the site.

Growth vigor and growth timing of the sources in Central Oregon were strongly related to temperature. Forty percent of the variation in growth potential were associated with temperature at source location. Growth potential was genetically fixed. Our results clearly support the hypothesis that elongation potential is strongly selected by temperature as has been hypothesized by SORENSEN (1994) who analyzed the same data using a different methodology. Moreover, the major pattern of adaptive variance, which reflected growth vigor and growth timing, strongly resembled the pattern of variation in growth vigor published by SORENSEN (1994). Our result demonstrate that response surface analysis, using canonical correlation procedures, provides an excellent statistical technique which may be utilized to simply asses the major patterns of adaptive variance and to easily illustrate these patterns in geographic space.

Based on metric traits, seed transfers over relatively limited distances within Southwest Oregon bear higher risks than transfers over the same distance within Central Oregon. Due to a rather ecotypic and rather complex pattern of adaptive variance, short distance transfers within Southwest Oregon may result in considerable mismatch between source locations and planting sites. In general, moving seed within Soutwest Oregon in a north-south direction is expected to cause less mismatch than moving it from west to east or vice versa. Our results suggest that the current seed zones in this area most likely do not adequately portray the major patterns of adaptive variance.

The Columbia River basin seems to possess highly distinct sources of ponderosa pine. Seed from this area should be used only locally and be moved neither to the west- nor to the east. Seed zones for Ponderosa pine in Central Oregon could be considerably larger in longitude and latitude than the current zones. Based on our map of relative seed transfer risk, 9 zones are maximally needed to guarantee a risk smaller than 30%. For practical reasons, 6 major zones appear to be sufficient. Although the existing zones seem to be too conservative in number and

size, they in general seem to reflect the major environmental gradients which are of adaptive importance rather well.

The trend-surface models described different proportions of additive genetic variance in the response variates of the metric traits and the allozyme data. As expected, metric traits showed a higher association with climate conditions than did the allozymes. However, in spite of the fundamental limitations of allozyme markers, multivariate analysis of allozymes at many loci may reveal adaptive patterns of genetic variation on the landscape. Although quantitative traits reflected much higher proportions of adaptive variance than allozymes, the adaptive patterns in allozymes clearly approximated the patterns for seedling traits, at least in Central Oregon.

Inferred from the allozyme pattern, changes in multilocus frequencies in Southwest Oregon were very rapid over short geographic distances, especially in the western part and in the northeastern corner of the sampling area. A complex pattern of adaptive variance was also evident in the contour map derived from the seedling traits. Gradients, however, were clearly less pronounced than those in the allozyme pattern. Visually, the major patterns did not seem to differ considerably. With a value of $R = 0.142$, the correlation among the predicted scores of the first allozyme vector and those of the first metric trait vector was rather low, however, indicating that the two trend surfaces differ in fact considerably. This comparatively low congruence of the patterns in Southwest Oregon may be due to several reasons. First of all, while overall trends may be similar (as the contour maps suggest) for both traits, the two surfaces may still differ locally due to the complex and small scale pattern of variaiton observed in this area. Secondly, both sets of traits (allozymes and metric traits) may be responding to the same selection pressures, but error of estimation (for individual points on the surface) is rather large, leading to the low congruence of the two surfaces. Another possible explanation is that the two sets of traits may be influenced to different degrees by the effects of different evolutionary forces. Traits may vary in the degree of response to selection pressure, as well as to other factors. Finally, both sets of traits may reflect adaptation to local environment, but they are responding to different selection pressures (even different metric traits may individually show quite different patterns). The fact that both sets of traits are associated with climate variables suggest that both may be responding to selection pressure, but that they either vary in the degree of their response or that they are affected to different degrees by different forces.

In Southwest Oregon, allozymes showed a higher association with moisture than did the metric traits. Allozymes thus appear to focus more on the moisture conditions of the site than do the studied traits. This may be a likely explanation for the low congruence of the two patterns in Southwest Oregon with respect to small scale, local variation. The general trends of the two surfaces, however, seem much more congruent. A visual inspection of the two contour maps reveales that both patterns portray the same major ecological gradients. Both vectors are mainly oriented in a west-east direction. Both patterns indicate that coastal populations differ from populations in the eastern part and that they are again different from populations in the center of the sampled area. Even if the contours differ in number, form and shape, the size, location and form of a distinct zone in the center appears very similar in both maps. This central zone coincides with dry habitat conditions. Both patterns thus clearly reflect a certain importance of moisture as a selective force. Moisture gradients, however, are represented in more detail in the adaptive response of the allozymes than in the response of the selected traits.

With a correlation coefficient of $R = 0.66$ between the predicted scores of the first allozyme variate and the first metric trait variate, the two patterns in Central Oregon resembled each other rather well. The contours were very similar except in the southeastern corner of the sampled area. Both patterns show the same northeast-southwest or southwest-northeast orientation of the contours. Both clearly delineate a central zone which appears to differ from the northern and the southern parts of the area. However, form and location of this central zone differ somewhat for the two patterns. While both show a clear minimum of scores in the area around seed zones 911 and 921, the second area with low scores extends slightly more to the north in the allozyme pattern (zones 672, 673, 911, 690) than in the trait pattern (zones 713, 690, 682). We do not have an explanation for this difference. It must be emphasized, however, that ponderosa pine occurs only in small and partly isolated stands in the eastern parts of the sampling area and sampling density was therefore rather small. For example, the large area of seed zone 952 is represented by 8 trees, and zones 712, 912 and 690 by 8, 4 and 9 individuals

only. Consequently, random effects may be responsible for these minor differences. Nevertheless, allozymes and metric traits clearly portray the same major ecological gradients of adaptive importance.

In the studied case, patterns of multivariate allozyme variation on the landscape thus bear a resemblance to those of metric traits. Since patterns based on quantitative traits make sense from an adaptation point of view (which is commonly agreed on), allozyme patterns seem to reflect adaptation to a certain extent. This conclusion is in perfect agreement with conclusions already reached from results of the other analyses. In the presented case, allozyme markers were thus able to contribute useful information which was highly compatible with that of seedling quantitative traits such as growth characteristics and phenology traits. Consequently, multilocus allozyme patterns could be directly used for practical purposes such as development of seed transfer guidelines or seed-zone designation.

This conclusion can, however, not be generalized. The question of whether multivariate patterns of allozyme variation are useful for describing adaptive variation in other species and in different environmental conditions cannot be answered. This is especially so because only a few studies have applied multivariate statisical techniques to allozymes. With the exception of these studies, whose results completely agree with our findings, little is known about the potential and utility of multivariate allozyme patterns for practical purposes. Moreover, our study as well as the earlier studies investigated allozyme variation in regions with rather pronounced environmental gradients. As suggested by our results from Southwest Oregon, the adaptive significance of allozymes may, however, vary as a function of the environmental variation found in the studied area and the complexity and interaction of the selective agents. Any generalization about the potential and utility of multilocus allozyme variation would thus be merely speculative. However, based on our results, we perfectly agree with the view of MILLAR and WESTFALL (1992), that certain statistical methods may allow allozyme diversity to be partitioned such that the portion of allozyme diversity that correlates with adaptive variation is revealed. We also agree with WESTFALL and CONKLE (1992) that significant, regular geographic patterns exist in allozyme loci in forest trees and that information on allozyme patterns can augment or even replace that of metric traits in certain situations.

Allozymes may, however, not only be useful traits with regards to describing adaptive variation. For conservation purposes, for example, all genetic information should be used together in order to assess patterns of variation important for conservation. Allozyme variation thus may provide useful and important additional information. Since we do not know what genes are or will be useful in future environments, all traits, including allozymes, should be weighted equally, independent of their role in adaptation. Multilocus allozyme variation and other traits may, however, give discordant patterns since the two traits may be reflecting different evolutionary forces. In this case, the species in question may be divided into the smallest reasonable units suggested by cummulative analysis of individual patterns as has been suggested and demonstrated for bishop pine (*Pinus muricata*) by MILLAR and WESTFALL (1992).

11. References

AAGAARD, J.E., SORENSEN, F., STRAUSS, S.H. 1993. Differentiation of RAPD markers between Pacific and North Plateau races of ponderosa pine. Poster presented at the annual meetings and joint conference of the Western Mensurationists Association and the Western Forest Genetics Association.

ADAMS, W.T. 1981. Applying isozyme analyses in tree breeding programs. In: Proceedings, symposium on isozymes of North American forest trees and forest insects. USDA Forest Service, Gen. Tech. Rep. PSW-48: 60-64.

ADAMS, W.T., CAMPBELL, R.K. 1981. Genetic adaptation and seed source specificity. Reforestation of skeletal soils. Proceedings of a workshop, November 17-19, 1981, Medford, Oregon: 78-85.

ADAMS, W.T., JOLY, R.J. 1980. Genetics of allozyme variants in loblolly pine (Pinus taeda L.). J. Heredity 71: 33-40.

ADAMS, W.T., NEAL, D.B., DOERKSEN, A.H., SMITH, D.B. 1990. Inheritance and linkage of isozyme variants from seed and vegetative bud tissue in costal Douglas-fir [Psedotsuga menziesii var. menziesii (Mirb.) Franco]. Silvae Genetica 39 (3-4): 153-167.

AGER, A.A., STETTLER, R.F. 1983. Local variation in seeds of ponderosa pine. Can. J. Botany 61: 1337-1344.

ALEXANDER, B.G., RONCO, F., FITZHUGH, E.L., LUDWIG, J.A. 1984. A classification of forest habitat types of the Lincoln National Forest, New Mexico. USDA Forest Service, Gen. Tech. Rep. RM-104, 29 p.

ALLENDORF, F.W., KNUDSEN, K.L., BLAKE, G.M. 1982. Frequencies of null alleles at enzyme loci in natural populations of ponderosa and red pine. Genetics 100: 497-504.

AMERICAN FORESTRY ASSOCIATION 1956. These are the champs. Part II. Amer. Forests 62: 33-40.

ANDERSON, S.M., MCDONALD, J.F. 1981. Effect of environmental alcohol dehydrogenase on in vivo properties of Drosophila alcohol dehydrogenase. Biochem. Genet. 19: 421-430.

ANTONOVICS, J, BRADSHAW, A.D., TURNER, R.G. 1971. Heavy metal tolerance in plants. Adv. Ecol. Res. 7: 1-85.

BAKER, C.M.A., MANWELL, C. 1977. Heterozygosity of the sheep: Polymorphism of "malic enzyme", isocitrate dehydrogenase (NADP+), catalase and esterase. Austr. J. Biol. Sci. 30: 127-140.

BALDWIN, E.M. 1976. Geology of Oregon. Kendall/Hunt Publishing Company, Dubuque, Iowa.

BARBUJANI, G., MILANI, F. 1986. Diversity of some gene frequencies in European and Asian populations. Effects of Longitude. J. Human Evolution 15: 61-69.

BASSMAN, J.H. 1987. Photosynthesis and water relations of ponderosa pine. In: Ponderosa pine: The species and its management. Symposium Proceedings Sept. 29 - Oct. 1 1987, Spokane, Washington: 47-58.

BECKMANN, J., MITTON, J.B. 1984. Peroxidase allozyme differentiation among succesional stands of ponderosa pine. Amer. Midland Naturalist 112 (1): 43-49.

BERGMANN, F. 1973. Genetische Untersuchungen bei Picea abies mit Hilfe der Isoenzym-Identifizierung. Silvae Genetica 22 (3): 63-66.

BERGMANN, F. 1974. Genetischer Abstand zwischen Populationen. II. Die Bestimmung des genetischen Abstandes zwischen europäischen Fichtenpopulationen (Picea abies) auf der Basis von Isoenzym-Genhäufigkeiten. Silvae Genetica 23 (1-3): 28-32.

BERGMANN, F. 1975a. Adaptive acid phosphatase polymorphism in conifer seeds. Silvae Genetica 24 (5-6): 175-177.

BERGMANN, F. 1975b. Herkunftsidentifizierung von Forstsaatgut auf der Basis von Iyoenzym-Genhäufigkeiten. Allg. Forst- u. J. Ztg. 146 (10): 191-195.

BERGMANN, F. 1978. The allelic distribution at an acid phosphatase locus in Norway spruce (Picea abies) along similar climatic gradients. Theor. Appl. Genet. 52: 57-64.

BERGMANN, F. 1988. Untersuchung der genetischen Differenzierung der Fichte in den Seetaler Alpen an Enzym-Genloci. Forum Genetik-Wald-Forstwirtschaft, 5. Arbeitstagung, Innsbruck, Austria: 163-172.

BERGMANN, F., GREGORIUS, H.R. 1993. Ecogeographical distribution and thermostability of isocitrate dehydrogenase (IDH) alloenzymes in European silver fir (Abies alba). Biochemical Systematics and Ecology 21 (5): 597-605.

BERGMANN, F., GREGORIUS, H.R., LARSEN, J.B. 1990. Levels of genetic variation in European silver fir (Abies alba). Are they related to the species' decline? Genetica 82: 1-10.

BERGMANN, F., SCHOLZ, F. 1987. The impact of air pollution on the genetic structure of Norway spruce. Silvae Genetica 36 (2): 80-83.

BEWLEY, J.D., BLAKE, M. 1985. Seed physiology of development and germination. Plenum Press, New York.

BIJLSMA, R. 1978. Polymorphism at the glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase loci in Drosophila melanogaster. Part 2. Evidence for interactions in fitness. Genet. Res. 31: 227-238.

BORCARD, D., LEGENDRE, P., DRAPEAU, P. 1992. Partialling out spatial components of ecological variation. Ecology 73 (3): 1045-1055.

BOTSTEIN, D., WHITE, R.L., SKOLNICK, K., DAVIS, R.W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphism. Amer. J. Human Genet. 32: 314-331.

BOYLE, T.J.B., MORGENSTERN, E.K. 1987. Some aspects of the population structure of black spruce in central New Brunswick. Silvae Genetica 36: 53-60.

BROTSCHOL, J.V., ROBERDS, J.H., NAMKOONG, G. 1986. Allozyme variation among North Carolina populations of Liliodendron tulipifera L..Silvae Genetica 35 (4): 131-138.

BROWN, A.H., MARSHALL, D.R., MUNDAY, J. 1976. Adaptivness of variants at an alcohol dehydrogenase locus in Bromus mollis L. (soft bromgrass). Austr. J. Biol. Sci. 29: 389-396.

BURGESS, I.P., BELL, J.C. 1983. Comparative morphology and allozyme frequncies of Eucalyptus grandis Hill. ex Maiden and E. saligna Sm. Austr. For. Res. 13: 133-149.

BURTON, R.S., FELDMAN, M.W. 1983. Physiological effects of an allozyme polymorphism: glutamate-pyruvate transaminase and response to hyperosmotic stress in the copepod Tigriopus californicus. Biochem. Genet. 21: 238-251.

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

BUSH, R.M., SMOUSE, P.E. 1992. Evidence of the adaptive significance of allozymes in forest trees. New Forests 6: 179-196.

BUSH, R.M., SMOUSE, P.E., LEDIG, F.T. 1987. The fitness consequences of multiple-locus heterozygosity: the relationship between heterozygosity and growth rate in pitch pine (Pinus rigida Mill.). Evolution 41 (4): 787-798.

CALLAHAM, R.Z., LIDDICOET, A.R. 1961. Altitudinal variation at 20 years in ponderosa and jeffrey pines. Journal of Forestry 59(11): 814-820.

CAMPBELL, R.K. 1974. A provenance-transfer model for boreal regions. Meddelelser fra Norsk Institutt for Skogforskning 31: 545-566.

CAMPBELL, R.K. 1979. Genecology of Douglas-fir in a watershed in the Oregon Cascades. Ecology 60 (5): 1036-1050.

CAMPBELL, R.K. 1984 (1983). Procedures for determining the biological limits of breeding zones in the Pacific Northwest. Proc. Progeny testing, servicewide genetics workshop, Charleston, South Carolina, Dec. 5 - 9, 1983: 24-33.

CAMPBELL, R.K. 1986. Mapped genetic variation of Douglas-fir to guide seed transfer in southwest Oregon. Silvae Genetica 35 (2-3): 85-96.

CAMPBELL, R.K. 1991. Soils, seed-zone maps, and physiography: Guidelines for seed transfer of Douglas-fir in southwestern Oregon. Forest Science 37 (4): 973-986.

CAMPBELL, R.K., PAWUK, W.A., HARRIS, A.S. 1989. Microgeographic genetic variation of Sitka spruce in southeastern Alaska. Can. J. For. Res. 19: 1004-1013.

CAMPBELL, R.K., RITLAND, S.M. 1982. Regulation of seed-germination timing by moist chilling in western hemlock. New Phytol. 92: 173-182.

CAMPBELL, R.K., SORENSEN, F.C. 1978. Effect of test environment on expression of clines and on delimitation of seed zones in Douglas-fir. Theor. Appl. Genet. 51: 233-246.

CAMPBELL, R.K., SUGANO, A.I. 1987. Seed zones and breeding zones for sugar pine in southwestern Oregon. USDA Forest Service, Research paper PNW-RP-383: 1-18.

CASHON, R.E., VAN BENEDEN, R.J., POWERS, D.A. 1981. Biochemical genetics of Fundulus heteroclitus (L.).VI. Spatial variation in gene frequencies of Idh-B, 6-Pgdh-A, and Est-S. Biochem. Genet. 19: 715-718.

CAVALLI-SFORZA, L.L 1974. The genetics of human populations. Sci. Am. 231 (3): 81-89.

CAVENER, D.R., CLEGG, M.T. 1981. Evidence for biochemical and physiological differences between enzyme genotypes in Drosophila melanogaster. Proc. Nat. Acad. Sci. U.S.A. 78: 4444-4447.

CHAMBERS, J.M., HASTIE, T.J. 1992. Statistical models in S. Chapter 8: Local regression models. Wadsworth and Brooks, Cole computer science series: 309-376.

CHARLES, D.J., LEE, C.Y. 1980. Biochemical and immunological characterization of genetic variants of phosphoglucose isomerase from mouse. Biochem. Genet. 18: 153-169.

CHELIAK, W.M., et al. 1984. Segregation of allozymes in megagametophytes of viable seed from a natural population of jack pine, Pinus banksiana Lamb. Theor. Appl. Genet. 69: 145-151.

CHEVERUD, J.M., WAGNER, G.P., DOW, M.M. 1989. Methods for the comparative analysis of variation patterns. Syst. Zool. 38: 201-213.

CHUNG, M.S. 1981. Biochemical methods for determining population structure in Pinus sylvestris. Acta For. Fenn. 173: 1-28.

CLARKE, B. 1975. The contribution of ecological genetics to evolutionary theory: detecting the direct effects of natural selection on particular polymorphic loci. Genetics 79 (6): 101-113.

CLEGG, M.T., ALLARD, R.W. 1972. Patterns of genetic differentiation in the slender wild oat species Avena barbata. Proc. Nat. Acad. Sci., U.S.A. 69: 1820-1824.

CLEVELAND, S.W., DEVLIN, S.J., GROSSE, E. 1988. Regression by local fitting. Methods, properties, and computational algorithms. Journal of Econometrics 37: 87-114.

CLEVELAND, W.S., GROSSE, E. 1991. Computational methods for local regression. Statistics and Computing 1: 47-62.

CLIFF, A.D., ORD, J.K. 1973. Spatial autocorrelation. Pion, London

COHEN, J. 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Measurem. 20: 37-46.

COMPS, B., BARRIERE, G., MERZEAU, D., LETOUZEY, J. 1987. La variabilité alloenzymatique des hêtraies dans les sous-domaines medio- et eu-atlantiques d' Europe. Can. J. For. Res. 17: 1043-1049.

CONKLE, M.T. 1973. Growth data for 29 years from the California elevational transect study of ponderosa pine. Forest Science 19 (1): 31-39.

CONKLE, M.T. 1979. Amount and distribution of isozyme variation in various conifer species. In: Tree improvement in the boreal forest: today and tomorow. Proceedings 17th meeting Canadian tree improvement association, part 2, August 27-30, 1979, Gander

CONKLE, M.T. 1981. Isozyme variation and linkage in six conifer species. USDA Forest Service, Gen. Tech. Rep. PSW-48: 11-17.

CONKLE, M.T., CRITCHFIELD, W.B. 1988. Genetic variation and hybridization of ponderosa pine. In: Ponderosa pine: The species and its management. Proceedings of symposium, Sept. 29 - Oct. 1, 1987, Spokane, Washington.

CONKLE, M.T., HODGSKISS, P.D., NUNNALLY, L.B., HUNTER, S.C. 1982. Starch gel electrophoresis of conifer seeds: a laboratory manual. USDA Forest Service, Gen. Tech. Rep. PSW-6, 21 p.

CONKLE, M.T., WESTFALL, R. 1984. *Evaluating beeding zones for Ponderosa Pine in California.* In: Progeny Testing: Proceedings of servicewide genetics workshop, Charleston, SC, December 5-9, 1983: 89-98.

COOPER, C.F. 1960. Changes in vegetation, structure, and growth of southwestern pine forests since white settlement. Ecol. Monogr. 30: 129-164.

COOPER, S., NEIMAN, K., STEELE, R. 1985. Forest habitat types of northern Idaho. USDA Forest Service, Intermountain Research Station, Missoula, 295 p.

COPES, D.L. 1981. Isoenzyme uniformity in western red cedar seedlings from Oregon and Washington. Can. J. For. Res. 11: 451-453.

CRITCHFIELD, W.B. 1984. Crossability and relationships of Washoe pine. Madrono 31 (3): 144-170.

CRITCHFIELD, W.B., LITTLE, E.L. 1966. Geographic distribution of the pines of the world. Map number 47, p. 80. Miscellaneous puplication 991, U.S. Department of Agriculture, Washington D.C.

CURTSINGER, J.W., LAURIE-AHLBERG, C.C 1991. Genetic variability of flight metabolism in Drosophila melanogaster. I. Characterization of power output during tethered flight. Genetics 98: 549-564.

DALY, K., CLARKE, B. 1981. Selection associated with alcohol dehydrogenase locus in Drosophila melanogaster: differential survival of adults maintained on low concentration of ethanol. Heredity 46: 219-226.

DANCIK, B.P., YEH, F.C. 1983. Allozyme variability and evolution of lodgpole pine (Pinus contorta var. latifolia) and jack pine (Pinus banksiana) in Alberta. Can J. Genet. Cytol. 25: 57-64.

DESPONTS, M., SIMON, J.P. 1987. Structure et variabilité génetique de populations d'epinette noire (Picea mariana (Mill.) B.S.P.) dans la zone hemiarctique de Nouveau-Quebec. Can. J. For. Res. 17: 1006-1012.

DIMICHELE, L., DIMICHELE, J.A., POWERS, D.A. 1986. Developmental and physiological consequences of genetic variation at enzyme sythesizing loci in Fundulus heteroclitus. Am. Zool. 26: 201-210.

DIMICHELE, L., PAYNTER, K., POWERS, D.A. 1991. Lactate dehydrogenase-B allozymes directly effect development of Fundulus heteroclitus. Science 253: 898-900.

DIMICHELE, L., POWERS, D.A. 1982a. Physiological basis for swimming endurance differences between Ldh-B genotypes of Fundulus heteroclitus. Science 216: 1014-1016.

DIMICHELE, L., POWERS, D.A. 1982b. LDH-B genotype-specific hatching times of Fundulus heteroclitus. Nature 296: 563-564.

DIMICHELE, L., POWERS, D.A. 1984. Developmental and oxygen cosumption differences between Ldh-B genotypes of Fundulus heteroclitus and their effect on hatching time. Physiol. Zool. 57: 52-56.

DIMICHELE, L., POWERS, D.A. 1991. Developmental heterochrony and differential mortality in the model teleost, Fundulus heteroclitus. Physiol. Zool. 64.

DICKINSON, T., KNOWLES, P., PARKER, W.H. 1988. Data set congruence in Northern Ontario Tamarack (Larix laricina, Pinaceae). Syst. Bot. 13 (3): 442-455.

DIEHL, W.J., GAFFNEY, P.M., MCDONALD, J.H., KOEHN, R.K. 1985. Relationship between weight-standardized oxygen consumption and multiple-locus heterozygosity in the mussel, Mytilus edulis. In: Gibbs, P.E. (ed.): Proceedings of the nineteenth European marine biology symposium, Cambridge, England: 529-534.

DORADO, G., BARBANCHO, M. 1984. Differential response in Drosophila melanogaster to environmental ethanol: modification of fitness components at the Adh locus. Heredity 53: 309-320.

DYRNESS, C.T., YOUNGBERG, C.T. 1966. Soil-vegetation relationships within the ponderosa pine type in the Central Oregon pumic region. Ecology 47 (1): 122-138.

ECHOLS, R.M., CONKLE, M.T. 1971. The influence of plantation and seed-source elevation on wood specific gravity of 29-year-old ponderosa pine. Forest Science 17 (3): 388-394.

ECKERT, R.T., JOLY, R.J., NEALE, D.N. 1981. Genetics of isozyme variants and linkage relationships among allozyme loci in 35 eastern white pine clones. Can. J. For. Res. 11: 573-579.

EDWARDS, M.D., STUBER, C.W., WENDEL, J.F. 1987. Molecular-marker-faciliated investigations of quantitative trait loci in maize. Genetics 116: 113-125.

ELLSTRAND, N.C., TORRES, A.M., LEVIN, D.A. 1978. Density and the rate of apparent outcrossing in Hellianthus annuus (Asteracea). Syst. Bot. 3: 403-407.

ELSTON, R.C., FORTHOFER, R. 1977. Testing for Hardy-Weinberg-equilibrium in small samples. Biometrics 33: 536-542.

EL-KASSABY, Y.A. 1981. Genetic interpretation of malate dehydrogenase isozymes in some conifer species. J. Heredity 72: 451-452.

EL-KASSABY, Y.A. 1982. Associations between allozyme genotypes and quantitative traits in Douglas-fir [Pseudotsuga menziesii (Mirb.) Franco]. Genetics 101: 103-115.

EL-KASSABY, Y.A., SZIKLAI, O. 1982. Genetic variation of allozyme and quantitative traits in a selected Douglas-fir (Pseudotsuga menziesii var. menziessii (Mirb.)) population. Forest Ecology and Managem. 4: 115-126.

EPPERSON, B.K. 1992. Spatial structure of genetic variation within populations of forest trees. New Forests 6: 257-278.

EPPERSON, B.K., ALLARD, R.W. 1989. Spatial autocorrelation analysis of the distribution of genotypes within populations of lodgepole pine. Genetics 121: 369-377.

ERIKSSON, G., ANDERSSON, S., EICHE, V., ET. AL 1980. Severity index and transfer effects on survival and volume production of Pinus sylvestris in Northern Sweden. Studia Forest. Suecica 156 (2): 28 p.

FALKENHAGEN, E.R. 1985. Isozyme studies in provenance research of forest trees. Theor. Appl. Genet. 69: 335-347.

FARRIS, M.A., MITTON, J.B. 1984. Population density, outcrossing rate, and heterozygote superiority in ponderosa pine. Evolution 38 (5): 1151-1154.

FERET, P.P. 1974. Genetic differences among three small stands of Pinus pungens. Theor. Appl. Genet. 44: 173-177.

FERET, P.P., BERGMANN, F. 1976. Gel electrophoresis of proteins and enzymes. In: Mischke, J.P. (ed.): Modern methods in forest genetics, Springer Verlag, New York: 49-77.

FOLTZ, D.W., NEWKIRK, G.F., ZOUROS, E. 1983. Genetics of growth rate in the American oyster: Absence of interactions among enzyme loci. Aquaculture 33: 157-165.

FOWELLS, H.A. 1965. Silvics of forest trees of the United States. USDA Forest Service, Agricult. Handbook No. 21, 762 p.

FOWLER, D.P., MORRIS, R.W. 1977. Genetic diversity in red pine: evidence for low genic heterozygosity. Can. J. For. Res. 7: 343-347.

FRANKLIN, J.F., DYRNESS, C.T. 1973. Natural vegetation of Oregon and Washington. USDA Forest Service, Gen. Tech. Rep. PNW-8.

FRYER, J.H. 1987. Agreement between patterns of morphological variability and isozym band phenotypes in pitch pine. Silvae Genetica 36 (5-6): 199-206.

FUJINO, K. 1978. Genetic studies on the Pacific abalone. II. Excessive homozygosity in deficiant animals. Bull. Jap. Soc. Fish. 44: 767-770.

FUJINO, K., SASAKI, K. 1984. Age association of genotypic proportions of isozymes in the Pacific abalone. Bull. Jap. Soc. Sci. Fish. 50: 11-15.

FUJIO, Y. 1971. Selective advantage of heterozygotes at B blood group locus in the chicken. Jap. J. Genet. 46: 181-189.

FUJIO, Y. 1982. A correlation of heterozygosity with growth rate in the Pacific oyster, Crassostrea gigas. Tohoku J. Agr. Res. 33: 66-75.

FURNIER, G.E., ADAMS, W.T. 1986. Mating systems in natural populations of jeffrey pine. Amer. J. Bot. 73 (7): 1002-1008.

FURNIER, G.R., ADAMS, W.T. 1986. Geographic patterns of allozyme variation in jeffrey pine. Amer. J. Bot. 73 (7): 1009-1015.

GÖMÖRY, D., VYSNY, J., COMPS, B., THIEBAUT, B. 1992. Geographical patterns of genetic differentiation and diversity in European beech (Fagus sylvatica L.) populations in France. Biologia (Bratislava) 47 (7)

GARTEN, C.T. 1976. Relationships between aggressive behavior and genetic heterozygosity in the old-field mouse, Peromyscus polionotus. Evolution 30: 59-72.

GARTON, D.W. 1984. Relationship between multiple locus heterozygosity and physiological energetics of growth in the estuarine gastropod Thais haemastoma. Physiol. Zool. 57: 530-543.

GARTON, D.W., KOEHN, R.K., SCOTT, T.M. 1984. Multiple-locus heterozygosity and the physiological energetics of growth in the coot clam, Mulinia lateralis, from a natural population. Genetics 108: 445-455.

GAUCH, H.G. 1982. Multivariate analysis in community ecology. Cambridge University Press, Cambridge, England, 298 p.

GEARY, R.C. 1954. The contiguity ratio and statistical mapping. The Incorp. Statistican 5: 115-145.

GEBUREK, TH. 1993. Zur Analyse der genetischen Architektur von Waldbeständen anhand räumlicher Autokorrelationen. Allg. Forst und Jagd Ztg. 164 (8): 137-144.

GIBSON, J. 1970. Enzyme flexibility in Drosophila melanogaster. Nature 227: 959-960.

GILLESPIE, J.H., LANGLEY, CH.H. 1974. A general model to account for enzyme variation in natural populations. Genetics 76 (4): 837-848.

GILLET, E.M. 1994. GSED, Genetic structures from electrophoretic data, version 1.0, user' s manual. Abteilung für Forstgenetik und Forstpflanzenzüchtung Universität Göttingen

GITTINS, R. 1985. Canonical analysis. A review with applications in ecology. Springer Verlag, New York, 351 p.

GONCHARENKO, G.G., PADUTOV, V.E., SILIN, A.E 1993. Allozyme variation in natural populations of Eurasian Pines. Silvae Genetica 42 (4/5): 237-246.

GOVINDARAJU, D.R., DANCIK, B.P. 1987. Environmental stress and the relationship among allozyme heterozygosity, biomass and biomass components in jack pine (Pinus banksiana Lamb.). Genetica 74: 173-179.

GRANT, M.C., LINHART, Y.B., MONSON, R.K. 1989. Experimental studies of ponderosa pine. II. Quantitative genetics of morphological traits. Amer. J. Bot. 76 (7): 1033-1040.

GRANT, M.C., MITTON, J.B. 1977. Genetic differentiation among growth forms of Engelmann spruce and Subalpine fir at tree line. Arctic and Apline Research 9 (3): 259-263.

GRAVES, J.E., ROSENBLATT, R.H., SOMERO, G.N. 1983. Kinetic and electrophoretic differentiation of lactate dehydrogenases of Teleost species-pairs from the Atlantic and Pacific coast of Panama. Evolution 37 (1): 30-37.

GRAVES, J.E., SOMERO, G.N. 1982. Electrophoretic and functional enzymic evolution in four species of eastern Pacific barracudas from different thermal environments. Evolution 36: 97-106.

GREEN, R.H., SINGH, S.M., HICKS, B., MCCUAIG, J. 1983. An arctic intertidal population of Macoma balthica (Mollusca, Pelecypoda): Genotypic and phenotypic components of population structure. Can. J. Fish. Aquat. Sci. 40: 1360-1371.

GREGORIUS, H.R. 1974a. Genetischer Abstand zwischen Populationen. Silvae Genetica 23 (1-3): 22-27.

GREGORIUS, H.R. 1974b. On the concept of genetic distance between populations based on gene frequencies. In: Proceedings Joint IUFRO meeting S.02.04(1-3). Session I. Department of forest genetics, royal college of forestry, Stockholm, Sweden: 17-26.

GREGORIUS, H.R. 1978. The concept of gentic diversity and its formal relationship to heterozygosity and genetic distance. Math. Biosci. 41: 253-271.

GREGORIUS, H.R. 1980. The probability of losing an allele when diploid genotypes are sampled. Biometrics 36: 643-652.

GREGORIUS, H.R. 1986. The importance of genetic multiplicity for tolerance of atmospheric pollution. Proceedings 18 th I.U.F.R.O. World Congress, Ljublijana, Vol I: 295-305.

GREGORIUS, H.R., KRAUHAUSEN, J., MUELLER-STARCK, G. 1986. Spatial and temporal genetic differentiation among the seed in a stand of Fagus sylvatica L.. Heredity 57: 255-262.

GREGORIUS, H.R., ROBERDS, J.H. 1986. Measurement of genetical differentiation among subpopulations. Theor. Appl. Genet. 71: 826-834.

GRIFFIN, A.R. 1978. Geographic variation in Douglas-fir from the coastal ranges of California. II. Predictive value of a regression model for seedling growth variation. Silvae Genetica 27: 96-101.

GRIFFIN, A.R., CHING, K.K. 1977. Geographic variation in Douglas-fir from the coastal ranges of California: I. Seed, seedling growth, and hardiness characteristics. Silvae Genetica 26: 149-157.

GROSSMAN, A. 1980. Analysis of genetic variation affecting the relative activnes of fast and slow ADH dimers in Drosophila melanogaster heterozygotes. Biochem. Genet. 18: 765-780.

GULLBERG, U., YAZDANI, R., RUDIN, D., RYMAN, N. 1985. Allozyme variation in Scots pine (Pinus sylvestris L.) in Sweden. Silvae Genetica 34 (6): 193-201.

GURIES, R.P. 1984. Genetic variation and population differentiation in forest trees. Proceedings of eight north American forest biology workshop, July 30-August 1, 1984, Logan, Utah: 119-131.

GURIES, R.P., LEDIG, F.T. 1978. Inheritance of some polymorphic isoenzymes in pitch pine (Pinus rigida Mill.). Heredity 40 (1): 27-32.

GURIES, R.P., FRIEDMAN, S.T., LEDIG, F.T. 1978: A megagametophyte analysis of genetic linkage in pitch pine (Pinus rigida Mill.). Heredity 40 (2): 309-314.

GURIES, R.P., LEDIG, F.T. 1981. Genetic structure of populations and differentiation in forest trees. In: Proccedings of the symposium on isozymes of North American forest trees and forest insects. USDA Forest Service, Gen. Tech Rep. PSW-48: 42-47.

HABERMAN, S.J. 1978. Analysis of qualitative data. Academic Press, New York, 368 p.

HAGNER, M. 1970. A genecological investigation of the annual rhythm of Pinus silvestris L. Studia Forest. Suec. 80: 1-40.

HAIR, J.F., ANDERSON, R.E., TATHAM, R.L. 1987. Multivariate data analysis, 2nd edition. Macmillan Publ. Co., New York, 449 p.

HALDANE, J.B.S. 1954. Heterozygote frequencies in small populations. Theor. Popul. Biol. 8: 31-48.

HALLER, J.R. 1965. The role of 2-needle fascicles in the adaptation and evolution of ponderosa pine. Britonia 17: 354-382.

HAMRICK, J.L, LINHART, Y.B., MITTON, J.B. 1979. Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. Ann. Rev. Ecol. Syst. 10: 173-200.

HAMRICK, J.L., BLANTO, H.M., HAMRICK, K.J. 1989. Genetic structure of geographically marginal populations of Ponderosa pine. Amer. J. Bot. 76 (11): 1559-1568.

HAMRICK, J.L., GODT, M.J.W. 1990. Allozyme diversity in plant species. In: Brown, A.H. et al.,1990: Plant population genetics, breeding, and genetic resources. Sinauer Ass., Sunderland: 43-63.

HAMRICK, J.L., GODT, M.J.W., SHERMAN-BROYLES, S. 1992. Factors influencing levels of genetic diversity in woody plant species. New Forests 6: 95-124.

HAMRICK, J.L., MITTON, J.B., LINHART, Y.B. 1981. Levels of genetic variation in trees: Influence of life history characteristics. USDA Forest Service, Gen. Tech. Rep. PSW-48

HARLOW, W.H., HARRAR, E.S. 1969. Textbook of dendrology, McGraw-Hill, New York, 512 S.

HARPER, J.L. 1977. Population biology of plants. Academic Press, New York

HARPER, R.A., AMSTRONG, F.B. 1973. Alkaline phosphatase of Drosophila melanogaster.II. Biochemical comparison among four allelic forms. Biochem. Genet. 10: 29-38.

HARRIS, R.J. 1975. A primer of multivariate statistics. Academic Press, New York, 332 p.

HARRY, D.E. 1983. Identification of a locus modifying the electrophoretic mobility of malate dehydrogenase isozymes in incense-cedar (Calcocedrus decurrens), and its implications for population studies. Biochem. Genetics 21 (5/6): 417-434.

HARTL, D.L., CLARK, A.G. 1989. Principles of population genetics. Sinauer, Sunderland, Mass., 682 p.

HATTEMER, H.H., GREGORIUS, H.R., ZIEHE, M., MÜLLER-STARCK, G. 1982. Klonzahl forstlicher Samenplantagen und genetische Vielfalt. Allg. Forst- u. J. Ztg. 153: 183-191.

HERMANN, R.K., LAVENDER, D.P. 1968. Early growth of Douglas-fir from various altitudes and aspects in southwest Oregon. Silvae Genetica 17: 143-151.

HIERL, H.F. 1976. Relationship between heterozygosity as estimated from genetic markers and fertility in cattle. II. Heterozygosity and fertility. Theor. Appl. Genet. 47: 77-84.

HOFFMAN, R. 1981. Evolutionary genetics of Metridium senile. I. Kinetic differences in phosphoclucose isomerase allozymes. Biochem. Genet. 19: 129-144.

HOLZER, K. 1982. Die Entwicklung der (forstlichen Saatgut-) Gesetzgebung im Zusammenhang mit der Herkunftsfrage. Forum Genetik-Wald-Forstwirtschaft, 2. Arbeitstagung, Göttingen, Germany: 77-78.

HUGHES, M.B., LUCCHESI, J.C. 1977. Genetic rescue of a lethal "null" activity allele of 6-phosphogluconate dehydrogenase in Drosophila melanogaster. Science 196: 1114-1115.

JACOBS, B.F., WERTU, C.R., GUTTMAN, S.I. 1984. Genetic relationships in Abies (fir) of eastern United States: an electrophoretic study. Can. J. Bot. 62: 609-616.

JENKINSON, J.L. 1974. Ponderosa pine progenies: Differential response to ultramafic and granitic soils. USDA Forest Service, Research paper PSW-101: 13 p.

JOHNSGARD, G.A. 1963. Temperature and water balance for Oregon weather stations. Agricultural Experiment Station, Oregon State University, Corvallis, 126 p.

JOHNSON, A.G., UTTER, F.M. 1975. Population differences of aspartate aminotransferase and peptidase in the bay mussel Mytilus edulis. Anim. Blood Groups Biochem. Genet. 6: 71-80.

JOHNSON, D.H. 1981. How to measure - a statistical perspective. In: Capen, D.E. (ed.); 1981: The use of multivariate statistics in studies on wildlife habitat. USDA Forest Service, Gen. Tech. Rep. RM-87

JOHNSON, F.M., SCHAFFER, H.E. 1973. Isozyme variability in species of the genus Drosophila. VII. Genotype environment relationships in populations of D. melanogaster from eastern United States. Biochem. Genet. 10: 149-163.

JOHNSON, F.M., SCHAFFER, H.E., GILLASPY, J.E., ROCKWOOD, E.S. 1969. Isozyme genotype-environment relationships in natural populations of the harvester ant, Pogonomyrmex barbatus, from Texas. Biochem. Genet. 3: 429-450.

JOHNSON, G.B. 1974. Enzyme polymorphism and metabolism. Polymorphism among enzyme loci is related to metabolic function. Science 184 (1): 28-37.

JOHNS, P.E, et al. 1979. Reproductive patterns, productivity and genetic variability in adjacent white-tailed deer populations. Proc Ann. Conf. Southeast Assoc. Fish Widlife Agenc. 31: 167-172.

JONES, J.S., SELANDER, R.K., SCHNELL, G.D. 1980. Patterns of morphological and molecular polymorphism in the land snail Cepaea nemoralis. Biol. J. Linn. Soc. 14: 359-387.

JUMARS, P.A. 1978. Spatial autocorrelation with RUM (Remote Underwater Manipulator): vertical and horizontal structure of a bathal bentic community. Deep Sea Res. 25: 589-604.

JUMARS, P.A., THISTLE, D., JONES, M.L. 1977. Detecting two dimensional spatial structure in biological data. Oecologia 28: 109-123.

KAHLER, A.L., WEHRHAN, C.F. 1986. Associations between quantitative traits and enzyme loci in the F2 population of a maize hybrid. Theor. Appl. Genet. 72: 15-26.

KAHLER, A.L., et al. 1980. Associations between isozyme phenotypes and environment in the slender wild oat (Avena barbata) in Israel. Theor. Appl. Genet. 56: 31-47.

KEEN, F.P. 1940. Longevity of ponderosa pine. Journal Forestry 38: 597-598.

KIMURA, M. 1968. Evolutionary rate at the molecular level. Nature 217: 624-626.

KIMURA, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, 367 p.

KIMURA, M., OTHA, T. 1971. Protein polymorphism as a phase of molecular evolution. Nature 229: 467-469.

KIM, Z.S. 1980. Veränderungen der genetischen Struktur von Buchenpopulationen durch Viabilitätsselektion im Keimlingsstadium. Diss. Georg-August Universität Göttingen, 84 p.

KING, D.P.F. 1985. Enzyme heterozygosity associated with anatomical character variance and growth in the herring (Clupea harengus L.). Heredity 54: 289-296.

KING, J.L., JUKES, T.H. 1969. Non-Darwian evolution: random fixation of selectively neutral mutations. Science 164: 788-798.

KITZMILLER, J.H. 1990. Genetic variation and adaptability of Douglas-fir in northwestern California. Joint Meeting of Western Forest Genetics Association and IUFRO working Parties S 2.02-05,06,12,14, August 20 - 24, 1990, Olympia, Washington, USA: Sect. 2:181.

KLECKA, W.R 1975. Discriminant analysis. In: Nie, H.H et al., 1975: SPSS: Statistical package for the social sciences. Second edition, Mc Graw-Hill, New York, 675 p.

KLECKA, W.R. 1980. Discriminant analysis. Sage University paper series quantitative applications in the social sciences, series number 07-019. Sage publications, Beverly Hills and London.

KNOWLES, P. 1984. Genetic variability among and within closely spaced populations of lodgpole pine. Can. J. Genet. Cytol. 26: 177-184.

KNOWLES, P., GRANT, M.C. 1981. Genetic patterns associated with growth variability in ponderosa pine. Amer. J. Bot. 68 (7): 942-946.

KNOWLES, P., MITTON, J.B. 1980. Genetic heterozygosity and radial growth variability in Pinus contorta. Silvae Genetica 29 (3-4): 114-118.

KOEHN, R.K. 1969. Esterase Heterogeneity: Dynamics of a polymorphism. Science 163: 943-944.

KOEHN, R.K. 1978. Physiology and biochemistry of enzyme variation: The interface of ecology and population genetics. In: Brussard, P.F. (ed.), 1978: Ecological genetics: The interface. Springer-Verlag, New-York: 51-72.

KOEHN, R.K., GAFFNEY, P.M. 1984. Genetic heterozygosity and growth rate in Mytilus edulis. Mar. Biol. 82: 1-7.

KOEHN, R.K., SHUMWAY, S.E. 1982. A genetic/physiological explanation for differential growth rate among individuals of the American oyster, Crassostrea virginica (Gmelin). Mar. Biol. Lett. 3: 35-42.

KOEHN, R.K., TURANO, F.J., MITTON, J.B. 1973. Population genetics of marine pelecypods. II. Genetic differences in microhabitats of Modiolus demissus. Evolution 27: 100-105.

KOEHN, R., IMMERMANN, F.W. 1981. Biochemical studies of aminopeptidase polymorphism in Mytilus edulis I. Dependence of enzyme activity on season, tissue and genotype. Biochem. Genet. 19: 1115-1142.

KONNERT, M. 1991. Die Fichte (Picea abies (L.) Karst.) im Schwarzwald: Genetische Variation und Korrelationen. Forstw. Cbl. 110: 84-94.

KONNERT, M. 1991. Vergleich der genetischen Struktur verschiedener Generationen zweier natürlich verjüngter Fichtenbestände des Schwarzwaldes. Silvae Genetica 40 (2): 60-65.

KONNERT, M. 1992. Genetische Untersuchungen in geschädigten Weisstannenbeständen (Abies alba Mill.) Südwestdeutschlands. Mitteilungen Forstliche Versuchs- und Forschungsanstalt Baden-Württemberg 167, 119 p.

KONNERT, M. 1993. Untersuchungen über die genetische Variation der Weisstanne (Abies alba Mill.) in Bayern. Allg. Forst.- u. Jagd Ztg. 164 (9-10): 162-169.

KONNERT, M., BERGMANN, F. 1995. The geographical distribution of genetic variation of silver fir (Abies alba, Pinaceae) in relation to its migration history. Pl. Syst. Evol. 196: 19-30.

KORSTIAN, C.F. 1924. A silvical comparison of the Pacific Coast and Rocky Mountain forms of western yellow pine. Am. J. Bot. 11: 318-324.

KREMER, A., et al.. 1991. Nuclear and organell gene diversity in Quercus robur and Q. petraea. In: Mueller-Starck, G., Ziehe, M.: Genetic variation of European populations of forest trees: 141-167.

KRUCKEBERG, A.R. 1967. Ecotypic response to ultramafic soils by plant species of northwestern United States. Brittonia 19: 133-151.

KRUCKEBERG, A.R. 1987. Serpentine endemism and rarity. Conservation of rare and endangered plants. Proceedings from a Conference of the California Native Plant Society, Sacramento, edited by T.S. Elias: 121-128.

LACHENBRUCH, P.A. 1975. Discriminant analysis. Hafner Press, New York, 128 p.

LAGERCRANTZ, U., RYMAN, N. 1990. Genetic structure of Norway spruce (Picea abies): Concordance of morphological and allozymic variation. Evolution 44 (1): 38-53.

LANGLET, O. 1936. Studien ueber die physiologische Variabilität der Kiefer und deren Zusammenhang mit dem Klima. Beiträge zur Kenntnis der Oekotypen von Pinus sylvestris L. Meddelanden fran Statens skogsfoersoeksanstalt 29: 219-470.

LATTER, B.D.H. 1980. Genetic differences within and between populations of the major human subgroups. Am. Nat. 116: 220-237.

LEDIG, F.T., CONKLE, M.T. 1983. Gene diversity and genetic structure in a narrow endemic, Torrey pine (Pinus torreyana Parry ex Carr.). Evolution 37 (1): 79-85.

LEDIG, F.T., FRYER, J.H. 1972. A pocket of variability in Pinus rigida. Evolution 26: 259-266.

LEDIG, F.T., GURIES, R.P., BONEFELD, B.A. 1983. The relation of growth to heterozygosity in pitch pine. Evolution 37 (6): 1227-1238.

LEDIG, F.T., ZOBEL, B.J., MATTHIAS, M.F. 1975. Geoclimatic patterns in specific gravity and tracheid length in wood of pitch pine. Can. J. For. Res. 5: 318-329.

LEDIG, F.T, LAMBETH, C.C., LINZER, D.I.H. 1976. Nursery evaluation of a pitch pine provenance trial. In: Proceedings 23rd northeast forest tree improv. conf., New Brunswick, N.J.: 93-108.

LEE, P.J. 1969. The theory and application of trend surfaces. J. Geology 77 (3): 303-318.

LEVENE, H. 1949. On a matching problem arsising in genetics. Ann. Math. Stat. 20: 91-94.

LEVINS, R. 1963. Theory of fitness in a heterogeneous environment. II. Developmental flexibility and nich selection. Am. Naturalist 97: 1093-1097.

LEVINS, R. 1969. Dormancy as an adaptive strategy. Symp. of Soc. for Exper. Biol. No. XXIII. Academic Press, New York.

LEVY, M., LEVIN, D.A. 1975. Genetic heterozygosity and variation in permanent translocation heterozygotes of the Oenothera biennis complex. Genetics 79: 493-512.

LEWONTIN, R.C. 1972. The apportionment of human diversity. Evol. Biol. 6: 381-398.

LEWONTIN, R.C. 1984. Detecting population differences in quantitative characters as oposed to gene frequencies. American Naturalist 123: 115-124.

LEWONTIN, R.C. 1985. Population genetics. Ann. Rev. Genet. 19: 81:102.

LINHART, Y.B., GRANT, M.C., MONTAZER, P. 1989. Experimental studies in Ponderosa pine. I. Relationship between variation in proteins and morphology. Am. J. Bot. 76 (7): 1024-1032.

LINHART, Y.B., MITTON, J.B. 1985. Relationships among reproduction, growth rates, and protein heterozygosity in ponderosa pine. Amer. J. Botany 72 (2): 181-184.

LINHART, Y.B., MITTON, J.B., BOWMAN, D.M., STURGEON, K.B. 1979. Genetic aspects of fertility differentials in ponderosa pine. Genet. Res. (Camb.) 33: 237-242.

LINHART, Y.B., MITTON, J.B., STURGEON, K.B., DAVIS, M.L. 1979. An analysis of genetic architecture in populations of ponderosa pine. USDA Forest Service, Gen. Tech. Rep. PSW-48: 53-59.

LINHART, Y.B., MITTON, J.B., STURGEON, K.B., DAVIS, M.L. 1981. Genetic variation in space and time in a population of ponderosa pine. Heredity 46 (3): 407-426.

LITTLE, E.L., JR. 1971. Atlas of United States trees, vol 1. Conifers and important hardwoods. Misc. Publ. 1146, U.S. Department of Agriculture, Forest Service, Washington, D.C.

LI, P., ADAMS, W.T. 1989. Range-wide patterns of allozyme variation in Douglas-fir (Pseudotsuga menziesii). Can. J. For. Res. 19 (2): 149-161.

LÖCHELT S. 1994. Bestimmung genetischer Merkmale von Fichten (Picea abies (L.) Karst.) mit unterschiedlich ausgeprägten Schadsymptomen auf baden-württembergischen Dauerbeobachtungsflächen. Allg. Forst- u. Jagd Ztg. 165 (2): 21-27.

LOOPSTRA, C. 1984. Patterns of genetic variation within and among breeding zones of Douglas-fir in southwest Oregon. M.Sc. thesis, Oregon State University, Corvallis, OR

LUNDKVIST, K. 1979. Allozyme frequency distribution in four Swedish populations of Norway spruce (Picea abies K.). Hereditas 90: 127-143.

LUNDKVIST, K., RUDIN, D. 1977. Genetic variation in eleven populations of Picea abies as determined by isozyme analysis. Herditas 85: 67-74.

MADSEN, J.L., BLAKE, G.M. 1977. Ecological genetics of ponderosa pine in the northern Rocky Mountains. Silvae Genetica 26 (1): 1-8.

MAKAVEEV, T., VENEV, I., BAULOV, M. 1977. Investigations on activity level and polymorphism of some blood enzymes in farm animals with different growth energy. Genet. Plant. Breed. (Sofia) 10: 229-237.

MANLY, B.F.J. 1985. The statistics of natural selection on animal populations. Chapman and Hall, London, 488 p.

MANOS, P.S., FAIRBROTHERS, D.E. 1987. Allozyme variation in populations of six northeastern american red oaks (Fagaceae: Quercus subg. Erythrobalanus). Syst. Bot. 12 (3): 365-373.

MANTEL, N. 1967. The detection of desease clustering and a generalized regression approach. Cancer research 27: 209-220.

MANTEL, N., VALAND, R.S. 1970. A technique of nonparametric multivariate analysis. Biometrics 26: 547-558.

MARSHALL, D.R., BROUE, P., PRYOR, A.J. 1973. Adaptiv significance of alcohol dehydrogenase isozyme in maize. Nature New Biology 244: 16-18.

MCCULLAGH, P., NELDER, J.A. 1983. Generalized linear models. Chapman and Hall, London and New York, 261 p.

MCLEAN, A. 1970. Plant communities of the Similkameen Valley, British Columbia, and their relationship to soils. Ecol. Monogr. 40 (4): 403-423.

MEJNARTOWICZ. L. 1983. Change in genetic structure of Scots pine (Pinus sylvestris L.) population affected by industrial emission of fluoride and sulphur dioxide. Genetica Polonica 24: 41-50.

MEJNARTOWICZ, L. 1979. Genetic variation in some isoenzyme loci in Scots pine (Pinus silvestris L.) populations. Arboretum Kornickie 24: 91-103.

MENOZZI, P., PIAZZA, A., CAVALLI-SFORZA, L.L 1978. Synthetic maps of human gene frequencies in Europe. Science 201: 786-792.

MERKLE, S.A., ADAMS, W.T. 1987. Patterns of allozyme variation within and among Douglas-fir breeding zones in southwest Oregon. Can. J. For. Res. 17: 402-407.

MERKLE, S.A., ADAMS, W.T., CAMPBELL, R.K. 1988. Multivariate analysis of allozyme variation patterns in costal Douglas-fir from southwest Oregon. Can. J. For. Res. 18: 181-187.

MERRITT, R.B. 1972. Geographic distribution and enzymatic properties of lactate dehydrogenase allozymes in the fathead minnow, Pimephales proleus. Am. Nat. 196: 173-184.

MIKOLA, J. 1982. Bud-set phenology as an indicator of climatic adaptation of Scots pine in Finland. Silva Fennica 16: 178-184.

MILLAR, C.I. 1985. Inheritance of allozyme variants in bishop pine (Pinus muricata D. Don). Biochem. Genetics 23 (11-12): 933-946.

MILLAR, C.I. 1989. Allozyme variation of bishop pine associated with pygmy-forest soils in northern California. Can. J. For. Res. 19: 870-879.

MILLAR, C.I., STRAUSS, S.H., CONKLE, M.T., WESTFALL, R.D. 1988. Allozyme differentiation and biosystematics of the Californian closed-cone pines (Pinus subsect. Oocarpae). Syst. Bot. 13 (3): 351-370.

MILLAR, C.I., WESTFALL, R.D. 1992. Allozyme markers in forest genetic conservation. New Forests 6: 347-371.

MILLER, S., PEARCY, R.W., BERGER, E. 1975. Polymorphism at the alphaglycerophosphate dehydrogenase locus in Drosophila melanogaster I. Properties of adult allozymes. Biochem. Genet. 13: 175-188.

MIROV, N.T., DUFFIELD, J.W., LIDDICOET, A.R. 1952. Altitudinal races of Pinus ponderosa - a 12-year progress report. Journal of Forestry 50 (11): 825-831.

MITTON, J.B. 1983. Conifers. In: Tanksley, S.D., Orton, T.J. (ed.), 1983: Isozymes in plant genetics and breeding, Part B, Elsevier, Amsterdam: 443-472.

MITTON, J.B., GRANT, M.C. 1980. Observations on the ecology and evolution of quaking aspen, Populus tremoloides, in the Colorado front range. Am. J. Bot. 67: 202-209.

MITTON, J.B., JEFFERS, R.M. 1989. The genetic consequences of mass selection for growth rate in Engelmann spruce. Silvae Genetica 38 (1): 6-12.

MITTON, J.B. et al. 1981. Associations between heterozygosity and growth rate variables in three western forest trees. USDA Forest Service, Gen. Tech. Rep. PSW-48: 27-34.

MITTON, J.B., KOEHN, R.K. 1975. Genetic organization and adaptive response of allozymes to ecological variables in Fundulus heteroclitus. Genetics 79: 97-111.

MITTON, J.B., LINHART, Y.B., HAMRICK, J.L., BECKMANN, J.S. 1977. Observation on the genetic structure and mating system of ponderosa pine in the Colorado Front Range. Theor. Appl. Genet. 51: 5-13.

MITTON, J.B., LINHART, Y.B., STURGEON, K.B., HAMRICK, J.L. 1979. Allozyme polymorphisms detected in mature needle tissue of ponderosa pine. J. Heredity 70: 86-89.

MITTON, J.B., STURGEON, K.B., DAVIS, M.L. 1980. Genetic differentiation in ponderosa pine along a steep elevational transect. Silvae Genetica 29 (3-4): 100-103.

MITTON, J.B., STUTZ, H.P., SCHUSTER, W.S., SHEA, K.L. 1989. Genotypic differentiation at PGM in Engelmann spruce from wet and dry sites. Silvae Genetica 38 (5-6): 217-221.

MOLLER, K. 1986. Genetische Untersuchungen bei der Tanne mit Hilfe von Enzym-Genmarkern. Allg. Forstzeitung (Wien) 97: 60-61.

MONSON, R.K., GRANT, M.C. 1989. Experimental studies in ponderosa pine. III. Differences in photosynthesis, stomatal conductance, and water-use efficiency between two genetic lines. Americ. J. Bot. 76 (7): 1041-1047.

MOON, T.W., HOCHACHKA, P.W. 1971. Temperature and enzyme activity in poikolotherms: Isocitrate dehydrogenases in rainbow-trout liver. Comp. Biochem. Physiol. 123: 695-705.

MOON, T.W., HOCHACHKA, P.W. 1972. Temperature and the kinetic analysis of trout isocitrate dehydrogenases. Comp. Biochem. Physiol. 124: 725-730.

MORAN, G.F., BELL, J.C., ELDRIDGE, K.G. 1988. The genetic structure an the conservation of the five natural populations of Pinus radiata. Can. J. For. Res. 18: 506-514.

MORAN, G.F., BELL, J.C., TURNBULL, J.W. 1989b. A cline in genetic diversity in river sheoak Casuarina cunninghamiana. Austr. J. Bot. 37: 169-180.

MORAN, G.F., HOPPER, S.D. 1987. Conservation of genetic resources of rare and widespread eucalypts in remnant vegetation. In: Saunders, D.A. et al. (eds.): Nature conservation: The role of remnants of nativ vegetation. Surrey Beatty, Sydney: 151-162.

MORAN, G.F., MUONA, O., BELL, J.C. 1989a. Breeding systems and genetic diversity in Accacia auriculiformis and A. crassicarpa. Biotropica 21.

MORAN, P.A.P. 1950. Notes on continuous stochastic phenomena. Biometrika 37: 17-23.

MORGANTE, M., VENDRAMIN, G.G., GIANNINI, R. 1993. Inheritance and linkage relationships of isozyme variants of Pinus leucodermis Ant. Silvae Genetica 42 (4-5): 231-237.

MORRISON, D.G. 1976. Multivariate statistical methods, 3rd edition. McGraw-Hill, New York, 322 p.

MORTON, J.R., GILMOUR, D.G., MCDERMIND, E.M., OGDEN, A.L. 1965. Association of blood-group and protein polymorphisms with embryonic mortality in the chicken. Genetics 51: 97-107.

MUELLER-STARCK, G. 1987. Genetic differentiation among seed samples from provenances of Pinus sylvestris. Silvae Genetica 36 (5-6): 232-238.

MUELLER-STARCK, G. 1989. Genetic implications of environmental stress in adult forest stands of Fagus sylvatica L. In: Scholz, F. et al. (eds.): Genetic effects of air pollutants in forest tree populations. Springer, Berlin: 127-142.

MUELLER-STARCK, G. 1991. Survey of genetic variation as inferred from enzyme gene markers. In: Mueller-Starck, G., Ziehe, M.: Genetic variation in European populations of forest trees: 21-37.

MUELLER-STARCK, G., GREGORIUS, H.R. 1986. Monitoring genetic variation in forest tree popultions. In: Donaubauer, E. (ed.): 18th IUFRO World Congress, Ljublijana, Yugoslavia, division 2, Vol II: 589-599.

MUELLER-STARCK, G., ZIEHE, M. 1991. Genetic variation in populations of Fagus sylvatica L., Quercus robur L. and Q. petraea Liebl. in Germany. In: Mueller-Starck, G., Ziehe, M.: Genetic variation of European populations of forest trees: 125-141.

MUELLER-STARCK, G., HATTEMER, H.H. 1989. Genetische Auswirkungen von Umweltstress auf Altbestände und Jungwuchs der Buche (Fagus silvatica L.). Forstarchiv 60 (1): 17-22.

MUKAI, T., YAMAZAKI, T. 1980. Test for selection on polymorphic isozyme genes using the population cage method. Genetics 96: 537-542.

MUONA, O. 1990. Population genetics in forest tree improvment. In: Brown, H.D., Clegg, M.T., Kahler, A.L., Weir, B.S.: Plant population genetics, breeding, and genetic resources, Sinauer Ass., Sunderland: 282-299.

MUONA, O., SZMIDT, E. 1985. A multilocus study of natural populations of Pinus sylvestris. In: Population genetics in forestry. Lectures in biomathematics 60: 226-240.

NAMKOONG, G. 1969. Nonoptimality of local races. In: Proceedings 10th southern For. Tree Improv. Conf., June 17-19,1969, Houston, Texas: 149-153.

NAMKOONG, G., KANG, H., BROUARD, J.S. 1988. Tree breeding: Principles and strategies. Springer Verlag, New York, 180 p.

NARISE, S. 1979. Biochemical differences between cytoplasmatic malate dehydrogenase allozymes of Drosophila virilis. Biochem. Genet. 17: 433-444.

NEI, M. 1972. Genetic distance between populations. Am. Naturalist 106: 283-292.

NEI, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Nat. Acad. Sci. U.S.A. 70: 3321-3323.

NEI, M. 1974. A new measure of genetic distance. In: Crow, J.F., Denniston, C. (eds.): Genetic distance. Plenum Press, New York and London: 63-76.

NEI, M. 1975. Molecular population genetics and evolution. North Holland, Publishing Inc., Amsterdam, 288 p.

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

NEI, M. 1987. Molecular evolutionary genetics. Columbia university press, New York, 512 p.

NEI, M., GRAUR, D. 1984. Extent of protein polymorphism and the neutral mutation theory. Evol. Biol. 17: 73-118.

NELDER, K.A., WEDDERBURN, R.W.M. 1972. Generalized linear models. J. Royal Stat. Soc., series A, 135: 370-384.

NEVO, E., BEILES, A., BEN-SHLOMO, R. 1984. The evolutionary significance of genetic diversity: ecological demographic and life-history correlates. In: Mani, G.S. (ed.): Evolutionary dynamics of genetic diversity. Springer Verlag, Berlin: 13-213.

NEVO, E., BEILES, A., KRUGMAN, T. 1988. Natural selection of allozyme polymorphisms: a microgeographic climatic differentiation in wild emmer wheat (Triticum dicoccoides). Theor. Appl. Genet. 75: 529-538.

NEVO, E., LAVIE, B., BEN-SHLOMO, R. 1983. Selection of allozyme. In: Ratayyi, M.C., Scandalios, J.G., White, G.S. (eds.): Isozymes: Current topics in biological and medical research. Liss Publ., New-York

NEVO, E., PEARL, T., BEILES, A., WOOL, D. 1981. Mercury selection of allozyme genotypes in shrimps. Experientia 37: 1152-1154.

NEVO, E., SHIMONY, T., LIBNI, M. 1977. Thermal selection of allozyme polymorphisms in barnacles. Nature 267: 699-701.

NEVO, E., SHIMONY, T., LIBNI, M. 1978. Pollution selection of allozyme polymorphisms in barnacles. Experientia 34: 1562-1564.

NEVO, E., et al. 1981. Microgeographic edaphic differentiation in allozyme polymorphisms of wild barley (Hordeum spontaneum, Poaceae). Plant Syst. Evol. 138: 287-292.

NEVO, E., et al.. 1980. Genetic structure as a potential monitor of marine pollution. 5em journees etudes pollution, Cagliari: 61-68.

NIEBLING, CH.R., CONKLE, M.T. 1990. Diversity of Washoe pine and comparison with allozymes of ponderosa pine races. Can. J. For. Res. 20: 298-308.

OAKESHOTT, J.G., et al.. 1982. Alcohol dehydrogenase and glycerol-3-phosphate dehydrogenase clines in Drosophila melanogaster on different continents. Evolution 36: 86-96.

O'BRIEN, S.J., MACINTYRE, R.J. 1972. The alpha-glycerophosphate cycle in Drosophila melanogaster. I. Biochemical and developmental aspects. Biochem. Genet. 7: 141-161.

O'MALLEY D., WHEELER, N.C., GURIES, R.P. 1980. A manual for starch gel electrophoresis. University of Wisconsin, Madison, Dept. Forestry, Staff paper series No 11, 16 p.

O'MALLEY, D.M., ALLENDORF, F.W., BLAKE, G.M. 1979. Inheritance of isozyme variation and heterozygosity in Pinus ponderosa. Biochem. Genet. 17 (3-4): 233-250.

O'REILLY, G.J., PARKER, W.H., CHELLIAK, W.M. 1985. Isozyme differentiation of upland and lowland Picea mariana stands in northern Ontario. Silvae Genetica 34 (6): 214-221.

PAPEL, I., et al. 1979. Drosophila alcohol dehydrogenase activity in vitro and in vivo: effects of acetone feeding. Biochem. Genet. 17: 553-563.

PAYNTER, K.T., DIMICHELE, L., HAND, S.C., POWERS, D.A. 1991. Metabolic implications of Ldh-B genotype during early development in Fundulus heteroclitus. J. Exp. Zool. 257: 24-33.

PELOQUIN, R.L. 1984. The identification of three-species hybrids in the ponderosa pine complex. The Southwestern Naturalist 29 (1): 115-122.

PERRY, D.J., KNOWLES, P. 1991. Spatial genetic structure within three sugar maple (Acer saccharum Marsh.) stands. Heredity 66: 137-142.

PFISTER, R.D., KOVALCHIK, B., ARNO, S.F., PRESBY, R.C. 1977. Forest habitat types of Montana. USDA Forest Service Gen. Tech. Rep. INT-34, 174 p.

PIGLIUCCI, F., VILLANI, F., BENEDETTELLI, S. 1990. Geographic and climatic factors associated with the spatial structure of gene frequencies in Castanea sativa Mill. from Turkey. J. Genetics 69 (3): 141-149.

PIGLIUCCI, M., BENEDETTELLI, S., VILLANI, F. 1991. Spatial patterns of genetic variability in Italian chestnut (Castanea sativa Mill.). Can. J. Bot. 68: 1962-1867.

PLACE, A.R., POWERS, D.A. 1979. Genetic variation and relative catalytic efficiencies: LDH-B allozymes of Fundulus heteroclitus. Proc. Nat. Acad. Sci. U.S.A. 76: 2354-2358.

PLACE, A.R., POWERS, D.A. 1984a. The lactate dehydrogenase (LDH-B) allozymes of Fundulus heteroclitus (Lin.): I. Purification and characterisation. J. Biol. Chem. 259: 1299-1308.

PLACE, A.R., POWERS, D.A. 1984b. The LDH-B allozymes of Fundulus heteroclitus. II. Kinetic analyses. J. Biol. Chem. 259: 1309-1318.

PLACE, A.R., POWERS, D. 1979. Genetic variation and relative catalytic efficiencies: lactate dehydrogenase B allozymes of Fundulus heteroclitus. Proc Nat. Acad. Sci. U.S.A. 76: 2534-2538.

PLESSAS, M.E., STRAUSS, S.H. 1986. Allozyme differentiation among populations, stands, and cohorts in Monterey pine. Can. J. For. Res. 16: 1155-1164.

POWELL, J.R. 1975. Protein variation in natural populations of animals. Evol. Biol. 8: 79-119.

POWERS, D.A., LAUERMANN, T., CRAWFORD, D., DIMICHELE, L. 1991. Genetic mechanisms for adapting to a changing environment. Annual Review Genet. 25: 629-659.

RAINEY, D.Y., MITTON, J.B., MONSON, R.K. 1987. Associations between enzyme genotypes and dark respiration in perennial ryegrass, Lolium perenne L. Oecologia 74: 335-338.

RAJORA, O.P., ZSUFFA, L., DANCIK, B.P. 1991. Allozyme and leaf morphological variation of eastern cottonwood at the northern limits of its range in Ontario. Forest Science 37 (2): 688-702.

RASMUSEN, B.A., HAGEN, K.L. 1973. The H blood-group system and reproduction in pigs. J. Anim. Sci. 37: 568-573.

READ, R.A. 1980. Genetic variation in seedling progeny of ponderosa pine provenances. Forest Science monograph 23, 59 p.

READ, R.A. 1983. Ten-year performance of ponderosa pine provenances in the great plains of North America. USDA Forest Service, Research paper RM-250, 16 p.

REDFIELD, J.A. 1974. Genetics and selection at the Ng locus in blue grouse (Dendragapus obscurus). Heredity 33: 69-78.

REHFELDT, D.G 1974. Genetic variation of Douglas-fir in the northern Rocky Mountains. USDA Forest Service, Res. Note INT-184

REHFELDT, G.E. 1979. Ecological adaptations in Douglas-fir (Pseudotsuga menziesii var. glauca) populations. Heredity 43 (3): 383-397.

REHFELDT, G.E. 1980. Genetic gains from tree improvement of ponderosa pine in southern Idaho. USDA Forest Service, Research paper INT-236, 9 p.

REHFELDT, G.E. 1983a. Adaptations of Pinus contorta populations to heterogeneous environments in northern Idaho. Can J. For. Res. 13: 405-411.

REHFELDT, G.E. 1983b. Ecological adaptations in Douglas-fir (Pseudotsuga menziesii var. glauca) populations. III Central Idaho. Can. J. For. Res. 13: 626-632.

REHFELDT, G.E. 1983c. Seed transfer in the Northern Rockies. In: Progeny testing. Proceedings of servicewide genetic workshop, Charleston, South Carolina December 5-9, 1983: 34-60.

REHFELDT, G.E. 1984. Microevolution of conifers in the northern Rocky Mountains: A view from common gardens. Proceedings of eigth North American forest biology workshop, July 30 - August 1, 1984, Logan, Utah: 132-146.

REHFELDT, G.E. 1986a. Adaptive variation in Pinus ponderosa from intermountain regions. I. Snake and Salmon River Basins. Forest Science 32 (1): 79-92.

REHFELDT, G.E. 1986b. Adaptive variation in Pinus ponderosa from intermountain regions: II. Middle Columbia river system. USDA Forest Service, Research paper INT-373, 9 p.

REHFELDT, G.E. 1989a. Ecological adaptation in douglas-fir (Pseudotsuga menziesii var. glauca): a synthesis. Forest Ecology and Managem. 28: 203-215.

REHFELDT, G.E. 1989b. Genetic variances and covariances in freezing tolerance of lodgepole pine during early winter acclimation. Silvae Genetica 38 (3-4): 133-137.

REHFELDT, G.E. 1990a. Genetic differentiation among populations of Pinus ponderosa from the upper Colorado River Basin. Bot. Gaz. 151 (1): 125-137.

REHFELDT, G.E. 1990b. Adaptability versus zone size: continuous zones for the Rocky Mountains (USA). Joint Meeting of Western Forest Genetics Association and IUFRO working Parties S 2.02-05,06,12,14, Aug. 20 - 24, 1990, Olympia, Washington.

REHFELDT, G.E. 1993. Genetic variation in the Ponderosae of the Southwest. Am. J. Bot. 80: 330-343.

RICHMOND, R.C., et. al. 1980. Esterase-6 and reproduction in Drosophila melanogaster. Science 287: 1483-1485.

RICK, C.M., FOBES, J.F., HOLLE, M. 1977. Genetic variation in Lycopersicon pimpinellifolium: evidence of evolutionary change in mating systems. Plant Syst. Evol. 127: 139-170.

ROPSON, I.J., POWERS, D.A. 1988. A novel dehydrogenase reaction: Mechanisms of hexose-6-phosphate dehydrogenase isolated from the teleost Fundulus heteroclitus. J. Biol. Chem. 263: 97-116.

ROPSON, I.J., POWERS, D.A. 1989. Allelic isozymes of hexose-6-phosphate dehydrogenase from the telesot Fundulus heteroclitus: Physical characteristics and kinetic properties. Mol. Biol. Evol. 6 (2): 171-185.

ROW, K.E, CHING, K.K 1973. Provenance study of Douglas-fir in the Pacific Northwest Region: II. Field performance at age nine. Silvae Genetica 22 (4): 115-119.

ROYALTEY, H.H., ASTRACHAN, E., SOKAL, R.R. 1975. Tests for patterns in geographic variation. Geographical analysis 7: 369-395.

RUDIN, D., ERIKSSON, G., EKBERG, I., RASMUSON, M. 1974. Studies of allele frequencies and inbreeding in Scots pine populations by the aid of the isozyme technique. Silvae Genetica 23: 10-13.

RUETZ, W.F., BERGMANN, F. 1988. Möglichkeiten der Identifizierung autochthoner Hochlagenbestände der Fichte. Forum Genetik-Wald-Forstwirtschaft, 5. Arbeitstagung, Innsbruck, Austria: 173-177.

RUETZ, W., BERGMANN, F. 1989. Möglichkeiten zum Nachweis von autochthonen Hochlagenbeständen der Fichte (Picea abies) in den Berchtesgadener Alpen. Forstw. Cbl. 108: 164-174.

SACKTOR, B. 1975. Biochemistry of insect flight. Part I. Utilization of fuels by muscle. In: Candy, D.J., Kilby, B.A. (eds.): Insect biochemistry and function, Chapman and Hall, 314 p.

SAMOLLOW, P.B., SOULE, M.E. 1983. A case of stress related heterozygote superiority in nature. Evolution 37: 646-649.

SCALTSOYIANNES, A., ROHR, R., PANETSOS, K.P., TSAKTSIRA, M. 1994. Allozyme frequency distributions in five European populations of black pine (Pinus nigra Arnold). Silvae Genetica 43 (1): 20-30.

SCHAAL, B.A., LEVIN, D.A. 1976. The demographic genetics of Litaris cylindrica Michx. (Compositae). Am. Nat. 110: 191-206.

SCHLEGER, W., MAYRHOFER, G., STUR, G. 1978. Relationships between marker heterozygosity and fitness in cattle. Z. Tierz. 94: 296-301.

SCHOLZ, F., BERGMANN, F. 1984. Selection pressure by air pollution as studied by isozyme-gene-systems in Norway spruce exposed to sulphur dioxide. Silvae Genetica 33 (6): 238-241.

SCHWARTZ, D. 1960. Genetic studies on mutant enzymes in maize: synthesis of hybrid enzymes by heterozygotes. Proc. Nat. Acad. Sci U.S.A. 46: 1210-1215.

SCHWARTZ, D. 1973. Single gene heterosis for alcohol dehydrogenase in maize: simple and compound loci. Theor. Appl. Genet. 43: 117-120.

SCHWARTZ, D., LAUGHNER, W.J. 1969. A molecular basis for heterosis. Science 166: 626-627.

SHEA, K.L. 1985. Mating system and population structure in Engelmann spruce and subalpine fir. Ph. D. thesis, University of Colorado, EPOB, Bolder

SHURKHAI, A., PODOGAS, A., ZHIVOTOVSKY, L. 1992. Alloyzme differentiation in the genus Pinus. Silvae Genetica 41 (2): 105-109.

SILEN, R.R., MANDAL, N.L. 1983. Clinal genetic growth variation within two Douglas-fir breeding zones. J. Forestry 81: 216-220.

SIMON, J.P., BERGERON, Y., GOGNON, D. 1986. Isozyme uniformity in populations of red pine (Pinus resinosa) in the Abitibi region, Quebec. Can. J. For. Res. 16: 1133-1135.

SINGH, S.M., ZOUROS, E. 1978. Genetic variation associated with growth rate in the American oyster (Crassostrea virginica). Evolution 32: 342-253.

SMITH, K.G. 1981. Canonical correlation analysis and its use in wildlife habitat studies. In: Capen, D.E. (ed.), 1981: The use of multivariate statistics in studies on wildlife habitat. USDA Forest Service, Gen. Tech. Rep. RM-87: 80-92.

SMITH, M.H., CHESSER, R.K. 1981. Rationale for conserving genetic variation of fish gene pools. Ecol. Bull. Stockholm 34: 13-20.

SMITH, P.J., FRANCIS, R.I.C.C. 1984. Glucosephosphate isomerase genotype frequencies, homozygote excess and size relationships in the sand flounder Rhombosolea plebeia. Mar. Biol. 79: 93-98.

SMITH, R.H. 1977. Monoterpenes of ponderosa pine xylem resin in Western United States. USDA Forest Service, Tech. Bull. 1532, 42 p.

SMITH, R.H. 1981. Variation in immature cone color of ponderosa pine (Pinaceae) in northern California and southern Oregon. Madrono 28: 272-275.

SMOUSE, P., NEEL, J.V. 1977. Multivariate analysis of gametic disequilibrium in the Yanomama. Genetics 85: 733-752.

SMOUSE, P., SPIELMAN, R.S., PARK, M.H. 1982. Multiple-locus allocation of individuals to groups as a function of the genetic variation within and differences among human populations. American Naturalist 119 (4): 445-463.

SNEATH, P.H.A., SOKAL, R.R. 1973. Numerical Taxonomy. The principles and practice of numerical classification. Freeman, San Francisco, 573 p.

SOKAL, R.R. 1978. Population differentiation: something new or more of the same?. In: Brussard, P.F. (ed.), Ecological genetics, the interface, Springer-Verlag, New York: 215-239.

SOKAL, R.R. 1979a. Testing statistical significance of geographic variation patterns. Syst. Zoology 28 (2): 227-232.

SOKAL, R.R. 1979b. Ecological parameters inferred from spatial correlograms. In: Patil, G.P., Rosenzweig, M.L. (eds.), 1979: Contemporary quantitative ecology and related ecometrics, Int. Coop. Publ. House, Fairland, MD: 167-196.

SOKAL, R.R., BIRD, J., RISKA, B. 1980. Geographic variation in Pemphigus populicaulis (Insecta: Aphididae) in Eastern North America. Biol. J. Linn. Soc. 14: 163-200.

SOKAL, R.R., MENOZZI, P. 1982. Spatial autocorrelation of HLA frequencies in Europe support demic diffusion of early farmers. American Naturalist 119 (1): 1-17.

SOKAL, R.R., ODEN, N.L. 1978a. Spatial autocorrelation in biology. 1. Methodology. Biol. J. Linn. Soc. 10: 199-228.

SOKAL, R.R., ODEN, N.L. 1978b. Spatial autocorrelation in biology. 2. Some biological implications and four applications of evolutionary and ecological interest. Biol. J. Linn. Soc. 10: 229-249.

SOKAL, R.R., WARTENBERG, D.E. 1981. Space and population structure. In: Griffith, D.A., McKonnon, R.D. (eds.): Dynamic spatial models. Sijthoff and Noordhoff, Alpen aan den Rijn. The Netherlands.

SOKAL, R.R., WARTENBERG, D.E. 1983. A test of spatial autocorrelation analysis using an isolation-by-distance model. Genetics 105: 219-237.

SOLTIS, D.E., SOLTIS, P.S. 1989. Isozymes in plant biology. Advances in plant sciences series, vol. 4, Timber Press, Portland, Oregon

SORENSEN, F.C. 1979. Provenance variation in Pseudotsuga menziesii seedlings from the var. menziesii - var. glauca transition zone. Silvae Genetica 28 (2-3): 96-103.

SORENSEN, F.C. 1983. Geographic variation in seedling Douglas-fir (Pseudotsuga menziesii) from the western Siskiyou Mountains of Oregon. Ecology 64: 696-702.

SORENSEN, F.C. 1994. Genetic variation and seed transfer guidlines for Ponderosa pine in Central Oregon. USDA Forest Service, Pacific Northwest Research Station, Research Paper PNW-RP-472, 24 p..

SQUILLACE, A.E. 1974. Average genetic correlations among offspring from open-pollinated forest trees. Silvae Genetica 23 (5): 149-156.

SQUILLACE, A.E., SILEN, R.R. 1962. Racial variation in ponderosa pine. Forest Science monograph 2, 27 p.

STALKER, H.D. 1976. Enzymes and reproduction in natural populations of Drosophila euronotus. Genetics 84: 375-384.

STEBBINS, G.L., HARTL, D.L. 1988. Comparative evolution: Latent potential for anagenetic advance. Proc. nat. Acad. Sci. U.S.A. 85: 5141-5145.

STEELE, R. 1987. Ecological relationships of Ponderosa pine. In: Ponderosa pine: The species and its management. Symposium Proceedings Sept. 29 - Oct. 1 1987, Spokane, Washington: 71-76.

STEELE, R., ARNO, S.F., GEIER-HAYES, K. 1986. Wildfire patterns change in central Idaho's ponderosa pine - Douglas-fir forest. West. J. Appl. For. 1: 16-18.

STIMM, B., BERGMANN, F. 1994. Genetische Untersuchungen an Fichtenrotten der subalpinen Waldstufe mit Hilfe von Isoenzympolymorphismen. Schweiz. Z. f. Forstw. 145 (5): 401-411.

STRAUSS, S.H. 1986. Heterosis at allozyme loci under inbreeding and crossbreeding in Pinus attenuata. Genetics 113: 115-134.

STRAUSS, S.H., BOUSQUET, J., HIPKINS, V.D., HONG, Y.P. 1992. Biochemical and molecular genetic markers in biosystematic studies of forest trees. New Forests 6: 125-158.

STRAUSS, S.H., CONKLE, M.T. 1986. Segregation, linkage, and diversity of allozymes in knobcone pine. Theor. Appl. Genet. 72: 483-493.

STRAUSS, S.H., LIBBY, W.J. 1987. Allozyme heterosis in radiata pine is poorly explained by overdominance. American Naturalist 130 (6): 879-890.

STURGEON, K.B. 1979. Monoterpene variation in ponderosa pine xylem resin related to western pine beetle predation. Evolution 33 (3): 803-814.

STUTZ, HP. 1990. Genetische Differenzierung und natürliche Selektionsfaktoren der Fichte (Picea abies (L.) Karst.) in einem zentralalpinen Gebiet der Schweiz. Dissertation Eidgenössische Technische Hochschule, Zürich, No. 9101, 128 p.

STUTZ, H.P., MITTON, J.B. 1988. Genetic variation in Engelmann spruce associated with variation in soil moisture. Arctic and Alpine Research 20 (4): 461-465.

ST. CLAIR, J.B 1989. Effects of seed characteristics and competitive environment on two-year performance of open-pollinated Douglas-fir families. Ph.D. Thesis, Oregon State University, Corvallis, 146 p.

SWOFFORD, D.L., SELANDER, R.B. 1981. BIOSYS-1: A FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. J. Heredity 72: 281-283.

SZMIDT, A.E. 1982. Genetic variation in isolated populations of stone pine (Pinus cembra). Silva Fennica 16: 196-200.

TER BRAAK, C.J.F. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67: 1167-1179.

THIEBAUT. M., LUMARET, R., VERNET, PH. 1982. The bud enzymes of Beech (Fagus sylvatica L.). Genetic distinction and analysis of polymorphisms in several French populations. Silvae Genetica 31 (2-3): 51-60.

THIEBAUT, B., COMPS, B., LEROUX, A. 1992. Relation hauteur-génotype dans une régénéreation naturelle de hêtre (Fagus sylvatica L), équienne et âgée de 18 ans. Ann. sci for 49: 321-335.

TIGERSTEDT, P.M.A. 1973. Studies on isozyme variation in marginal and central populations of Picea abies. Hereditas 75: 47-60.

TIGERSTEDT, P.M.A. 1974. Genetic structure of Picea abies populations as determined by the isozyme approach. Proceedings joint IUFRO meeting S.02.04.1-3, Stockholm: 283-291.

TURESSON, G. 1923. The scope and importance of genecology. Herditas 4: 171-176.

VAN BENEDEN, R.J., WEISGRAM, P.A., ROSENTHAL, E. 1989. Glucosephosphate isomerase allozymes from the teleost Fundulus heteroclitus. Mol. Biol. Evol. 6 (2): 155-170.

VAN DEVENDER, T.R., BETANCOURT, J.L., WIMBERLEY, M. 1984. Biogeographic implications of a packrat midden sequence from the Sacramento Mountains, south-central New Mecixo. Quat. Res. 22: 344-360.

VAN HAVERBEKE, D.F. 1986. Genetic variation in ponderosa pine: A 15-year test of provenances in the great plains. USDA Forest Service, Res. paper RM-265, 16 p.

VAN HOOSER, D.D., KEEGAN, CH.E. 1987. Distribution and volumes of ponderosa pine forests. In: Ponderosa pine: The species and its management. Symposium Proceedings Sept. 29 - Oct. 1 1987, Spokane, Washington: 1-6.

VIGUE, C.L., WEISGRAM, P.A., ROSENTHAL, E. 1982. Selection at the alcohol dehydrogenase locus of Drosophila melanogaster: effects of ethanol and temperature. Biochem. Genet. 20: 681-688.

VILLANI, F., BENDEDETTELINI, S., CHERUBINI, M. 1991. Genetic differentiation among Turkish chestnut (Castanea sativa Mill.) populations. Heredity 66: 131-136.

VON WUEHLISCH, G., KRUSCHE, D. 1991. Single and multilocus genetic effects on diameter growth in Picea Abies (Karst.). In: Fineschi, S. et al.: Biochemical markers in the population genetics of forest trees. Academic Publ., The Hague: 77-86.

WALKER, R.B. 1954. The ecology of serpentine soils. II. Factors affecting plant growth on serpentine soils. Ecology 35: 259-266.

WARTENBERG, D. 1985. Canonical trend surface analysis: A method for describing geographic patterns. Syst. Zool. 34 (3): 259-279.

WATT, W.B. 1977. Adaptation at specific loci.I. Natural selection on phosphoglucose isomerase of Colias butterflies: Biochemical and population aspects. Genetics 87: 177-194.

WATT, W.B., CASSIN, R.C., SWAN, M.S. 1983. Adaptation at specific loci. III. Field behavior and survivorship differences among Colias PGI genotypes are predictable from in vitro biochemistry. Genetics 103: 725-739.

WATT, W., CARTER, P.A., BLOWER, S.M 1985. Adaptation at specific loci. IV. Differential mating success among glycolytic allozyme genotypes of Colias butterflies. Genetics 109: 157-175.

WEBER, J.C. 1988. Geographic variation in central Oregon ponderosa pine (Pinus Ponderosa Laws.): Seed germination; seed, wing and cone morphology; seed color. PhD Thesis, Oregon State University, Corvallis.

WEBER, J., STETTLER, R. 1980. Isoenzyme variation among ten populations of Populus trichocarpa Torr. et Grey in the Pacific Northwest. Silvae Genetica 30: 82-87.

WEIDMAN, R.H. 1939. Evidence of racial influence in a 25-year old test of ponderosa pine. J. Agric. Res. 59 (12): 855-887.

WELLS, O.O. 1964a. Geographic variation in ponderosa pine. I. The ecotypes and their distribution. Silvae Genetica 13 (4): 89-103.

WELLS, O.O. 1964b. Geographic variation in ponderosa pine. II. Correlation between progeny performance and characteristics of the native habitat. Silvae Genetica 13 (5): 125-132.

WESTFALL, R.D. 1992. Developing seed transfer zones. In: Fins, L. et al. (eds.): Handbook of quantitative forest genetics. Kluwer Academic Publishers, Dordrecht, Netherlands: 313-398.

WESTFALL, R.D., CONKLE, M.T. 1992. Allozyme markers in breeding zone designation. New Forests 6: 279-309.

WHEELER, N.C., GURIES, R.P. 1982. Population structure, genetic diversity, and morphological variation in Pinus contorta Dougl. Can. J. For. Res 12: 595-606.

WHEELER, N.C., GURIES, R.P. 1982. Biogeography of lodgpole pine. Can. J. Bot. 60: 1805-1814.

WHEELER, N.C., GURIES, R.P., O'MALLEY, D.M. 1983. Biosystematics of the Genus Pinus, subsection Contortae. Biochem. Syst. Ecol. 11: 333-340.

WHITE, T.L. 1981. Genecology of Douglas-fir from southwestern Oregon. Ph.D. thesis, Oregon State University, Corvallis, Oregon

WILLIAMS, B.K. 1981. Discriminant analysis in wildlife research: Theory and applications. In: Capen, D.E. (ed.): The use of multivariate statistics in studies on wildlife habitat. U.S. Forest Service, Gen. Tech. Rep. RM-87: 59-71.

WILLIAMS, B.K., TITUS, K. 1988. Assessment of sampling stability in ecological applications of discriminant analysis. Ecology 69: 1172-1174.

WILLIAMS, J.G., et al.. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl. Acid Res. 18: 6531-6535.

WOODS, J.H., BLAKE, G.M., ALLENDORF, F.W. 1983. Amount and distribution of isozyme variation in ponderosa pine from eastern Montana. Silvae Genetica 32: 151-156.

WOODS, J.H., BLAKE, G.M., ALLENDORF, F.W. 1983. Amount and distribution of isozyme variation in ponderosa pine from eastern Montana. Silvae Genatica 32 (5-6): 151-157.

WORKMAN, P.L., NISWANDER, J.D. 1970. Population studies on southwestern Indian tribes. II. Local genetic differentiation in the Papago. Am. J. Human Genet. 22: 24-49.

WRIGHT, J.W. 1976. Introduction to Forest Genetics. Academic Press, New York, 463 p.

WRIGHT, S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19: 395-420.

WRIGHT, S. 1978. Evolution and the genetics of populations variability within and among natural populations. Vol. 4, Chicago 1978, 580 p.

WRIGLEY, N. 1985. Categorical data analysis for geographers and environmental scientists. Longman, London and New York, 392 p.

XIE, Y.CH., KNOWLES, P. 1992. Associations between allozyme phenotypes and soil nutrients in a natural population of jack pine (Pinus banksiana). Biochemical Systematics and Ecology 20 (2): 179-185.

YEH, F.CH., O'MALLEY, D. 1980. Enzyme variations in natural populations of Douglas-fir, Pseudotsuga menziesii (Mirb.) Franco, from British Columbia. 1. Genetic variation patterns in costal populations. Silvae Genetica 29 (3-4): 83-92.

YEH, F.C. 1981. Analyses of gene diversity in some species of conifers. USDA Forest Service, Gen. Tech. Rep. PSW-48: 48-52.

YEH, F.C., CHELIAK, W.M., DANCIK, B.P. 1985. Population differentiation in lodgepole pine, Pinus contorta spp. latifolia: a discriminant analysis of allozyme variation. Can. J. Genet. Cytol. 27: 210-218.

YEH, F.C., EL-KASSABY, Y.A. 1980. Enzyme variation in natural populations of sitka spruce (Picea sitchensis). 1. Genetic variation patterns among trees from 10 IUFRO provenances. Can. J. For. Res. 10: 415-422.

YEH, F.C., KHALIL, M.A., EL-KASSABY, Y.A., TRUST, D.C. 1986. Allozyme variation in Picea mariana from Newfoundland: genetic diversity, population structure, and analysis of differentiation. Can. J. For. Res. 16: 713-720.

YEH, F.C., LAYTON, C. 1979. The organisation of genetic variability in central and marginal populations of lodgpole pine Pinus contorta spp. latifolia. Can. J. Genet. Cytol. 21: 487-503.

YOUNGBLOOD, A.P., MAUK, R.L. 1985. Coniferous forest habitat types of central and southern Utah. USDA Forest Service, Gen. Tech. Rep. INT-187, 89 p.

ZAMER, W.E., HOFFMAN, R.J. 1989. Allozymes of glucose-6-phosphate isomerase differentially modulate pentose-shunt metabolism in the sea anemone Metritium senile. Proc. Nat. Acad. Sci. U.S.A. 86: 2737-2741.

ZANETTO, A., KREMER, A., LABBE, T. 1993. Difference of genetic variation based on isozymes of primary and secondary metabolism in Quercus petraea. Ann. Sci. For. 50, Suppl. 1: 245-252.

ZOBEL, B., TALBERT, J. 1984. Applied tree improvement. John Wiley and Sons, New York, 505 p.

ZOUROS, E., FOLTZ, D.W. 1987. The use of allelic isozyme variation for the study of heterosis. Isozymes, current topics in biological and medical research 13: 1-59.

ZOUROS, E., SINGH, S.M., FOLTZ, D.W., MALLET, A.L. 1983. Post-settlement viability in the American oyster (Crassostera virginica): An overdominant phenotype. Genet. Res. 41: 259-270.

ZOUROS, E., SINGH, S.M., MILES, H.E. 1980. Growth rate in oyster: An overdominant phenotype and its possible explanations. Evolution 34: 856-867.

Appendix I: Buffers and stains used for electrophoresis of Ponderosa pine

Buffers:

Gel buffers (and tray buffers for systems D and E):

Tris citrate pH 8.3:	62.0 g	Trizma base (Sigma corp.) 7-9
System A	16.0 g	Citric acid (anhydrous)
	10.0 l	dH_2O
Tris citrate pH 8.8:	121.1 g	Trizma base 7-9
System B	14.5 g	Citric acid (anhydrous)
	10.0 l	dH_2O
Morpholine citrate pH 6.1:	76.85 g	Citric acid (anhydrous)
System D	10.00 l	dH_2O titrated to pH 6.1 with 1 M (3-aminopropyl)morpholine
Morpholine citrate pH 8.1:	77.0 g	Citric acid (anhydrous)
System E	10.0 l	dH_2O titrated to pH 8.1 with 1 M (3-aminopropyl)morpholine

Tray buffers:

Lithium borate pH 8.3:	21.0 g	Lithium hydroxide
System A	118.9 g	Boric acid
	10.0 l	dH_2O
Sodium borate pH 8.0:	20.0 g	Sodium borate
System B	185.5 g	Boric acid
	10.0 l	dH_2O

Extraction buffer:

10 ml	0.2 M Phosphate buffer pH 7.5 (see stain buffers)
10 ml	5 % succrose solution
100 mg	Ascorbic acid
50 mg	D-glucose-6-phosphate
50 mg	Bovine albumin
10 mg	Dithiothreitol

Stain buffers:

1 M Tris HCL buffer:	74.0 g	Trizma base 7-9
pH 8.0	61.4 g	Trizma hydrochloride
	1.0 l	Distilled water
1 M Tris HCL buffer:	16.0 g	Trizma base 7-9
pH 7.0	137.4 g	Trizma hydrochloride
	1.0 l	dH_2O
ACP buffer:	1.21 g	Sodium acetate, trihydrate
pH 4.0	2.35 g	Acetic acid, glacial
	2.50 ml	Magnesium chloride
	0.50 l	dH_2O titrated to pH 4.0 with 1 M natrium hydroxid solution

Appendix I: Buffers and stains used for electrophoresis of Ponderosa pine

PHOS buffer:	3.84 g	Sodium phosphate, monobasic
pH 7.5	23.86 g	Sodium phosphate, dibasic
	1.00 l	dH_2O
Tris Malate buffer:	24.2 g	Trizma base 7-9
pH 3.8	23.2 g	Maleic acid
		filled to 100 ml with dH_2O

Stock solutions for stain components:

0.2 M L-Aspartic acid solution (pH 7.5):	5.3 g	Aspartic acid
	200 ml	dH_2O
		titrated to pH 7.5 with 2 M KOH
3 % Bovin albumin solution:	3.0 g	Bovine albumin
	100 ml	dH_2O
0.1 M a-ketoglutarate solution:	3.0 g	a-Ketoglutaric acid
	200 ml	dH_2O
0.2 M Citric acid solution:	38.4 g	Citric acid
	1.0 l	dH_2O
0.25 M Fumaric acid solution:	40.0 g	Fumaric acid
	1.0 l	dH_2O
D - Glucose - 1,6 - diphosphate solution:	5.0 mg	D-Glucose-1,6-diphosphate
	50 ml	dH_2O
G6PDH solution:	1.0 g	Bovine albumine
	100 ml	PHOS buffer
	900 ml	dH_2O
	4000 units	Glucose-6-phosphate dehydro-genase
GOT solution:	42.5 ml	0.2 M Aspartic acid solution
	10.0 ml	3 % Bovin albumin solution
	12.5 ml	0.1 M a-ketoglutarate solution
	5.0 ml	0.5 % Pyridoxal-5-phosphate
2 M KOH:	28.0 g	KOH
	250 ml	dH_2O
4 M NaOH:	40 g	NaOH
	250 ml	dH_2O
1 M NaOH:	10 g	NaOH
	250 ml	dH_2O
1 M (3-aminopropyl) morpholine:	28.48 g	(3-aminopropyl)morpholine
	200 ml	dH_2O
1 % $MgCl_2$ solution:	2.5 g	$MgCl_2$
	250 ml	dH_2O
1 M $MgCl_2$ solution:	23.8 g	$MgCl_2$
	250 ml	dH_2O

Appendix I: Buffers and stains used for electrophoresis of Ponderosa pine

Malic acid solution (pH 7.0):	67.5 g	Malic acid
	40.0 g	NaOH
	500 ml	dH2O
		titrated to pH 7.0 with 4 M NaOH
NADP or NAD solution:	1 g	B-Nicotamidee adenine nucleotide (NADH) or dinucleotide (NAD)
	1 g	Nitro blue tetrazolium
	500 mg	Phenazine methosulfat
	250 ml	dH_2O
PGI solution:	250 mg	Phosphoglucoseisomerase (5 units/mg)
	62.5 ml	dH_2O
0.5 % Pyridoxal-5-phosphate:	100 mg	Pyridoxal-5-phosphate
	20 ml	dH_2O
Cis-aconitic acid solution:	1.5 g	Cis-aconitic acid
	20 ml	dH_2O
		titrated to pH 7.0 with 1 M NaOH
MTT solution:	5 mg	(3-[4,5-dimethylthiazole-2-yl]-2,5-diphenyltetrazolium bromide)
	1 ml	dH_2O
PMS solution:	3 mg	Phenazine methosulfate
	1 ml	dH_2O

Fixation of gels:

Fixation solution:	5 parts	Ethanol
	5 parts	dH_2O
	1 part	Acetic acid glacial

Seed stratification:

1 % Hydrogen peroxide:	10 ml	H_2O_2 (30%)
	290 ml	dH_2O

Stain recipies:

The quantities for the following stain recipies are all given for the staining of 4 slices

Appendix I: Buffers and stains used for electrophoresis of Ponderosa pine

MNR:	**Menadione reductase:**
E.C. Number:	1.6.99.2
Buffer System:	A (E)
Inheritance:	Strauss und Conkle, 1986
Stain recipie:	
Buffer:	20 ml 1 M Tris HCl pH 7.0 80 ml dH_2O
Components:	50 mg NADH 50 mg Menadione 2 mg Nitro blue tetrazolium
Procedure:	Add components to warm

PEP:	**Peptidase**
E.C. Number:	3.4.13.1
Buffer System:	A
Inheritance:	Strauss und Conkle, 1986 Niebling Conkle, 1990 Conkle, 1981 O'Maley et al. , 1979
Stain recipie:	
Buffer:	20 ml 1 M Tris HCl pH 7.0 80 ml dH_2O
Components:	80 mg Glycyl-L-leucine 80 mg Leucyl-L-alanine 60 mg Leucyl-L-tryosine 20 mg Valyl-L-tyrosine 20 mg Peroxidase 20 mg snake venome 100 mg 3-amino-9-ethyl carbazole 2 ml N,N-dimethyl formamide
Procedure:	Combine first six with buffer, add amino-carbazole dissolved in dymethyl formamide just before staining

MPI:	**Mannose phosphate isomerase**
E.C. Number:	5.3.1.8
Buffer System:	A
Inheritance:	Strauss and Conkle, 1986 Furnier and Adams, 1986 O'Malley et. al., 1979 Niebling and Conkle, 1990
Stain recipie:	
Buffer:	20 ml 1 M Tris HCl pH 8.0 20 ml dH_2O
Components:	70 ml G6PDH solution 50 ml PGI solution 2 ml 1% $MgCl_2$ solution 4 ml NADP solution 100 mg Mannose-6-phosphate
Procedure:	add components to warm buffer before staining

LAP:	**Leucine aminopeptidase**
E.C. Number:	3.4.11.1
Buffer System:	A
Inheritance:	Strauss and Conkle, 1986 Furnier and Adams, 1986 Niebling and Conkle, 1990 Conkle, 1981
Stain recipie:	
Buffer:	66 ml Tris malate buffer pH 3.8 6.6 ml 1 M NaOH 27.4 ml dH_2O
Components:	34 ml dH_2O 26 mg L-leucine-B-naphthyl-amide 26 mg Black K salt
Procedure:	dissolve components in water staining

Appendix I: Buffers and stains used for electrophoresis of Ponderosa pine

GOT:	**Glutamate oxaloacetate transaminase**
E.C. Number:	2.6.1.1
Buffer System:	B
Inheritance:	Strauss and Conkle, 1986 Conkle, 1981 Niebling and Conkle, 1990 Furnier and Adams, 1986
Stain recipie:	
Buffer:	50ml PHOS buffer pH 7.5
Components:	22.4 ml GOT solution 100 mg Fast Blue BB Salt
Procedure:	add components just before staining

ACP:	**Acid phosphatase**
E.C. Number	3.1.3.2
Buffer System:	B
Inheritance:	Strauss and Conkle, 1986 Conkle, 1981
Stain recipie:	
Buffer:	100 ml ACP buffer pH 4.0
Components:	100 mg a-naphthyl acid phosphate 100 mg Black K Salt 2 ml 1% $MgCl_2$ solution
Procedure:	add components to buffer, mix well before staining

G6P:	**Glucose-6-phosphate dehydrogenase**
E.C. Number:	1.1.1.49
Buffer System:	B
Inheritance:	Strauss and Conkle, 1986 Niebling and Conkle, 1990 Conkle, 1981 O'Malley et al., 1979 Furnier and Adams, 1986
Stain recipie:	
Buffer:	20 ml 1 M Tris HCl pH 8.0 80 ml dH_2O
Components:	400 mg D-glucose-6-phosphate 2 ml 3% Bovine albumin solution 4 ml NADP solution 2 ml 1% $MgCl_2$ solution 2 ml PMS solution 2 ml MTT solution
Procedure:	add components to warm buffer before staining

GDH:	**Glutamate dehydrogenase**
E.C. Number:	1.4.1.3
Buffer System:	B
Inheritance:	Strauss und Conkle, 1986 Conkle, 1981 Furnier and Adams, 1986 Niebling and Conkle, 1990
Stain recipie:	
Buffer:	7. 5 ml 1 M Tris HCl pH 8.0 140 ml dH_2O
Components:	3 ml 1 M NaOH 4 g L-glutamic acid 5 ml NAD solution 100 mg CaCl2 2 ml PMS solution 2 ml MTT solution
Procedure:	add components to warm buffer before staining

Appendix I: Buffers and stains used for electrophoresis of Ponderosa pine

IDH:	**Isocitrate dehydrogenase**
E.C. Number:	1.1.1.42
Buffer System:	D
Inheritance:	Strauss and Conkle, 1986 Conkle, 1981 Niebling and Conkle, 1990 O'Malley et al., 1979 Furnier and Adams, 1986
Stain recipie:	
Buffer:	20 ml 1 M Tris HCl pH 8.0 80 ml dH_2O
Components:	200 mg Isocitric acid 2 ml NADP solution 2 ml 1 % $MgCl_2$ solution
Procedure:	add components to warm buffer

ACO:	**Aconitase**
E.C. Number:	4.2.1.3
Buffer System:	D
Inheritance:	Strauss and Conkle,1986 Conkle, 1981 Niebling and Conkle, 1990 Furnier and Adams, 1986
Stain recipie:	
Buffer:	20 ml 1M tris HCl PH 8.0 80 ml dH_2O
Components:	4 ml Cis-aconitic acid pH 7.0 2 ml Isocitrate dehydrogenase 2 ml 1 % $MgCl_2$ solution 6 ml NADP solution
Pocedure:	add components to warm buffer

SKD:	**Shikimate dehydrogenase**
E.C. Number:	1.1.1.25
Buffer System:	D
Inheritance:	Strauss and Conkle, 1986 Niebling and Conkle, 1990
Stain recipie:	
Buffer:	10 ml 1 M Tris HCl pH 8.0 4 ml 1 M NaOH 100 ml dH_2O
Components:	200 mg Shikimic acid 20 mg NADP 4 ml MTT solution 2 ml PMS solution
Procedure:	add all components to warm buffer

PGM:	**Phosphoglucomutase**
E.C. Number:	2.7.5.1
Buffer System:	D
Inheritance:	Strauss and Conkle, 1986 Conkle, 1981 Niebling and Conkle, 1990 O'Malley et al., 1979 Furnier and Adams, 1986
Stain recipie:	
Buffer:	20 ml 1M Tris HCl pH 8.0 80 ml dH_2O
Components:	20 ml G6PDH solution 2 ml 1% $MgCl_2$ solution 4 ml NADP solution 2 ml 3% Bovine albumin solution 4 ml D-Glucose-1,6-diphophate solution 100 mg a-D-glucose-1-phosphate
Procedure:	add all components to warm buffer

Appendix I: Buffers and stains used for electrophoresis of Ponderosa pine

PGI:	**Phosphoglucose isomerase**
E.C. Number:	5.3.1.9
Buffer System:	E
Inheritance:	Strauss and Conkle, 1986 Conkle, 1981 Niebling and Conkle, 1990 O'Malley et al., 1979 Furnier and Adams, 1986
Stain recipie:	
Buffer:	10 ml 1 M Tris HCl pH 8.0 120 ml dH_2O
Components:	10 ml G6PDH solution 2 ml 1% $MgCl_2$ solution 2 ml NADP solution 100 mg Fructose-6-phosphate
Procedure:	mix components before staining develop at room temperature watch for over-staining

ADH:	**Alcohol dehydrogenase**
E.C. Number:	1.1.1.1
Buffer System:	E
Inheritance:	Strauss and Conkle, 1986 Conkle, 1981 Niebling and Conkle, 1990 O'Malley et al., 1979 Furnier and Adams, 1986
Stain recipie:	
Buffer:	20 ml 1 M Tris HCl pH 8.0 80 ml dH_2O
Components:	2 ml 95 % ethylalcohol (ethanol) 2 ml NAD solution
Procedure:	add components to warm buffer

MDH:	**Malate dehydrogenase**
E.C. Number:	1.1.1.37
Buffer System:	E
Inheritance:	Strauss and Conkle, 1986 Conkle, 1981 Niebling and Conkle, 1990 O'Malley et al.,1979 Furnier and Adams, 1986
Stain recipie:	
Buffer:	20 ml 1 M Tris HCl pH 8.0 80 ml dH_2O
Components:	20 ml malic acid solution 4 ml NAD solution 2 ml MTT solution 1 ml PMS solution
Procedure:	add components to warm buffer

FDP:	**Fructose diphosphatase**
E.C. Number:	3.1.3.11
Buffer System:	E
Inheritance:	Strauss and Conkle, 1986 Niebling and Conkle, 1990
Stain recipie:	
Buffer:	20 ml 1 M tris HCl pH 8.0 20 ml dH_2O
Components:	30ml G6PDH solution 10ml PGI solution 2 ml 1% $MgCl_2$ solution 2 ml NADP solution 200 mg Fructose-1,6-diphosphate
Procedure:	add components to warm buffer

Appendix I: Buffers and stains used for electrophoresis of Ponderosa pine

UGP:	UDP-glucose pyrophospho-rylase	FUM:	Fumarase
E.C. Number:	2.7.7.9	E.C. Number:	4.2.1.2
Buffer System:	E	Buffer System:	E
Inheritance:	Strauss and Conkle, 1986	Inheritance:	Strauss and Conkle, 1986 Niebling and Conkle, 1990
Stain recipie:		Stain recipie:	
Buffer:	20 ml 1 M Tris HCl pH 8.0 80 ml dH_2O	Buffer:	20 ml PHOS buffer pH. 7.5 80 ml dH_2O
Components:	2 ml Glucose-1,6-diphosphate solution 30 ml G6PDH solution 2 ml 1 % $MgCl_2$ solution 5 ml NADP solution 6 drops Phosphoglucomutase 40 mg D-Pyrophosphate 120 mg Uridine-5-di-phospho-glucose	Components:	60 mg NAD 400 mg Fumaric acid 2 ml MTT solution 2 ml PMS solution 10 drops Malic dehydro- 2 ml NAD solution
Procedure:	add components to warm buffer	Procedure:	add components to warm buffer

Acknowledgments

This study would not have been possible without the help and support of several persons to whom I am deeply grateful.

First of all I thank professor J.-Ph. Schütz who enthusiastically supported my interests in forest genetics and gene-ecology. Thanks to him I had the possibility to spend a year in Corvallis, where I did the data collection and followed courses in genetics and statistics. Moreover, he agreed to act as examiner for this thesis. Professor W.T. Adams, Oregon State University, Corvallis was the driving force who made this study possible. For his support and friendship, his cooperation to use his laboratory, his many helpful suggestions and his willingness to act as co-examiner I am very grateful.

Several other persons contributed to the successful completion of this work. R. K. Campbell and F. Sorensen, Pacific Northwest Research Station, Corvallis had the idea for this study. They provided the seed for the allozyme analysis, all the location data and kindly accepted to let me use the data of the common garden studies. Their enthusiasm and support was essential for this work. I also thank M. T. Conkle and P. Hodgskiss, Pacific Southwest Forest and Range Experiment Station, Berkeley, for their advise regarding electrophoretic procedures and for the use of their maps of zymograms for ponderosa pine. D. Delany, Pacific Southwest Forest and Range Experiment Station, Berkeley, shared all her knowledge and experience regarding electrophoresis and scoring of zymograms with me. I also thank R.D. Westfall, Pacific Southwest Forest and Range Experiment Station, Berkeley, for his advise in statistics. His assistance and numerous suggestions regarding the analyses of the complex data structure were deeply appreciated. For several month, Ch. Spriggs, graduate student, helped me in the laboratory. He became an expert in pouring and slicing gels. Thanks to his ideas, many improvments in the procedures could be realized. I also appreciated the helpful hints and cooperation of V. Hipkins who shared the laboratory with me. G. Lienkaemper, Oregon State University, Corvallis, helped me digitize the seed zone map, transform coordinates, assign trees to seed zones and provided a lot of assistance using GIS. I also thank M. Irmay, Swiss Federal Institute of Technology, Zürich, who helped me produce several overlay maps with the help of GIS.

Curriculum

Born in Kreuzlingen, Switzerland on September 1, 1953 as the second son of Paul and Meta Rotach-König from Berg, TG.

1960 - 1966:	Primary School in Berg, TG, Switzerland
1966 - 1969:	Secondary School in Berg, TG
1969 - 1973:	Gymnasium in Frauenfeld, TG - Matura degree
1973 - 1974:	Military Service and assignments as a teacher in different secondary schools
1974 - 1979:	Study of Forestry Sciences at the Swiss Federal Institute of Technology (ETH), Zürich - Diploma degree (dipl. Forsting. ETH)
1977 - 1978:	Trainee year in two different forest districts in Switzerland (Ins, BE and Locarno, TI) - Federal certificate of eligibility for appointments in Federal or Cantonal Forest Services
1979 - 1986:	Teaching assistant at the Chair of Silviculture, Swiss Federal Institute of Technology (ETH) Zürich
1986 - 1990:	Senior assistant at the Chair of Silviculture, Swiss Federal Institute of Technology (ETH) Zürich
1990 - present:	Scientific collaborator at the Chair of Silviculture, Swiss Federal Institute of Technology (ETH) Zürich, teaching - and research assistant for silviculture
1991	Visisting scientist at Oregon State University, Department of Forest Science, in Corvallis, Oregon, U.S.A. Attending courses in genetics, forest genetics and breeding, statistics and research methods. Start of thesis with an introduction to electrophoresis and isozyme techniques and the completion of the isozyme analysis of ponderosa pine
1992 - present:	Teaching a course in forest genetics, breeding and reproduction material;different scientific projects, completion of thesis