AN ABSTRACT OF THE THESIS OF

Valerie D. Hipkins for the degree of Master of Science

in Forest Science presented on October 28, 1988

Title: Biochemical Markers of Quality in Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco] Seedlings: The Relationship Between Survival, Root Growth Potential, Freezing and Drying Damage, and Time of Lift

Abstract approved: _____
Joe B. Zaerr

The physiological condition of a seedling often determines its ability to survive when planted. However, physiological damage is difficult to determine from external measurement or observation, because a healthy looking seedling can be of poor quality (quality is defined as the ability of a seedling to survive when outplanted). Methods currently used to assess quality rely primarily on the growth performance of a seedling. The most widely used quality evaluation method tests a seedlings ability to initiate and elongate roots, commonly called root growth potential (RGP). Other evaluation methods measure seedling frost hardiness, plant water potential, and speed of bud break. However, these indirect measurements of seedling physiology are often inaccurate evaluators of quality and are slow to yield results. To obtain a direct measure of the physiological condition of seedlings, and hence a direct measure of seedling quality, biochemical markers were sought for Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco] plants of varying vigor.

Markers of field survival, growth room survival, RGP (number and total length of new roots), exposure to freezing and drying conditions, and time of lifting were identified by measuring a large number of compounds from methanol extracts of the apical shoots of 300 seedlings using high performance liquid chromatography (HPLC). Data were acquired and processed on a computer, and analyzed with five different multivariate statistical techniques. Principal component analysis was used as an exploratory technique to investigate the structure of the data. Stepwise multiple regression and cluster analysis were used to explore grouping among the treatments and to identify markers. The nature of many of the resulting clusters was not easily determined using this method. Unknown factors in addition to survival, RGP, and exposure to freezing and drying conditions, appeared to control treatment similarity. However, it was possible to cluster treatments accurately based on the time seedlings were lifted from the nursery.

Results from stepwise and canonical discriminant analysis imply that markers identified with these procedures are capable of separating treatments of varying field survival, growth room survival, RGP, or time of lift. Markers were unable to confirm a clear separation between seedlings that had been damaged by exposure to freezing or drying conditions. A strong relationship was observed between survival and RGP as indicated by shared markers.

Biochemical markers can be used to predict the survival of seedlings by classifying plants of unknown survival into predetermined groups. This technique promises to be a rapid, reliable, and quantitative means of evaluating seedling quality.

Biochemical Markers of Quality in Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco] Seedlings: The Relationship Between Survival, Root Growth Potential, Freezing and Drying Damage, and Time of Lift

by

Valerie D. Hipkins

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Completed October 28, 1988

Commencement June 1989

APPROVED:

Professor of Tree Physiology in charge of major

Head of Department of Forest Science

Dean of Graduate School

Date thesis is presented October 28, 1988

ACKNOWLEDGEMENT

I would like to gratefully acknowledge the Southwest Oregon Forestry Intensified Research Program for their funding that made this research possible.

Many people have provided me with invaluable statistical and computer advice and assistance. I wish to thank: Tom Bell for writing programs to manipulate my data in Symphony and transforming it into a SAS data set; Tom Sabin and Dr. Susan Stafford, my much needed and appreciated statistical advisors; Gody Spycher who managed to get my large data set onto the mainframe SAS system, solve many of my computer problems, and continually provide me with additional computer space when I needed it; and Dr. Joe Zaerr and Jon Zaerr who spent many hours writing the BASIC programs that were necessary for accurate identification of the HPLC peaks.

I would also like to thank Joe Zaerr for the support he has given me since I arrived at Oregon State. Thanks also to Keith Forry for all of his help during the early part of my work and Nan Vance for the advice and encouragement.

Most of all, a heartfelt thank you to my husband, John Curless. Not only did he help me kill and grind up the buds of 300+ seedlings and take field bud status measurements in the scorching summer sun, but he also was always there with constant love, support, and encouragement when the going got tough and I no longer wanted to get going.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES
APPENDIX D

Biochemical Markers of Quality in Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco] Seedlings: The Relationship Between Survival, Root Growth Potential, Freezing and Drying Damage, and Time of Lift

INTRODUCTION

A crucial need to reforest cut-over lands in the Pacific Northwest was first realized in the early 1900's. Yet it was not until the 1972 Oregon Forest Practices Act that reforestation became a legal obligation in this state (Cleary et al. 1982). Regeneration is now a critical and often costly step in the forestry cycle. The annual production of seedlings from American and Canadian nurseries was estimated in 1972 to exceed one billion (Hermann et al. 1972). In the Pacific Northwest alone, bareroot nurseries were estimated to produce approximately 278 million seedlings annually (Duryea and Landis 1984).

Although no accurate survival data exist, plantation failures and poor seedling performance in the field are common occurrences. Failures can be attributed to competing vegetation, animal browse, and insect or disease damage. However, there are many instances when failures are related to the physiological condition of nursery stock at the time of planting. Seedlings can become damaged while still in nursery beds (such as frost or insect damage) and during the lifting, handling, storage, or planting processes. Tissues can become damaged when exposed to extremely high or low temperatures such as those encountered by equipment failure during cold storage. Physical injury, especially to the root system, may result from rough handling or lifting from heavy, wet soils (McCreary and Zaerr 1987).

Because the need to reforest harvested lands rapidly represents such a large investment, there exists a need to plant only seedlings with a high potential for survival and growth in the field. In the Pacific Northwest the need to evaluate a seedlings quality

before planting is especially critical because of the narrow midwinter 'lifting window' and the need for extensive cold storage to accommodate late-spring planting at high elevations (Ritchie 1982).

While many definitions have been applied to seedling quality, it is perhaps best defined in terms of reforestation success. Seedling quality is dependent on both physiological and morphological characteristics, yet the physiological condition of a seedling, perhaps more than morphological factors, determines its ability to survive when planted (Duryea and McClain 1984). However, physiological damage is difficult to determine from external measurement or observation since a healthy looking seedling can be of poor quality.

Although there is a greatly expanded level of knowledge of the physiological requirements of seedlings, not all factors and conditions governing subsequent seedling survival and growth are understood. Seedlings are subject to a large array of stresses that do not manifest themselves in a visible manner and, therefore, remain undetected. For example, desiccation of roots during lifting can cause serious damage. Such seedlings may appear healthy, yet fail to survive when planted in the field.

Improvements in cultural practices and a better understanding of seedling physiological requirements have made it possible for nurseries to produce stock of higher quality than in the past. Nursery managers can manipulate various exogenous factors to control seedling growth and development. Nutrients, moisture, density, and temperature can interact with plant physiological factors such as hormone levels and carbohydrate reserves to yield seedlings of maximum quality (Lavender 1984). However, because seedling physiology changes seasonally through dormancy, a seedlings physiological cycle must be in phase with environmental signals for high quality stock to result (Lavender 1985).

Forest nurseries may, therefore, evaluate seedlings to assess their cultural practices and improve stock quality. Test results may also provide a basis for determining when to lift and store nursery stock and for culling poor quality seedling lots. On the other hand, foresters and landowners may evaluate seedling quality to insure that only high quality planting stock is used on their lands, to correct storage and handling practices, and to match seedlings to specific sites.

Seedling quality evaluation methods have been organized by Ritchie (1984) into two classes-- material and performance. Material attributes are those that are directly measurable such as morphology and carbohydrate status. Measurements such as root growth, frost hardiness, and vigor characterize the performance of the entire seedling and are, therefore, performance attributes.

The most widely used and studied evaluation techniques are root growth potential, frost hardiness, and morphology (Duryea 1985). Even though these and other methods have received a great deal of attention and study, there still does not exist a procedure to predict accurately how well a seedling will survive and grow when planted in the field. The demand today is not only for higher quality stock but also for improved quality evaluation methods. What foresters, landowners, and nursery managers have in common is the need for a quality evaluation method that rapidly yields results and is inexpensive, reliable, and quantitative.

The primary objective of this thesis research was to identify biochemical markers of seedling quality (defined as the ability of a seedling to survive when planted). A secondary objective was to identify markers of factors related to seedling survival. Markers related to field survival, growth room survival, freezing and drying damage, root growth potential (number and total length of new roots), and date of lifting were identified, though not chemically characterized. Because it is thought that RGP measurements are directly related to seedling survival, markers were identified for

number of new roots and total length of new roots. Seedling damage due to exposure to freezing and drying conditions is an important forestry problem in the Pacific Northwest. Markers were identified in order to predict damage of this kind. The time that seedlings are lifted plays an important role in their subsequent field survival. One treatment of December-lifted seedlings was included in this study by request of the nursury providing the seedlings. Douglas-fir seedlings are purchased by foresters and landowners in January, which puts a large demand on the nursery over a short period of time. Markers for time of lift were sought that would determine if seedlings such as these December-lifted seedlings were of equal quality as the January-lifted seedlings. It was our goal to study the importance that various compounds had in determining seedling quality and shed light on the biochemical interrelationship between survival and related processes such as date of lifting, freezing and drying damage, and root growth potential.

In the future, identified markers can be used in a predictive capacity to actually classify individual seedlings with respect to survival. This approach has the potential to provide a rapid, reliable, and quantitative means of predicting seedling quality as measured by growth room or field survival.

LITERATURE REVIEW

Successful establishment of plantations depends on the availability of high quality forest tree seedlings. The ability to plant only seedlings of the highest quality reduces plantation loss and the risk of regeneration delays brought about by the need to replant or interplant. For over 60 years the importance of determining superior seedlings has been realized (Paton 1929). Various attempts have been made to evaluate seedlings using morphological and, more recently, physiological characteristics (Duryea 1985; Sutton 1979). Initially, seedling quality was defined by a plants physical appearance (Duryea and McClain 1984). With the trend toward physiological characterization of planting stock, which began with the observations of Wakeley in the 1940's and 50's (Wakeley 1948; Wakeley 1954), a high quality seedling is said to be one that is able to survive and grow on a particular planting site (Duryea 1984; Ritchie 1984).

The earliest attempts at developing standards to test seedling quality were based on morphological characteristics such as root:shoot ratio, shoot height, stem diameter, and root mass (Duryea 1984). A number of studies in the 1940's and 50's, using species of the genus *Pinus*, showed that seedling size is related to field performance (Chapman 1948; Pomeroy et al. 1949; Fowells 1953). Though larger seedlings generally appear to perform better in the field than smaller seedlings, Chapman (1948) found that the shortest seedlings had the greatest survival and growth in some environments. Additionally, Wakeley (1948) reported that large seedlings occasionally were of lower quality than smaller seedlings, and Zaerr and Lavender (1976) showed that seedlings from medium-sized classes had higher survival than those from large-sized classes.

Determining seedling quality based on morphology alone did not prove to be particularly accurate or reliable. The use of physiological characteristics to determine seedling quality began when Wakeley (1954) observed that outplanting performance depended not only on the visual or physical attributes of a seedling but, more importantly, on a seedlings physiological condition. Though seedling morphology is important to seedling performance, the physiological condition likely has an overriding effect on growth and survival (Ritchie 1984).

Planting failures have been linked to various physiologically controlled responses, such as a seedlings inability to regenerate new roots, take up water, and maintain a suitable nutrient status (Duryea and McClain 1984). In the 1940's Wakeley (1948) and Schopmeyer (1940) suggested that a seedlings ability to regenerate new roots immediately after planting could explain variation in survival observed among plants. Wakeley concluded that a seedlings ability to grow and extend its root system depended less on its morphological characteristics than on its physiological condition. A number or studies followed in the 1950's and 60's that revealed the existence of variation in root growth among nursery stock (Burdett 1987). Stone (1955) then began assessing root growth potential (RGP), or root growth capacity (RGC), as a measure of stock quality under the assumption that root growth immediately after planting was indicative of seedling establishment in the field.

Many studies have shown RGP to be correlated with early field performance (Ritchie 1985; Stone and Jenkinson 1971; Burdett 1979a; Larsen et al. 1986; van den Driessche 1983; Ritchie and Dunlop 1980; Sutton 1980; Burdett et al. 1983; Feret and Kreh 1985; Sutton 1987; Kormanik 1986). However, RGP measurements appear to be more strongly correlated to field survival than to field growth (McCreary and Duryea 1987). Although there is a physiological basis to the predictive ability to RGP, the relationship between root growth and seedling establishment remains unclear (Burdett

1987; Johnson et al. 1988). While it is probable that RGP measurements are actually predicting a seedlings ability to avoid desiccation or perhaps to take up nutrients (Nambier 1979; Burdett 1987), Ritchie (1985) suggests that RGP may instead be correlated with cold hardiness or other environmental stresses.

Although RGP appears to be under relatively strong genetic control (Jenkinson 1975; Johnson et al. 1988), it is also affected by environmental and physiological conditions. Studies indicated that RGP of coniferous seedlings is influenced by lifting date (Stone et al. 1962; Dewald and Feret 1987; Jenkinson and Nelson 1978), cold storage (Stone and Schubert 1959b; Feret et al. 1984), culturing practices (Krugman et al. 1965), soil moisture levels (Stone and Jenkinson 1970), and carbohydrate reserves (Duryea and McClain 1984; Puttonen 1986).

Several studies have addressed the importance of carbohydrate reserves to RGP. Shiroya et al. (1966) suggested that carbohydrates were allocated to actively growing roots and that levels in the roots were high in times of rapid growth. However, a cause-and-effect relationship between carbohydrate reserves and root growth was not shown and others have been unable to find such a correlation (van den Driessche 1978; Ritchie 1982).

Because carbohydrate reserve levels do not appear to be the only factor influencing root growth, Ritchie (1982) proposed that the status of bud dormancy and the supply of current photosynthate from the shoot may be alternate physiological factors controlling RGP. In a study using Douglas-fir and Sitka spruce seedlings labeled with $^{14}CO_2$, van den Driessche (1987) found that new roots were highly radioactive, indicating the importance of current photosynthate to new root growth. However, no relationship between new root growth and net photosynthesis was observed. Zaerr (1967), looked at another physiological factor, auxin, and its affect on root growth. Though auxin applied to disbudded ponderosa pine seedling shoots

caused a small increase in root initiation, auxin concentration had little to do with root growth potential.

Although the relationship between RGP and field performance is not clear, root growth capacity does provide a measure of seedling quality. In fact, RGP is the most commonly and widely used seedling quality evaluation technique (Ritchie 1985). Procedures for measuring RGP have basically remained unchanged since the methods development by Stone and his colleagues in the 1950's (Stone 1955; Stone and Schubert 1959a). Techniques of measuring root growth usually consist of counting or measuring the length of new roots after a period of growth under standard laboratory conditions. The time allowed for root growth is usually four weeks (Burdett 1987; Ritchie 1985). Alternative procedures have been suggested in order to decrease test cost and time required to obtain results. Burdett (1979a) modified the RGP method to yield results in one week by raising the temperature to increase root growth. A disadvantage is the requirement for a controlled environment chamber. Also developed by Burdett are root number classes of geometrically increasing size to which seedlings are assigned. This method has been shown to be effective for only some species (Ritchie 1985; Burdett 1979a). Volumetric determinations of RGP give an estimation of volume gained after a specified period of growth. However, results appear to be inconsistent and prone to error (Burdett 1979b; Ritchie 1985). A hydroponic system has been used with some success and has several advantages over potting seedlings. Root assessment is more accurate (roots are not broken and lost during unpotting), greater uniformity in root growth is obtained, and less maintenance and space is required (Winjum 1963; Ritchie 1985). It remains to be seen if the methods that speed assessment and lower costs are 'better' predictors of field performance than are complete root counts.

Another method used to evaluate seedling quality monitors survival and bud activity of seedlings growing under controlled conditions. This evaluation technique,

the vigor assessment method, was developed at Oregon State University and simulates stresses that newly outplanted seedlings are likely to experience (Hermann and Lavender 1979; McCreary and Duryea 1985). The stress treatment consists of placing half the seedlings under drying conditions for 15 minutes (32°C and 30% relative humidity) and measuring mortality and speed of bud burst over a two month period. The other half of the seedlings are potted and placed in the growth room as controls. A positive correlation between survival of stressed seedlings and field survival has been reported by Lavender et al. (1980) and McCreary and Duryea (1985).

A third approach used to evaluate stock quality measures the water potential, or plant moisture stress (PMS), of seedlings with a pressure chamber soon after planting. A damaged seedling will be unable to maintain a sufficient water balance and will show a decrease in water potential because of its inability to take up water (Ritchie and Hinckley 1975). Though plant moisture stress measurements are commonly used for scheduling irrigation (Zaerr et al. 1981), PMS immediately after planting has been used by McCreary and Duryea (1987) to help assess first- and second-year field survival and height growth. Unlike RGP and vigor assessment that characterize the overall physiological quality of seedlings, PMS measures a specific aspect of seedling physiology and is, therefore, a material attribute.

Another seedling material attribute used as a quality evaluator is root respiration rate (Johnson-Flanagan and Owens 1986). McCreary and Zaerr (1987) reported that while respiration rate was unable to detect damage caused by freezing, hot storage, or heating, it may help to assess damage caused by desiccation. However, as an indicator of overall quality this measurement did not appear promising.

Several nondestructive evaluation methods are being developed that measure a specific physiological condition of the seedling. Weatherspoon and Laacke (1985) reported the use of thermography to measure plant temperature. The basic premise is

that plant temperature indicates important physiological attributes that may be related to seedling quality. Though thermography was reported to be promising as an estimator of dormancy status, no further information on this technique has been published to our knowledge.

Vidaver and his colleagues at Simon Fraser University have developed a technique called variable chlorophyll fluorescence that is based on molecular phytochemistry and chlorophyll fluorescence quenching (Vidaver et al. 1981; Schreiber et al. 1977). This technique uses a fluorescence probe to measure the emission of excited chlorophyll molecules of a leaf mass and is being applied to operational forestry problems to allow for the direct assessment of photosynthetic activity of seedlings (Vidaver and Binder 1987). The basic premise of the chlorophyll fluorescence method is that photosynthetic activity, as measured through chlorophyll fluorescence induction, is indicative of a seedlings physiological status. The technique is said to provide a means to determine defects in photosynthetic activity due to damage sustained during storage and incomplete hardening off (Schreiber et al. 1975; Vidaver and Binder 1987). Although chlorophyll content is assuredly important to a seedlings overall photosynthetic activity, the role of photosynthesis in determining quality is obscure. Photosynthesis, and hence chlorophyll content, is not indicative of all types of damage that affect quality, such as impaired water uptake and certain nutrient deficiencies and diseases.

Biochemical measurements are also finding their way into quality assessment (Zaerr, 1985). Plant growth regulators (PGR), or hormones, have generally been found to be poor indicators of seedling quality. This is largely due to the difficulty of their measurement and the unknown nature of their physiological roles. Starch, though difficult to measure, may play a role in seedling quality prediction. However, further testing is needed. No one class of compounds, whether PGRs, starch, enzymes, or

sugars, is able to predict seedling quality. It is difficult, if not impossible, to choose an appropriate compound when so little is understood about the physiological processes involved in seedling survival and growth.

An approach suggested by Zaerr (1985) to circumvent this difficulty is to measure a large number of compounds and obtain a 'biochemical fingerprint' of a seedling. This 'biochemical fingerprint', or marker, approach has been used successfully in other areas. Schaefer and Hanover (1986) identified twenty-two monoterpenes from the cortical oleoresin of blue and Engelmann spruce that had diagnostic potential in the taxonomic studies of these species. Coastal and Rocky Mountain varieties of Douglas-fir were distinguished by the relative amounts of a number of terpenes from volatile leaf oils (von Rudloff 1971), and the terpene content of the essential oil of wood and foliage aided in population and hybrid identification of bristlecone and foxtail pines (Zavarin et al. 1976). Twenty flavonoid markers were found by Asen and Griesbach (1983) to distinguish cultivars of geranium. Lodgepole pine seed origin was identified by X-ray energy spectrometric mineral profiles (El-Kassaby and McLean 1985), and grand fir, white fir, red fir, noble fir, and inland and coastal Douglas-fir seed lots were identified from each other by analysis of terpene compounds isolated from seedcoats (Zavarin et al. 1979). Finally, phenolic compounds were found to be useful as biochemical markers of juvenility in walnut (Juglans sp.) (Jay-Allemand et al. 1987).

Of the seedling quality evaluation methods discussed in this literature review, none are able to predict accurately a seedlings field performance. It is difficult to characterize the physiological condition of a seedling because of the complex metabolic systems that control growth and survival. It is likely that successful prediction of seedling performance under various growing conditions will depend on the use of several quality evaluation methods.

MATERIALS AND METHODS

QUALITY REDUCING TREATMENTS

Douglas-fir seedlings were obtained from D.L. Phipps State Forest Nursery, Elkton, Oregon. Seedlings were 2+0 stock, lot #A92, from seed zone 491, elevation 1500 feet, and had been sown in 1984.

Over a period of seven months (October 1985-April 1986) seedlings were treated, planted in the growth room and field, and alcohol extracts of apical shoots prepared. Treatments were designed to create a complete range in seedling quality from which field and growth room performance could be determined using various evaluation methods (Bud Status, Root Growth Potential (RGP), and Survival). Four components that simulated possible harmful conditions experienced by seedlings during the lifting through outplanting process were manipulated to create thirty treatments (Table 1). The intent of the experimental design was not to study directly the effects of freezing, drying, lifting, or storage on seedling quality, but instead to use these factors to create the complete range of vigor.

1) Lift Date. An early lift in October and a late lift in March were used to introduce stress and decrease seedling vigor. The January lift, considered to be within the Douglas-fir lifting window for this lot of plants (Stone and Schubert 1959a; Jenkinson 1984), was not intended to decrease seedling quality but rather to yield seedlings of intermediate quality upon further treatment.

2) Freezing Temperatures. Exposure to freezing temperatures can occur during storage, in the nursery before lifting, or in the field after outplanting. The use of freezing temperatures in this study was designed to reduce seedling quality (the lower the temperature to which seedlings were exposed, the greater the reduction in quality) (Weiser 1970).

Table 1. Thirty treatments were created by manipulating four components: lift date (month), exposure to freezing (°C) or drying (minutes) conditions, and cold storage (months).

TREATMENTS

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
LIFT	Oct.	X	X	X	X	X	X	X	X	X																					
	Dec.										X																				
	Jan.											X	X	X	X	X	X	X	X	X	X	X	X	X							
	March																								X	X	X	X	X	X	X
FREEZING	-9		X										X							X						X				
	-12			X										X													X			
	-15				X										X						X							X			
DRYING	15					X										X						X						X		
	30						X										X						X						X	
	60							X										X						X						X
STORAGE	2								X										X	X	X	X	X	X							
	4									X																					

3) Drying conditions. During the lifting and planting processes roots can often be exposed, dry out, and become damaged, thus reducing quality. Seedlings were exposed to three levels of drying that simulated the damage that can occur to unprotected roots.

4) Cold Storage. A period of cold storage can deplete the carbohydrate reserves of seedlings or induce mold development (Lavender and Hermann 1976; Ritchie 1982). Some October and January-lifted seedlings were stored for 2 to 4 months at +2 to +4°C to create low to intermediate quality (Hermann et al. 1972).

LIFTING

Seedlings were lifted on four dates: 25 October 1985 (1,000 seedlings), 5 December 1985 (250 seedlings), 17 January 1986 (1,200 seedlings), and 10 March 1986 (800 seedlings). October, January, and March seedlings were lifted with a shovel by hand, immediately sealed in plastic bags and placed in a +2 to +4°C cold room within two hours. December seedlings were lifted by nursery personnel using routine lifting procedures.

Within four days of lifting, seedlings were culled to eliminate trees with multiple tops, major structural root damage, cranberry root girdler damage, and small plants (diameter at root collar of less than 4 mm). Roots were washed of mud and pruned to eight inches (December-lifted seedlings were root-pruned at the nursery before shipping). Within the October, December, and March lifts seedlings were randomly placed in groups of 70, sealed in plastic bags and stored in the cold room. January lifted seedlings were placed in groups of 140 seedlings.

Groups of 70 seedlings constituted a single treatment. After treatment of the 70 seedlings, twenty were potted and placed in the growth room for the root growth potential (RGP) test, another twenty were potted, also placed in the growth room and used to measure bud status and survival in the growth room, twenty were planted in the

field to determine one-year field survival, and ten were used in the search for potential seedling quality markers (Fig. 1).

FREEZING CONDITIONS

Groups of 30 seedlings were sealed in plastic bags and placed in a programmable freezer that lowered in temperature 2°C per hour starting from an initial temperature of +1°C. After reaching a temperature of -9°C, -12°C, or -15°C, seedlings were taken out of the freezer and allowed to thaw slowly at +2 to +4°C.

DRYING CONDITIONS

Drying treatments consisted of hanging seedlings with roots exposed for 15, 30, or 60 minutes in a controlled environment maintained at 32°C and 30% relative humidity. Before entering this environment, all roots were uniformly dampened. After exposure to the drying conditions, roots were placed in a bucket of water for five minutes to rehydrate them before being either potted and kept in the growth room or sealed in plastic bags and placed at +2 to +4°C until planting.

COLD STORAGE

Seedlings were sealed in double plastic bags and stored at +2 to +4°C for either two or four months. Seedlings were checked periodically for mold and roots were moistened if dry.

SEEDLING EVALUATION

GROWTH ROOM

Growth room evaluation was determined by three methods: Bud Status, Survival, and Root Growth Potential (RGP). Seedlings were potted in 4-liter pots containing forest soil. Twenty seedlings per treatment were chosen randomly, potted five seedlings/pot (four pots/treatment), and placed in a growth room at 21°C and 16

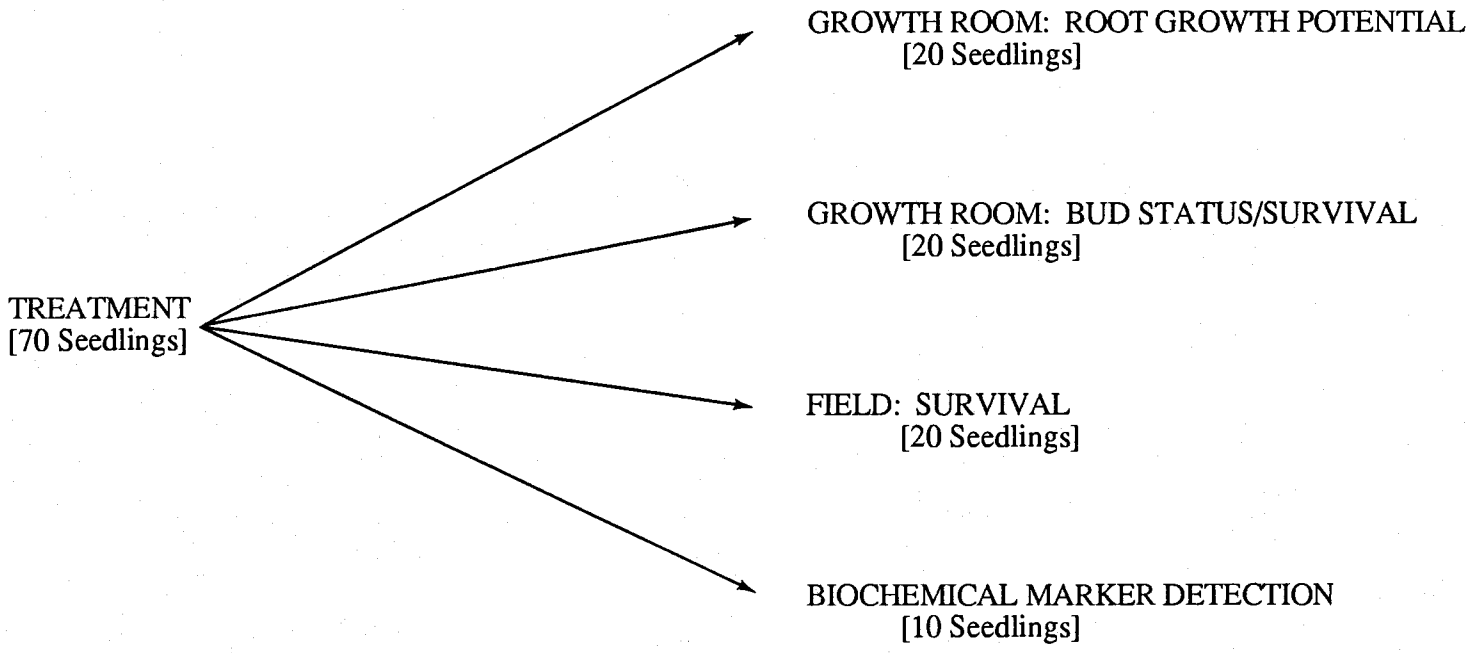

Figure 1. Distribution of seedlings within each treatment.

hour photoperiods. These seedlings were evaluated for bud status and survival. Another group of twenty seedlings per treatment , also randomly chosen, had all new roots removed and were then planted similiarly. These seedlings were evaluated with the RGP test. The position of pots in the growth room was changed weekly. Pots were watered twice a week to field capacity.

Bud status measurements began approximately one to two weeks after seedlings were potted and placed in the growth room. The terminal bud and most advanced lateral bud were observed from one to three times a week, depending on the general rate of activity. A numerical code was used to indicate the state of the bud (a '3' indicated a broken bud, a '5' designated bud flush). Observations continued for up to four months or until all buds had either flushed or died.

Six weeks after planting, the percent survival was determined for each treatment. Survival was indicated by the general state of the entire seedling-- dead buds, brown, desiccated, brittle needles, and a desiccated, shriveled stem were used as indicators of mortality.

After thirty days in the growth room, RGP seedlings were removed from their pots and the roots washed. All new root growth (white roots) longer than 0.5 cm were measured to the nearest 0.5 cm. Number of new roots and total length of new roots on each seedling was recorded.

FIELD

Twenty randomly chosen seedlings from each treatment were outplanted between the months of November 1985 to April 1986 at Peavy Arboretum Genetics Nursery located approximately 11 km north of Corvallis, Oregon. The outplanting plot, located in a rich clay soil, measured 30'x48', and was kept free of other vegetation. Seedlings were planted randomly by row at a 1'x1' spacing.

Seedlings were kept in an ice chest until planting began early in the day. Cool, overcast to light rainy days were chosen on which to plant. While seedlings were out of the ice chest their roots were kept in buckets of water.

Because competing grass and weeds flourished at the outplanting site, hand weeding was done in June 1986. Big Game Repellent (Powder-BGR-P; Deer-Away, Minneapolis, MN) and Gopher Bait (ORCO, Eugene, OR) were applied several times during spring 1986.

One-year survival was determined for each seedling in early November 1986. Survival was based on the general appearance of the seedling (as in growth room survival). If seedling survival could not be determined based on general appearance, the cambium near the root collar was used as an indicator-- a brown, desiccated cambium indicated mortality.

BIOCHEMICAL MARKER DETECTION

SAMPLE PREPARATION

Extraction

The terminal two centimeter section of the apical shoot (containing the terminal bud) of each seedling (10 seedlings per treatment, 30 treatments), was diced with a razor blade, transferred to a 10 ml plastic centrifuge tube (Sarstedt, West Germany) and homogenized in 2.5 ml of HPLC grade methanol (J.T. Baker Chemical Co.) with a Brinkman Polytron for 30 seconds.

Samples were centrifuged for 5 minutes at top speed using a DYNAC Centrifuge. One ml of the supernatant was transferred to a 1.5 ml eppendorf tube (VWR Scientific, Inc.) and dried under vacuum in a Savant Instruments Speed-Vac Concentrator for approximately three hours. Samples were stored at -80°C until purification.

Purification

To each dried sample 0.5 ml methanol was added and the tube mixed with a Vortex Mixer for five to ten minutes until the sample was dissolved. A 0.1 volume (0.05 ml) of the dissolved sample was applied to a piece of weighing paper, allowed to dry, and then weighed to determine the dry weight. To the remaining sample, 0.5 ml of OmniSolv Hexane (EM Science) was added, the tube mixed for one to two minutes, and the top hexane layer (containing chlorophyll and non-polar lipids) aspirated off. This hexane extraction was repeated six additional times. After removal of the hexane layer for the final time, samples were dried in the Speed-Vac for less than 30 minutes and stored overnight at -20°C. The following morning, 1.0 ml HPLC grade methanol was added and the sample mixed until dissolved. A 0.1 ml aliquot was transferred to a 0.5 ml eppendorf tube (VWR) and diluted with 0.1 ml HPLC grade methanol. Tubes were covered with parafilm and loaded into the automatic sampler to begin analysis. Samples were purified in groups of eight to twelve.

HIGH PERFORMANCE LIQUID CHROMATOGRAPHY

An automated chromatography system was assembled that consisted of a Perkin Elmer ISS-100 Automatic Sampling System, Varian Model 5000 High Performance Liquid Chromatograph, Beckman 164 Variable Wavelength Ultraviolet (UV) Detector, Perkin Elmer 650-10S Fluorescence Spectrophotometer, Varian Model 9176 Recorder, and an IBM-XT with a Keithley Series 500 data acquisition and control system (Keithley Data Acquisition and Control, Inc., Boston, MA) and Maxima Chromatography Software (Dynamic Solutions, Ventura, CA) (Fig. 2). A C_{18} reverse-phase column (Spherical ODS, 0.46 x 25 cm; particle size 5 μm, Burdick & Jackson Labs, Inc., Muskegan, MI) was used. All samples were detected with both the UV-absorption and fluorescence detectors and were recorded using the Varian recorder and Maxima Chromatography Software.

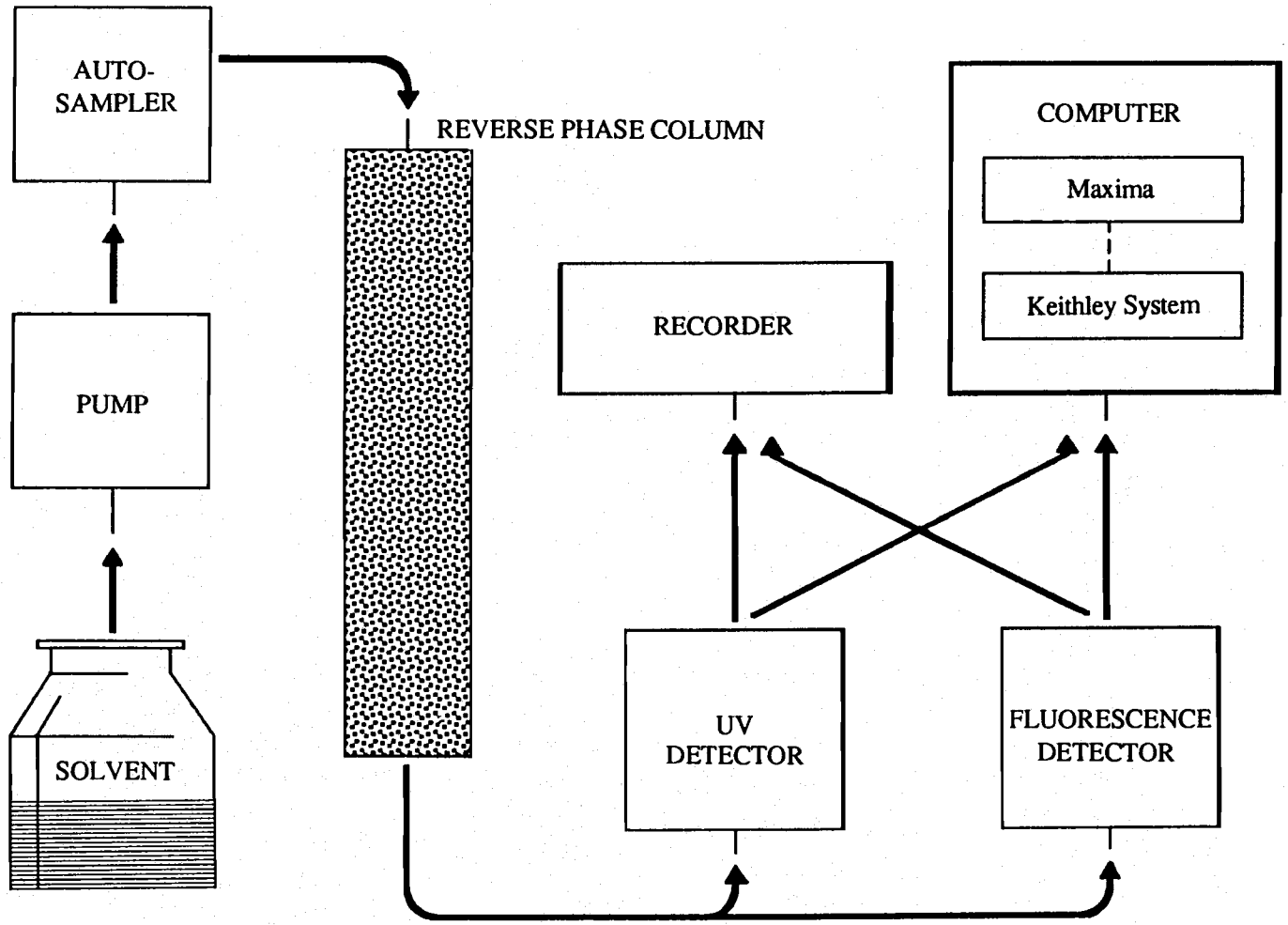

Figure 2. Automated chromatography system used to detect biochemical markers.

All analyses were performed on 25 µl of sample using a mobile phase flow rate of 1 ml/min. Pump B contained HPLC grade methanol. Pump A contained 20mM TEA (Tri-ethylamine) acetate buffer (glacial acetic acid, J.T. Baker), pH 3.37, filtered through a C_{18} column before use and degassed by bubbling helium through the solution. The UV detector was set at 254 nm and adjusted to a sensitivity of 0.05 AUFS and the fluorescence detector set to an emission wavelength of 360 nm, an excitation wavelength of 290 nm, and adjusted to a range of 1.0.

The gradient consisted of a 40 minute, four step linear solvent program (10% methanol at time 0, 65% at 20 min., 85% at 25 min., 100% at 27 min.) which permitted the best separation of a mixture containing an extremely large number of compounds. The system was allowed to equilibrate for 20 minutes at starting conditions between samples.

DATA PROCESSING

Keithley DAS 500

The Keithley Series 500 control system was used in conjunction with the IBM-XT computer for data collection. This system uses a bus extension to communicate with the host computer and is controlled through the Maxima software. Data acquisition and control tasks (analog input and output and signal processing) are executed by a series of Maxima subroutines and prompts.

Maxima

The auto-acquire option of Maxima was used for data acquisition. Two methods were employed, one to collect and integrate data detected by the UV detector and the other for data obtained using the fluorescence detector. The 'UV Method' was set at a range of 0 to 0.02, used 2.941 Hz (data points per second), and a y-axis value of 0 to 2 volts. Settings for the 'Fluorescence Method' included: range 0 to 0.01, y-axis 0 to 2 volts, and 2.941 Hz. Both methods used the same external trigger, run times of 40

minutes, and one minute pre-acquisition delays to eliminate pre-solvent front data points. Peak integration parameters for the UV Method were 18 baseline points, 17 filter window points, a coarse integration sensitivity of 5.333, a fine integration sensitivity of 2.698, a skim ratio of 8, and a rejection criteria for minimum peak area, height, and width equal to zero. The Fluorescence Methods peak integration section used the same criteria except fine sensitivity was set at 0.5509 and coarse sensitivity at 2.5. 'Auto-acquire' allowed for the fluorescence and UV chromatograms (those chromatograms resulting from detection by the Fluorescence and UV detectors, respectively) to be integrated and saved to floppy disks.

The sample queue was configured for 20 samples. Chromatograms were integrated, archived, and custom results (peak number, component name, retention time, and peak area) were written out to text files. An example of a representative Maxima table of custom results is presented in Appendix A. After all 300 samples were analyzed on the HPLC, a UV chromatogram was chosen that appeared to contain the largest number of peaks compared to other UV chromatograms. This chromatogram (peak numbers and retention times) was loaded into the Component Table and each of its peaks given a permanent name. Eight reference peaks were chosen out of the component table. Each UV chromatogram was compared to this component table in the Verify-Component Match option of Maxima. Peaks were identified on the basis of their retention times and given the permanent names from the component table. If a chromatogram contained a peak not appearing in the component table it would be labeled 'unknown'. The component table reference peaks aided in identification since retention times can change with column aging. After this procedure was repeated for the fluorescence chromatograms, there existed 600 text files (300 UV and 300 fluorescence).

After peak identification was completed, the resultant text files (the information from each chromatogram-- peak number, peak name, retention time, and peak area) were adjusted using three programs written in compiled BASIC (Appendix B).

The first program, 'TEST', aided in detecting any misidentified peaks by subtracting the retention times of the peaks in the component table from the retention times of the corresponding peaks in each text file (any peaks labeled 'unknown' in a text file are ignored). For example, if Peak D in the component table had a retention time of 2.8 minutes and Peak D in text file #11 had a retention time of 2.2 minutes, all peaks in the component table should precede all corresponding peaks in text file #11 by approximately 0.6 minutes. If all peaks (Peaks A-Z) had a difference of approximately 0.6 minutes except for one peak, Peak L, which had a difference of 1.8 minutes, that peak would likely be misidentified. This program was run once for each text file. The UV text files (300) were subtracted from the UV component table and the fluorescence text files (300) from the fluorescence component table. Any misidentified peaks were manually corrected in Maxima.

The purpose of the second program, 'FIX', was to create a master file (a file containing peak names and retention times of every peak appearing in any text file). A separate master file was created for the UV and fluorescence data (Appendix C). The following procedure is for the UV data. Fluorescence data were treated similarly .

Each text file contained unknown peaks, except for the text file (chromatogram) that was used to create the component table. The unknown peaks needed to be given permanent names that were consistent among all the text files. To illustrate: the second peak in text file #8 (an 'unknown') and the fourth peak in text file #32 (also an 'unknown') may actually be the same peak. What was needed was a master list of all peaks and their retention times before names could be given to the unknowns. To begin this program the information in the component table was used as a reference file or

starting point. Each time an unknown peak was encountered in a text file it was given a name, based on its retention time, and that name added to the reference file. For example, the reference file contains Peaks B, C, and D, a text file being analyzed contains Peak B, followed by two unknowns, then Peak D. The program takes the difference in retention time from Peak B to the first unknown in the text file and inserts its new name into the appropriate place in the reference file (either as Peak B-2 or as Peak C-2 depending on the difference in retention times between Peaks B, C, and D in the component table). The second unknown peak name would be inserted as either B-3, C-3, or C-2 if the first unknown was B-2. To further illustrate: a second text file contained Peak B, followed by an unknown, then Peak C. If the reference file contained Peak B-2 (added from the previous text file), and the unknown peak fell between Peaks B and B-2, based on retention time, the reference file would insert the unknown peak and its retention time as B-2 and rename the old B-2 to B-3. The final reference file is refered to as the master file.

'FILL', the last program in the series, compares each text file against the master file (UV text files against the UV master file and the fluorescence text files against the fluorescence master file). An 'unknown' peak in a text file is given a permanent name from the master file based on the peaks retention time. Peak names appearing in the master file but not in a text file are inserted into the text file and given an area value of ten. Because most peak areas are on the order of 50,000 to several million microvolt*seconds, an area of 10 is small enough to be considered insignificant.

Final text files were imported into a Symphony (Lotus Development Corporation, Cambridge, MA) spreadsheet and arranged into 300 rows of samples (10 seedlings per 30 treatments) by over 300 columns of peaks (UV detected peaks appended to fluorescence detected peaks). All peaks in each row were divided by the appropriate sample dry weight.

STATISTICAL ANALYSIS

GROWTH ROOM AND FIELD DATA

The ability of three seedling quality evaluation methods (root growth potential, bud status, and growth room survival) to predict field survival was determined by regressing the data obtained from each method separately against field survival. RGP data consisted of number of new roots greater than 0.5 cm per treatment (adjusted with a square root transformtion) and total length of new roots measured to the nearest 0.5 cm per treatment. Days to budbreak and budflush for both the terminal and most advanced lateral buds (also adjusted with a square root transformation) comprised the bud status data set. Six-week growth room survival values and one year field survival values were transformed with an arcsin(SQRT(percent/100)) equation.

All statistical analyses were performed using SAS programs. The predictive ability of the evaluation methods were determined using a SAS regression procedure. Residuals were plotted to determine the necessity of further transformation. To determine if the treatments applied to the seedlings created a full range of quality as indicated by survival and RGP, ANOVA's were performed using the SAS GLM procedure.

HIGH PERFORMANCE LIQUID CHROMATOGRAPHY DATA

The Symphony data file was converted to a SAS (SAS Institute Inc., Cary, North Carolina) file and peak areas rounded to the nearest integer value. Peaks judged to be randomly located among less than 20 samples were removed from the SAS file. All 230 variables were assessed for normality using the Shapiro-Wilk statistic (Shapiro and Wilk 1965) and adjusted with a square root transformation. One hundred twenty four peaks remained non-normally distributed after transformation.

Because the statistical analyses assume variables are normally distributed, various data sets were used to detect seedling quality markers. Data set 1 consisted of 230 peak areas for each of 300 seedlings . The second data set contained 230 peak areas averaged over the 30 treatments. When assessed for normality, only 43 peaks were considered to be non-normally distributed. The number and degree of non-normality in the data set was small enough to be considered acceptable for use in further analysis (19% of the variables had Shapiro-Wilk statistics < 0.8). The third data set consisted of the frequency of a peaks presence in each treatment. Peaks that were present in all 300 samples were dropped resulting in a data set with 209 peaks.

Data set 2 was collapsed to treatment means because survival, bud status, and RGP values were available only on a per treatment basis, not an individual seedling basis, and also to overcome the problem of extensive non-normality in the data. The third data set avoided much of the normality problem. However, by collapsing the data, all information concerning peak magnitude was lost and only that pertaining to peak presence or absence remained.

Principle component analysis (PCA) was performed as an exploratory technique using all three data sets. Two approaches were used for marker detection (Fig. 3). One was based on cluster analysis, the other on discriminant analysis. Cluster analysis groups individuals into unknown groups on the basis of similarity, while discriminant analysis classifies individuals into previously specified groups. Each approach was taken to identify markers of field survival, growth room survival, RGP (number and total length of new roots), freezing and drying damage, and time of lifting (these will now be referred to as the dependent variables). Data sets 1 and 2 were analyzed using both approaches. Bud status markers were not looked for since bud activity was unable to predict field survival.

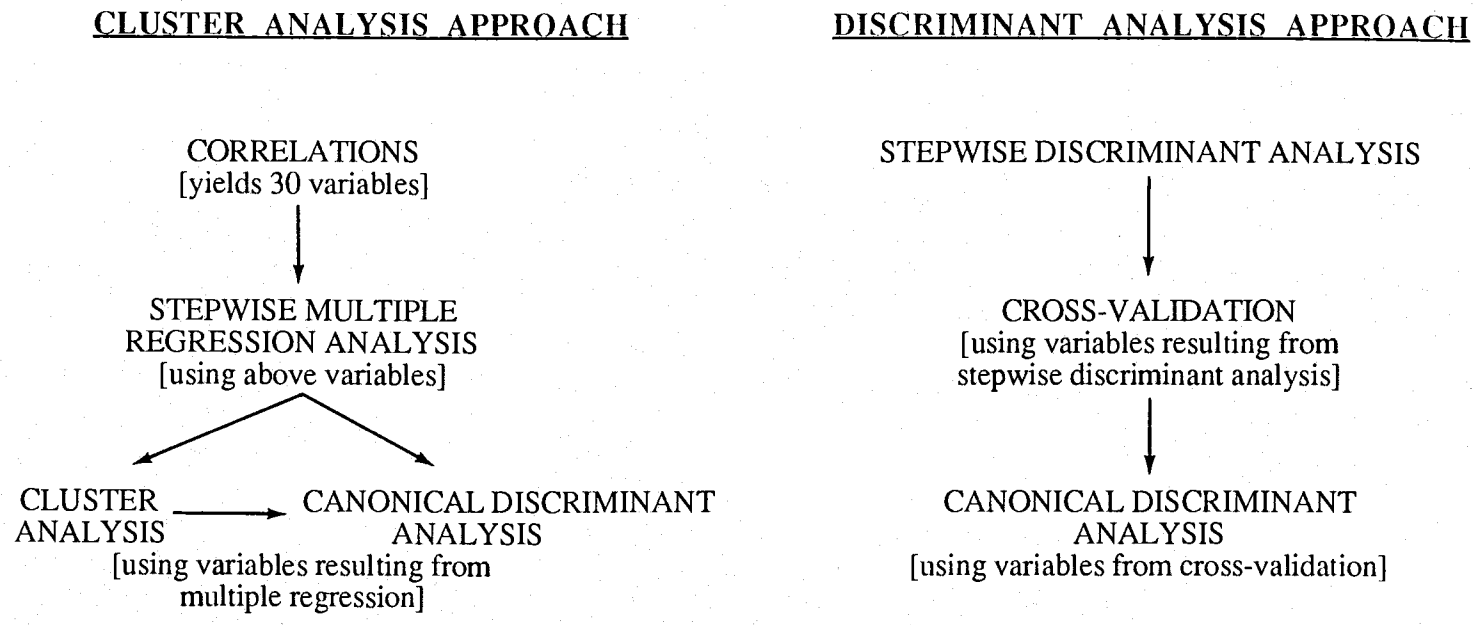

Figure 3. Two multivariate statistical approaches used to detect biochemical markers.

Cluster Analysis Approach

The first step in the cluster approach was to identify the 30 peaks most strongly correlated, negatively or positively, with the dependent variables (growth room survival, field survival, number or roots, total length of roots, freezing and drying damage, and date of lifting). These sets of 30 peaks were then used as variables in multiple regression analyses. Linear prediction equations were established that used the subsets of variables that explained a significant proportion of the variance of the dependent variable. Treatments were then grouped together by cluster analysis using the variables obtained from the multiple regression procedure. This series of analyses was repeated for each dependent variable using data set 2 (peak frequency data).

Discriminant Analysis Approach

The discriminant analysis approach to marker detection used stepwise discriminant analysis to identify those variables able to discriminate the specified groups. The crossvalidation option in the discriminant analysis procedure calculated the probability of misclassification of observations. Canonical variables were obtained using canonical discriminant analysis and plotted for visual representation. These analyses were repeated for both data sets, all dependent variables, and various *a priori* group specifications. For example, field survival was examined in two group arrangements (0-40%, 41-80%, 81-100% survival and 0%, 1-20%, 21-50%, 51-80%, 81-100% survival).

RESULTS AND DISCUSSION

CREATING A FULL RANGE OF SEEDLING QUALITY

In order to identify biochemical markers and to determine how well quality evaluation methods predicted survival, a broad range of quality among seedlings had to be created. Analyses of variance indicated that treatments did create a full range of quality as indicated by growth room survival, field survival, and root growth potential. In all cases there was a significant difference between treatments ($p \leq 0.0001$) (Fig. 4). Multiple comparison tests (least significant differences (LSD's)) indicated that a full range of quality was created among treatments by comparing every pair of treatment means (data not shown).

PREDICTING FIELD SURVIVAL WITH RGP, GROWTH ROOM SURVIVAL, AND BUD STATUS

Number of new roots and total length of new roots were good predictors of field survival. The linear relationship of number of new roots (transformed) against field survival was highly significant ($p \leq 0.0001$) and had an R^2 value of 0.54 (Fig. 5). Residuals showed no relationship, therefore no stabilizing transformation was made or additional variable terms added. Total length of new roots regressed against field survival was also highly significant ($p \leq 0.0002$) and had an $R^2=0.53$ (Fig. 5). Residuals did not indicate the need for transformation or variable addition. MS_E (mean square error) for number of roots equaled 0.062 (20 df) while the total length of roots had a MS_E of 0.06 (20 df). Therefore, the independent variable, number of roots, was slightly 'better' at predicting field survival than total length of roots. Number of new roots can also predict growth room survival ($p \leq 0.0001$, $R^2=0.87$, and $MS_E=0.17$). The linear relationship found between RGP data and field survival is not consistent with

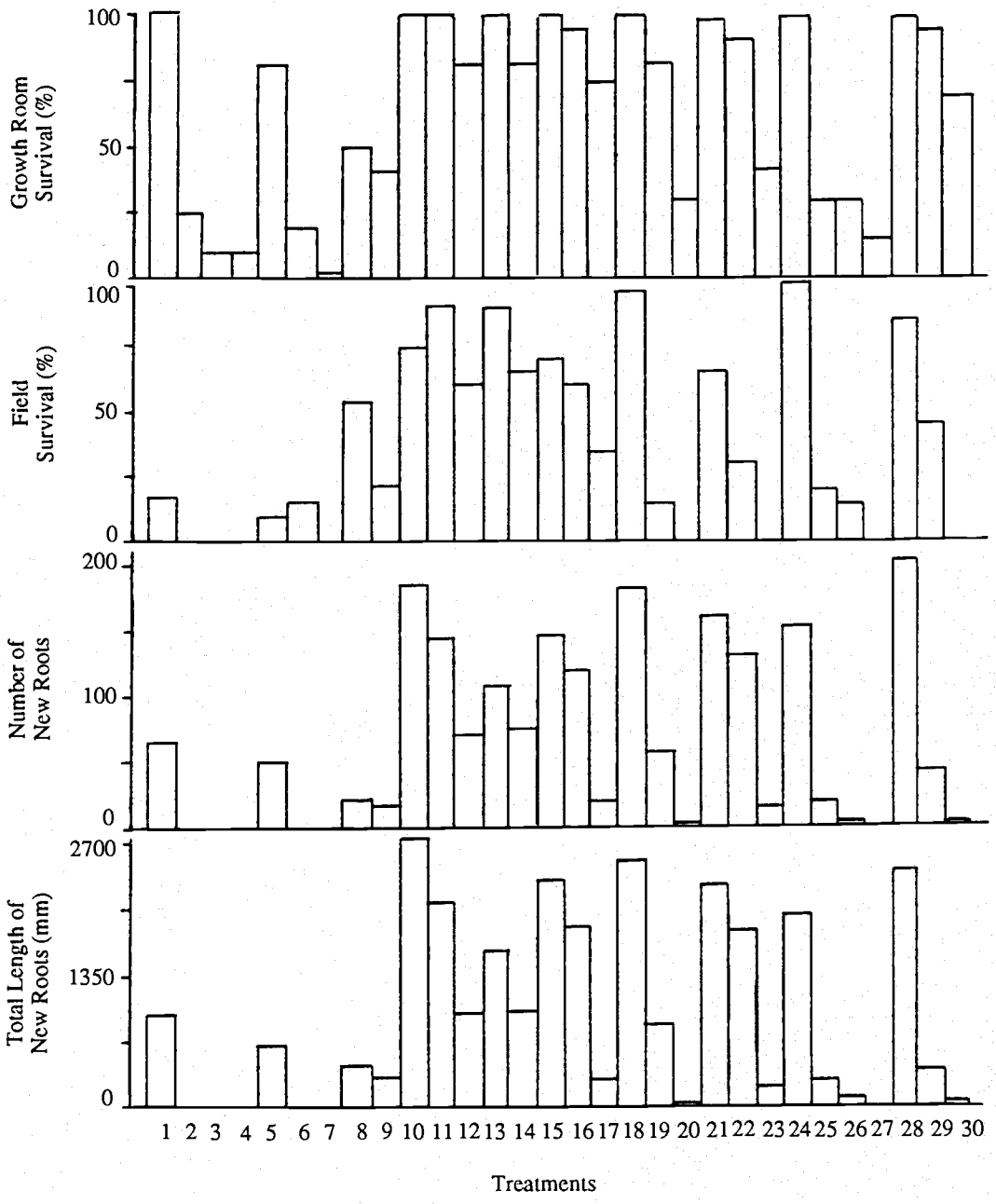

Figure 4. Histograms of seedlings subjected to quality reducing treatments. Each treatment value is an average of twenty seedlings. See Table 1 for explanation of treatments.

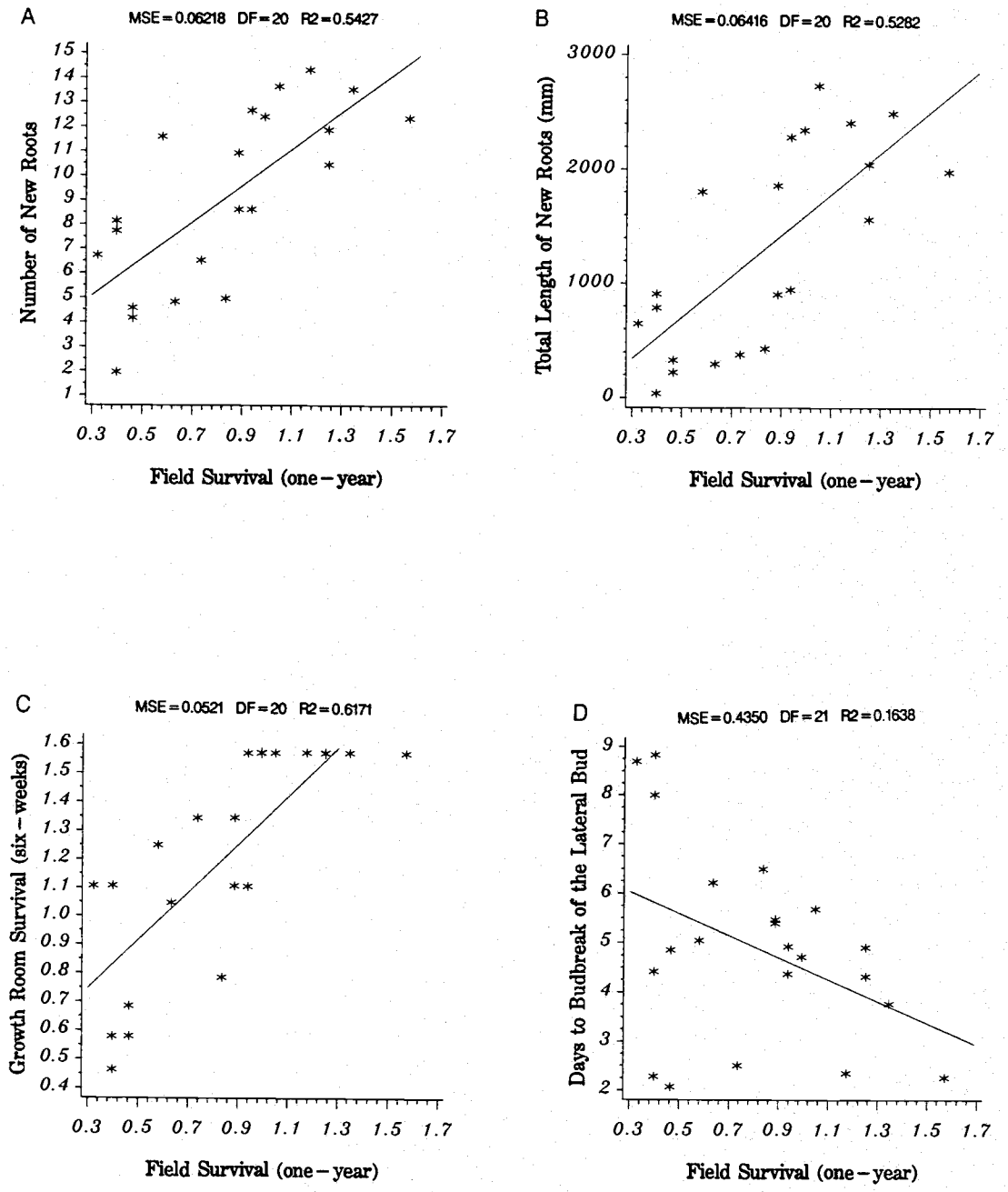

Figure 5. Transformed number of new roots (A), total length of new roots (B), transformed six-week growth room survival (C), and transformed days to budbreak of the most advanced lateral bud (D) regressed against one-year field survival.

that reported by others. Burdett et al. (1983) and McCreary and Duryea (1987) describe the relationship between RGP and field survival to be curvilinear rather than linear.

The linear relationship between growth room survival and field survival was highly significant ($p \leq 0.0001$) and had an R^2 value of 0.62 (Fig. 5). Residuals showed no relationship when plotted.

Bud status data consisted of days to budbreak and budflush (adjusted with a square root transformation) for the terminal and most advanced lateral buds. All four independent variables were very poor predictors of field and growth room survival (Table 2). The largest R^2 value, using days to budbreak of the lateral bud regressed on field survival, was only 0.16 (Fig. 5). All values were so low that no further conclusions will be made using this data.

HIGH PERFORMANCE LIQUID CHROMATOGRAPHY

Six hundred chromatograms were obtained using UV and fluorescence detection. A representative chromatogram detected with 254 nm UV light is presented in Figure 6 and a chromatogram resulting from fluorescence detection (360 nm emission, 290 nm excitation) is presented in Figure 7. The peak integration baselines and methanol gradient profiles are shown. Because these are only two of 600 chromatograms, not every peak will be present.

MULTIVARIATE STATISTICAL ANALYSIS

Multivariate analysis usually refers to all statistical methods that analyze multiple variables for each individual studied (Afifi and Clark 1984). However, to be truly multivariate, variables "must be random and interrelated in such ways that their different effects cannot meaningfully be interpreted separately" (Hair et al. 1987). Specific

Table 2. Results of bud status data regressed against field and growth room survival. Bud status data was adjusted with a square root transformation and survival data was transformed with an arcsin(SQRT(percent/100)) equation. Data are averages of values from twenty seedlings for each of thirty treatments.

Independent Variables		Dependent Variables: Field Survival			Dependent Variables: Growth Room Survival		
		R^2	MS_E	P	R^2	MS_E	P
Days to Budbreak	Lateral Bud	0.1638	0.4350	0.0326	0.0659	0.4319	0.1874
	Terminal Bud	0.0034	0.4647	0.7860	0.0572	0.4026	0.2605
Days to Budflush	Lateral Bud	0.1612	0.4357	0.0342	0.0342	0.4392	0.3463
	Terminal Bud	0.0026	0.4649	0.8118	0.0321	0.4080	0.4023

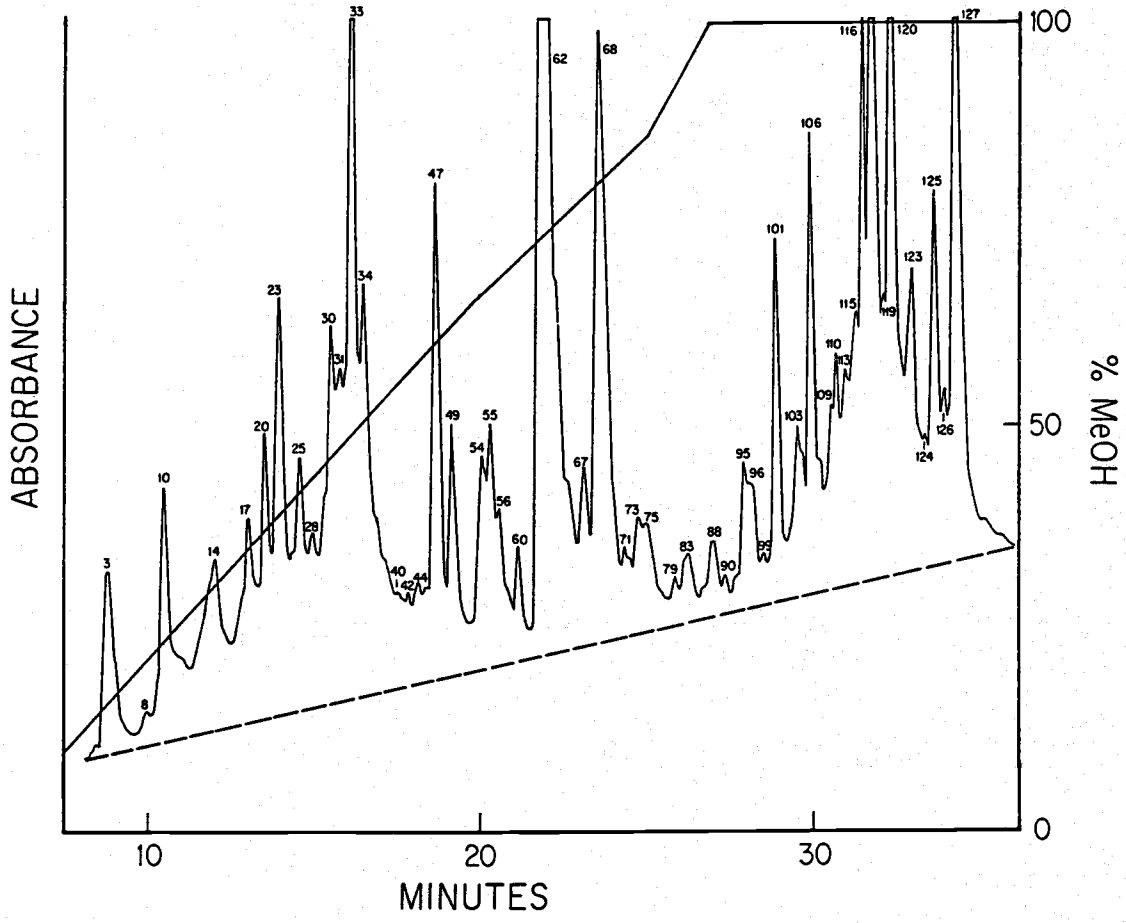

Figure 6. HPLC chromatogram showing the separation of a 25µl aliquot Douglas-fir apical shoot extract detected with a UV detector. The linear solvent program is indicated by the solid line. The dashed line represents the approximate baseline used for peak integration.

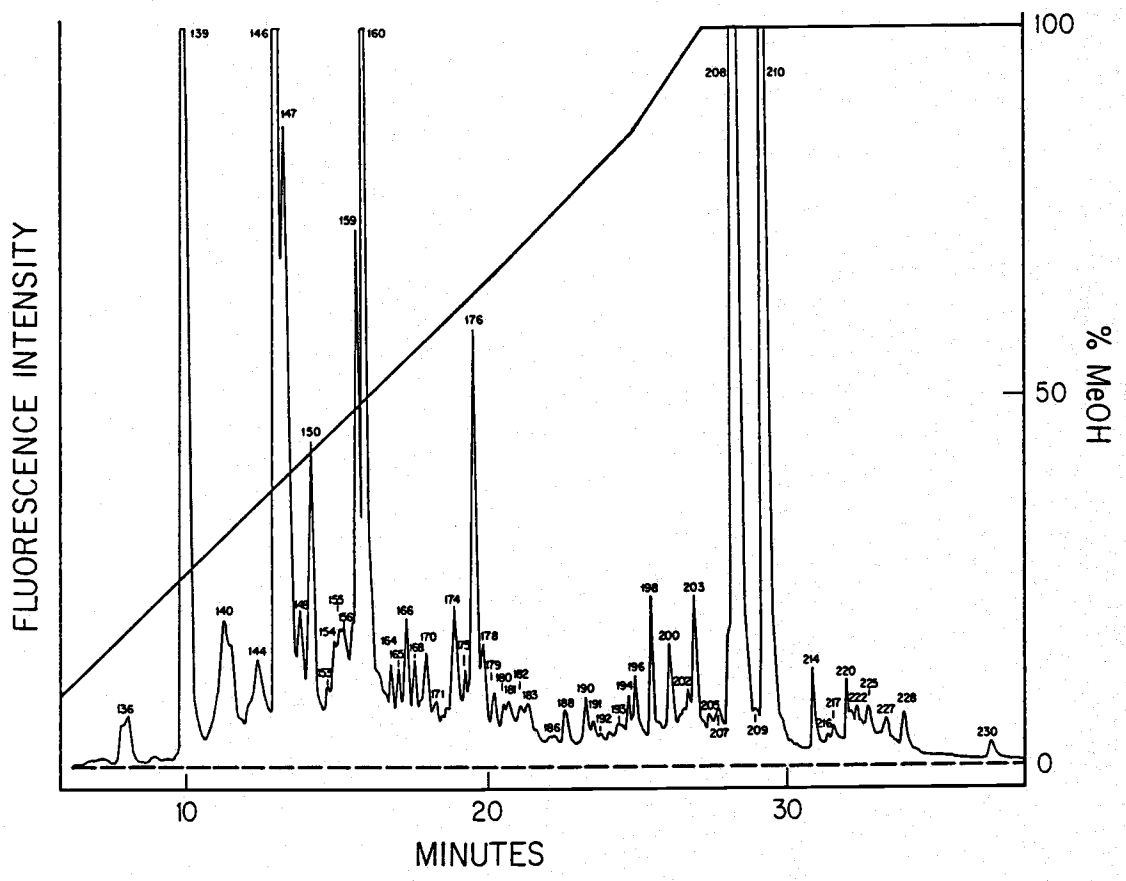

Figure 7. HPLC chromatogram showing the separation of a 25µl aliquot Douglas-fir apical shoot extract detected with a fluorescence detector. The linear solvent program is indicated by the solid line. The dashed line represents the approximate baseline used for peak integration.

multivariate techniques used to analyze the biochemical data were principal component analysis, multiple regression, cluster analysis, and multiple discriminant analysis. Multiple regression actually is a univariate procedure. However, for organizational purposes it will be included in the multivariate statistical analysis section of this chapter. Principal component analysis was used as an exploratory technique, while the other analysis procedures determined the relative contributions of the independent variables to a classification of observations or treatments.

PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) allows the structure of data to emerge without imposed constraints because there is no *a priori* placing of samples, or observations, into mutually exclusive groups (Flury and Riedwyl 1988). Variables were not eliminated, which is a common purpose of PCA, but rather transformed into new, uncorrelated variables called principal components. Principal components are linear combinations of the original variables and are arranged in order of decreasing variance with the most informative being the first (Afifi and Clark 1984).

PCA was performed using three data sets-- the initial peak area data set containing 230 variables and 300 observations (data set 1), data set 2 (peak area data averaged over 30 treatments), and the peak frequency data set (data set 3). Resultant components from the analyses were plotted to visually assess the structure of the data.

A well defined separation of observations was obtained using peak frequency data (Fig. 8). Groupings were strongly based on time of lifting and less strongly on freezing, drying, and storage treatments. A plot of components 1 and 2 split January and March-lifted observations into subgroups based on treatment exposure to freezing and drying conditions. A stronger separation by date of lifting is observed when plotting components 1 and 3 (Fig. 9). Treatment 9 (seedlings lifted in October and

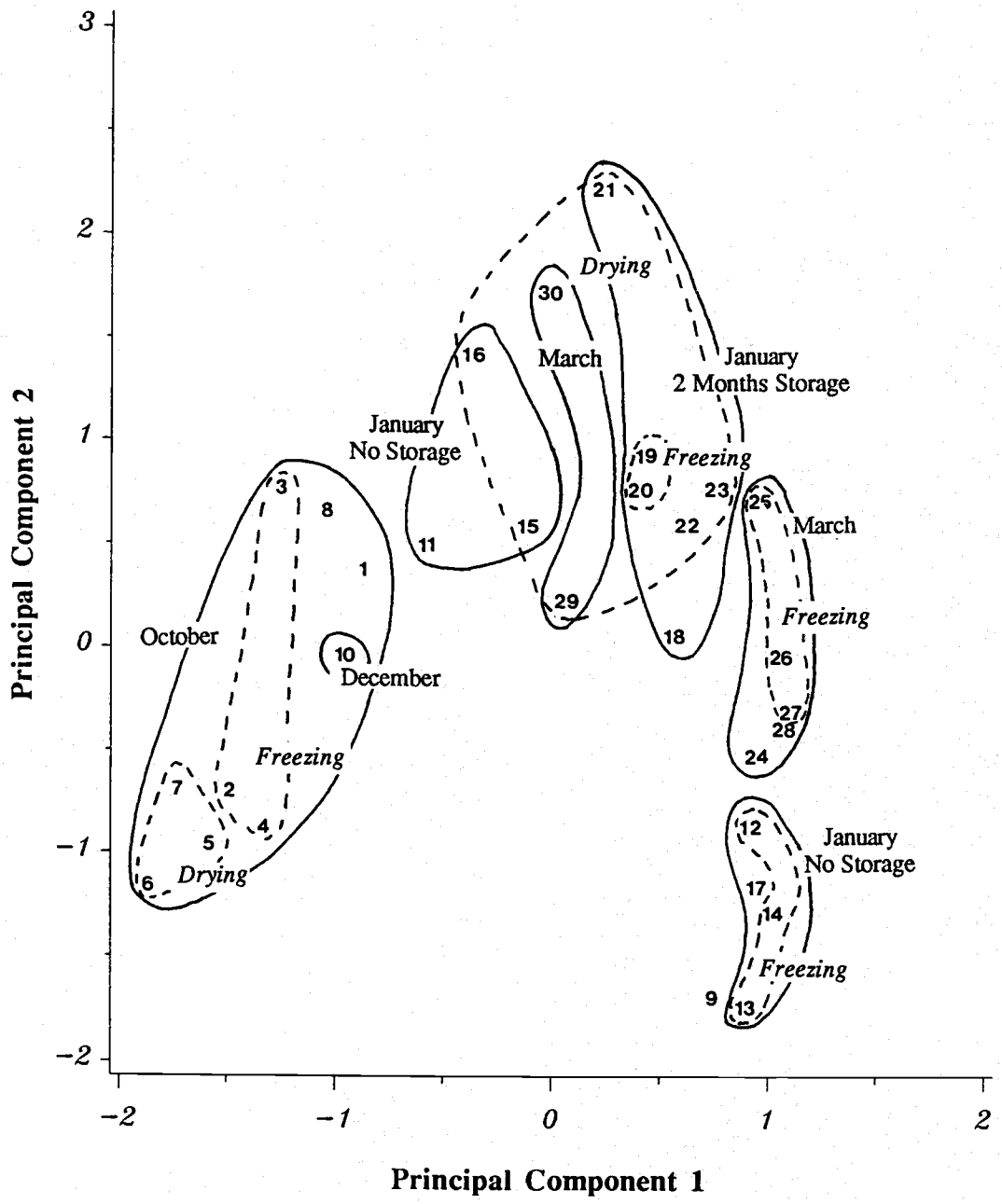

Figure 8. Scatterplot of the first and second principal components derived from peak frequency data for thirty treatments. Components 1 and 2 explain 49.31% and 16.18% of the total variation, respectively. See Table 1 for explanation of treatments.

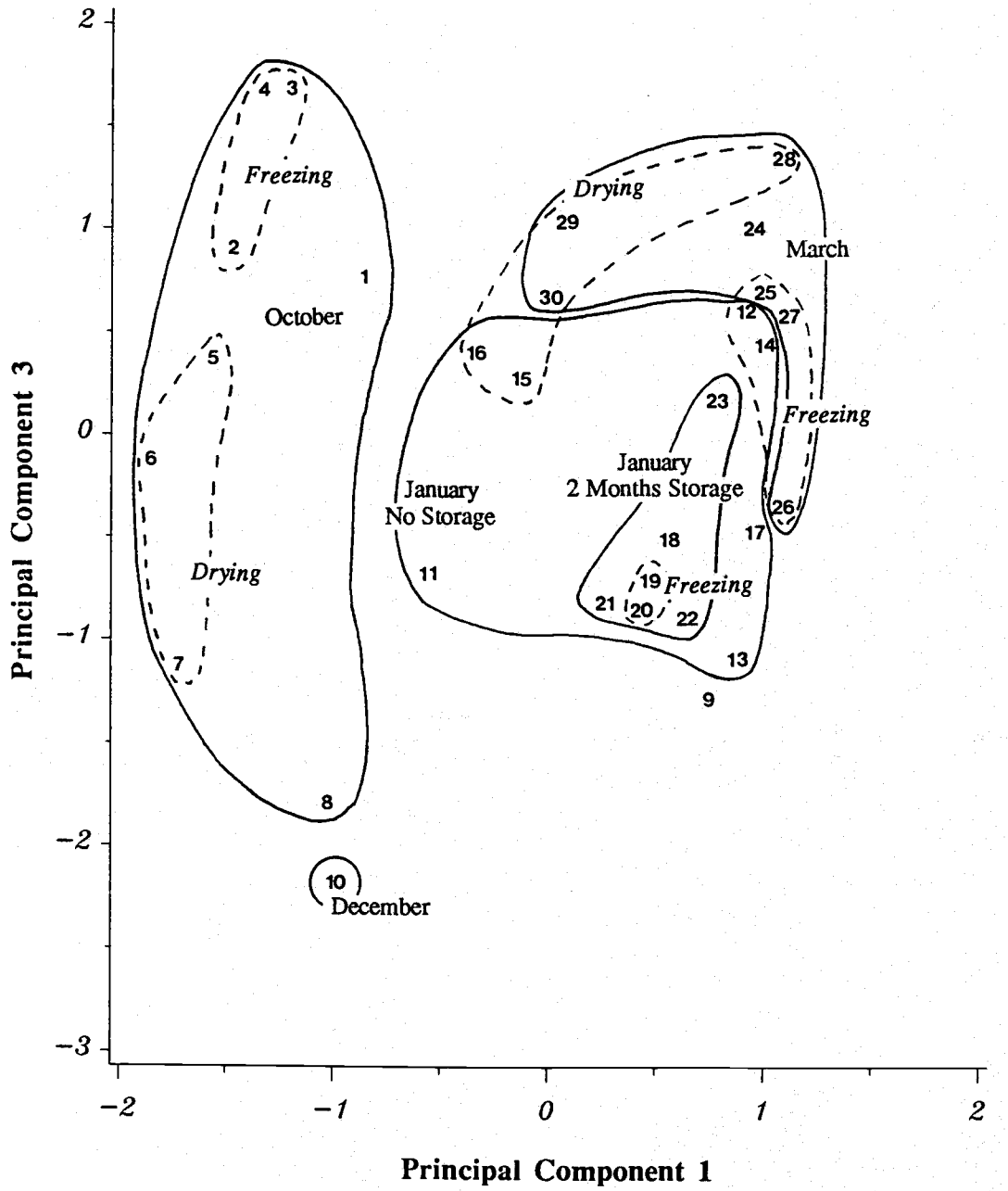

Figure 9. Scatterplot of the first and third principal components derived from peak frequency data for thirty treatments. Components 1 and 3 explain 49.32% and 13.18% of the total variation, respectively. See Table 1 for explanation of treatments.

stored 4 months) is well removed from all other October-lifted seedlings. Treatment 10, the only treatment lifted in December, is grouped with treatments lifted in October. The first three principal components explained 78% of the total variation. Principal component analysis of peak areas averaged over treatments (data set 2) yielded similar results (Fig. 10). Again, treatment 9 is well removed from the other October-lifted seedlings and the December-lifted treatment is grouped with those from October. Variation expained by the first three principal components was 95% of the total.

Data set 1 was analyzed to visualize the structure of the data on a seedling, not treatment average, basis. When plotting the first two principal components, observations were shown to group slightly on time of lifting, with October-lifted seedlings showing the strongest separation (Fig.11). There was no apparent distinction between unstored and stored January-lifted seedlings, or between seedlings exposed to freezing or drying conditions. However, 63 of the 300 observations remained hidden so no firm conclusions can be drawn from the figure. Only 49% of the total variation was explained by the first three principal components. Figure 11 does indicate that seedling variability is fairly low. Therefore, a treatment average based on ten seedlings is more than adequate.

Principal component analysis indicated a strong contribution made by date of lifting on the structure of the data. To overcome the strong influence of lift date, cluster analysis and discriminant analysis procedures were taken to identify biochemical markers.

CLUSTER ANALYSIS APPROACH

Cluster analysis groups individuals, or observations, into unknown groups and provides a measure of similarity between any two groups. Resulting groups should show high within-cluster homogeneity and high between-cluster heterogeneity (Hair et

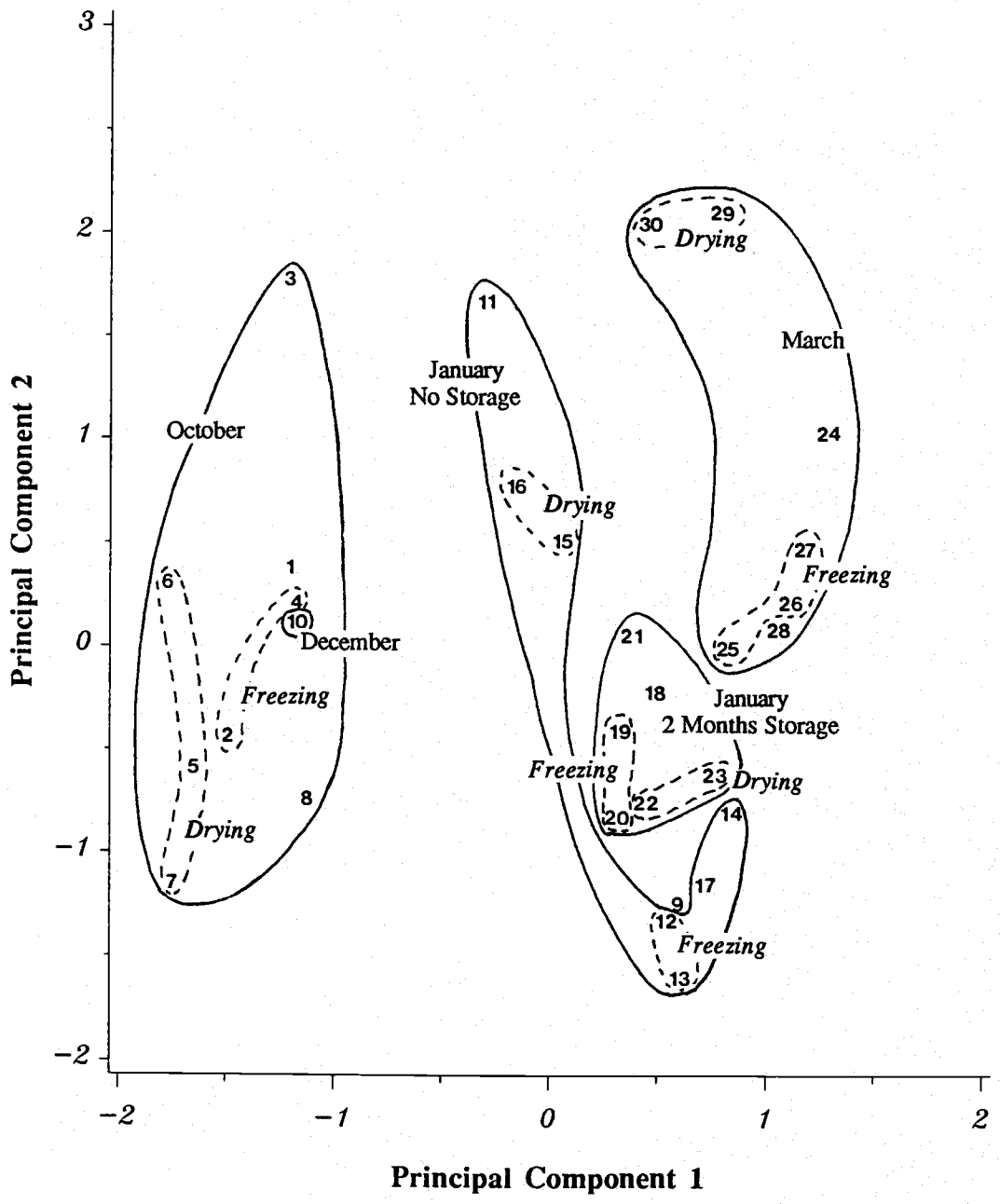

Figure 10. Scatterplot of the first and second principal components derived from mean peak area data for thirty treatments. Components 1 and 2 explain 55.10% and 24.73% of the total variation, respectively. See Table 1 for explanation of treatments.

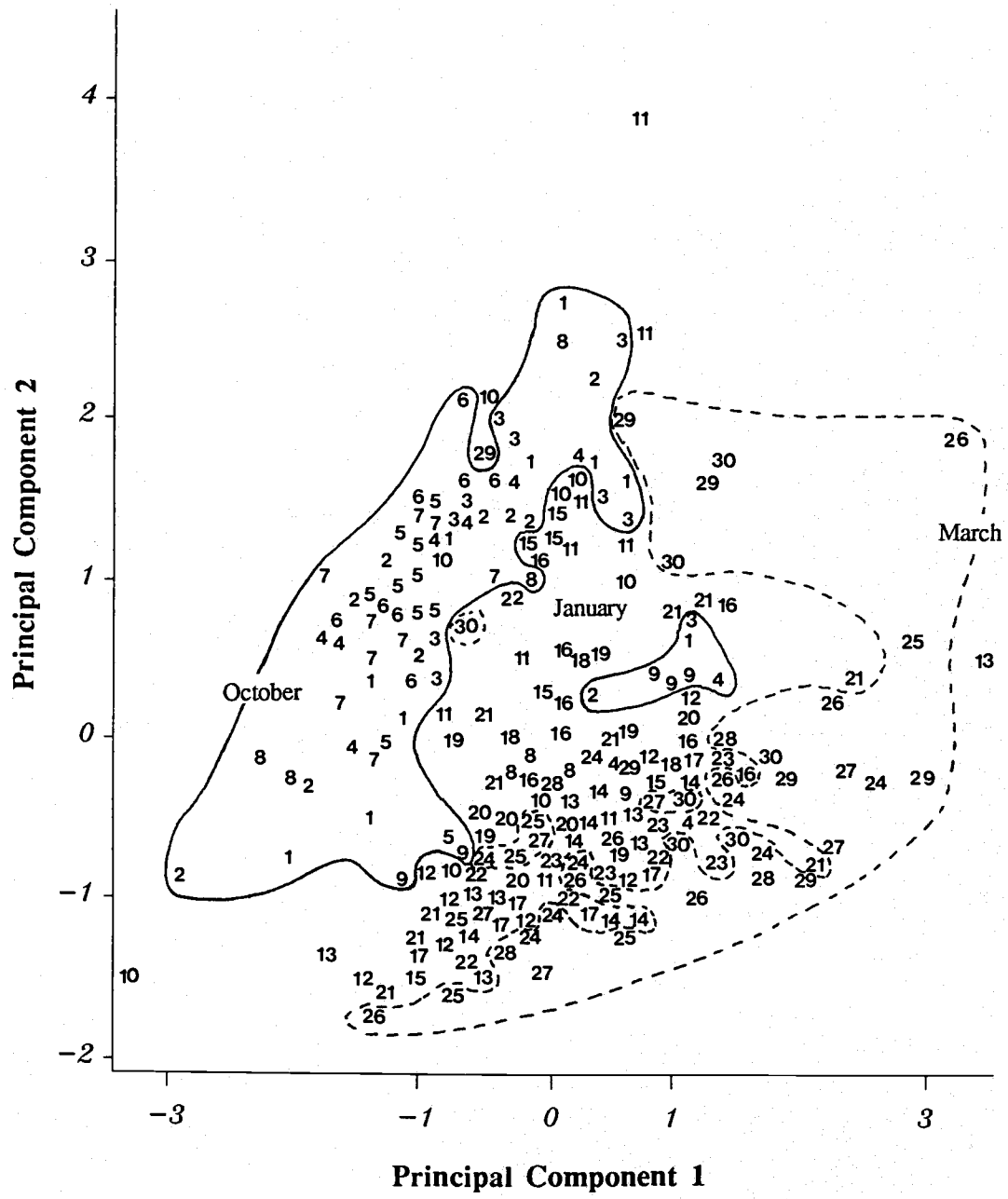

Figure 11. Scatterplot of the first and second principal components derived from peak area data for 300 seedlings. Components 1 and 2 explain 24.13% and 15.41% of the total variation, respectively. There are ten seedlings per treatment. See Table 1 for explanation of treatments.

al. 1987). The cluster analysis approach was taken as a means of identifying biochemical markers without classification of treatments into *a priori* defined groups. This approach was employed because group boundaries for field survival, growth room survival, and root growth potential were unknown. For example, it was not known if a treatment with 80% field survival was more similar, and hence belonged in the same group, to a treatment with 70% survival or 90% survival. Cluster analysis was expected to provide the appropriate group classification for field survival, growth room survival, and RGP, while aiding in marker identification for each dependent variable.

Preliminary analysis using all independent variables resulted in clusters based strongly on time of lifting (data not shown). This is, perhaps, not surprising considering the strong effect time of lifting had in the principal component analysis. However, to be able to obtain clusters on the other dependent variables, those peaks responsible for the overriding separation of treatments on lift date had to be removed from the analysis. The rationale used was to keep in each analysis only those peaks most strongly correlated with the appropriate dependent variable. This would then allow treatments to be clustered based on that dependent variable. Analyses were performed using data sets 2 (mean areas) and 3 (frequencies). However, the results obtained using data set 2 were inconclusive and, therefore, not presented.

Thirty peaks that were most strongly correlated to each dependent variable were identified and used as independent variables in a stepwise multiple regression procedure. Multiple regression was used to analyze the relationship between a dependent variable and multiple independent variables (Hair et al. 1987). The regression procedure determined which of the thirty independent variables should be included in a regression model. Regression models and R^2 values for each dependent variable are presented in Table 3. All variables met a 0.1500 significance level for entry into the model. Though this regression technique helps explain the relationships

Table 3. Linear multiple regression equations obtained by stepwise multiple regression analysis for each dependent variable using the peak frequency data set.

DEPENDENT VARIABLES	REGRESSION MODEL	R^2
Field Survival	Y= 7.18 + 0.07(**peak 4**) + 0.06(**peak46**) + 0.16(**peak78**) + 0.07(**peak 85**) - 0.73(**peak 103**)	0.72
Growth Room Survival	Y= -1.67 - 0.07(**peak 35**) + 0.08(**peak 78**) + 0.05(**peak 79**) + 0.24(**peak 119**) + 0.09(**peak 155**)	0.70
Number of New Roots	Y= -211.60 + 7.21(**peak 79**) - 6.88(**peak 172**) + 20.46(**peak 78**) + 19.05(**peak 93**) + 17.92(**peak 41**) + 9.36(**peak 138**) + 6.81(**peak 12**)	0.71
Total Length of New Roots	Y= 10067.16 + 99.43(**peak 12**) + 180.53(**peak 41**) + 180.18(**peak 46**) + 386.71 (**peak 78**) - 216.26(**peak 94**) - 1153.89(**peak 144**)	0.71
Time of Lifting	Y= 3.31 - 0.28(**peak 20**) - 0.07(**peak 50**) - 0.10(**peak 61**) - 0.17(**peak 111**) + 0.10(**peak 152**)	0.91
Exposure To Freezing and Drying Conditions	Y= -22.93 - 0.64(**peak 27**) + 0.93(**peak 29**) + 0.60(**peak 45**) - 0.31(**peak 57**) + 0.38(**peak 97**) + 0.48(**peak 104**) - 1.52(**peak 117**) + 0.49(**peak 122**) + 0.13(**peak 150**) + 1.35(**peak 180**) + 0.18(**peak 192**) + 1.21(**peak 207**) + 0.81(**peak 224**)	0.94

between the dependent and independent variables, it does not always yield the 'best' model or the model with the largest R^2 (Afifi and Clark 1984).

The variables that were included in the regression model were those then used in the final cluster analysis procedure. A hierarchical clustering algorithm, based on Ward's method, was used to calculate the distance between two clusters by adding, over all variables, the sum of squares between the two clusters (Everitt 1977). Dendrograms were constructed using output data from the cluster analysis. Canonical variables were next created with canonical discriminant analysis and plotted. Canonical variables are linear combinations of the independent variables and are derived in such a way that they best show the differences among the groups (Hair et al. 1987). The number of canonical variables equals either the number of variables or k-1 (where k equals the number of groups), whichever is smaller (Afifi and Clark 1984). F statistics are provided to test the equality of the group means. A plot of the canonical variables is useful in observing how well the groups are separated. Groups used in the canonical discriminant analysis were determined from the dendrograms. A dendrogram and corresponding canonical plot were constructed for each dependent variable using the peak frequency data set.

Markers of Field Survival

Treatments were clustered into four field survival groups (Fig. 12). Groups B and D generally contain treatments with the highest field survival while A and C contain those with the lowest. Several treatments with particularly low survival were classified into D (treatments 9, 27, and 22, with 20%, 0%, and 30% survival, respectively). Mean survival values seem to indicate two groups: A+C ($11.33 \pm 13.95\%$ survival) and B+D ($64.00 \pm 28.67\%$ survival). Seedlings classified into the A+C group would have a 0 to 25% chance of survival while seedlings found in the B+D group would have a chance of survival anywhere from 35 to 93%. The canonical plot (Fig. 13) has.

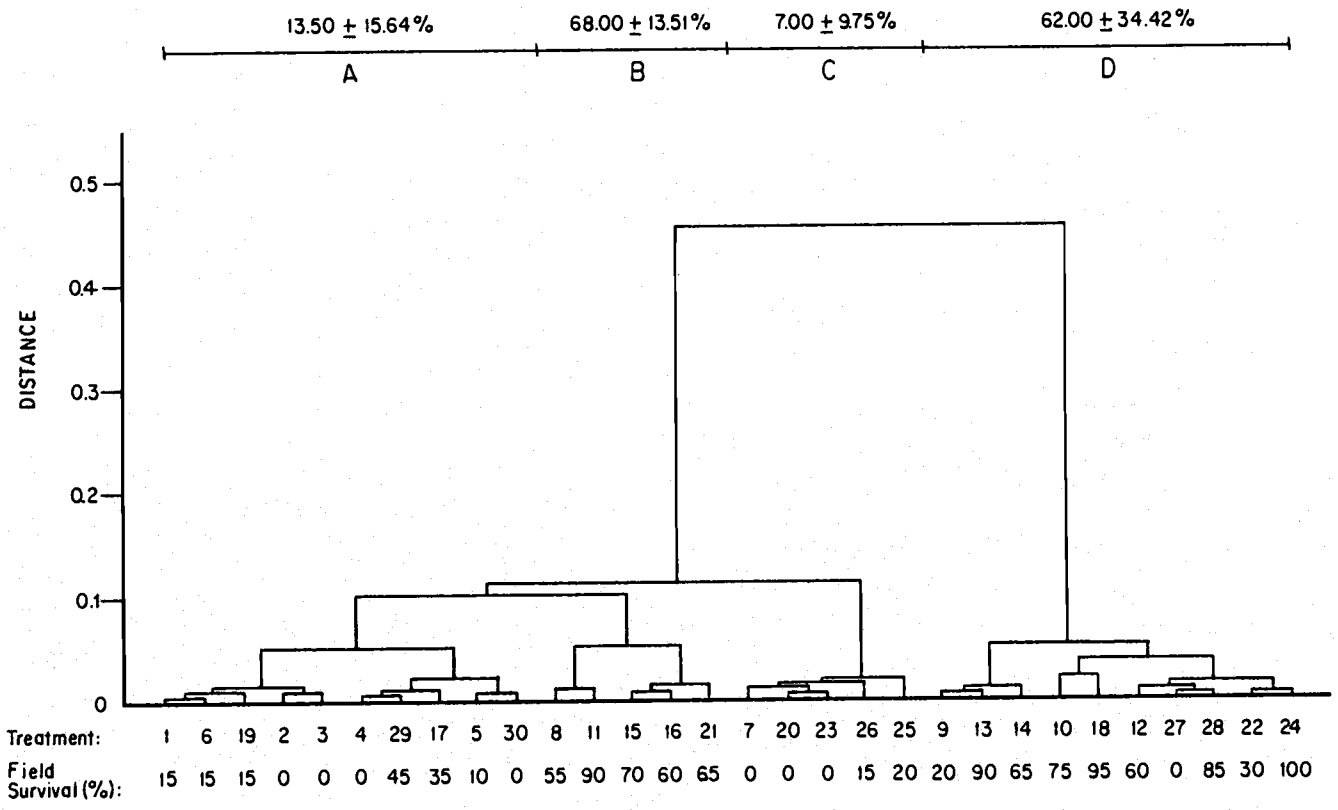

Figure 12. Dendrogram based on squared euclidean distance measures of thirty treatments of varying field survival. See Table 1 for explanation of treatments.

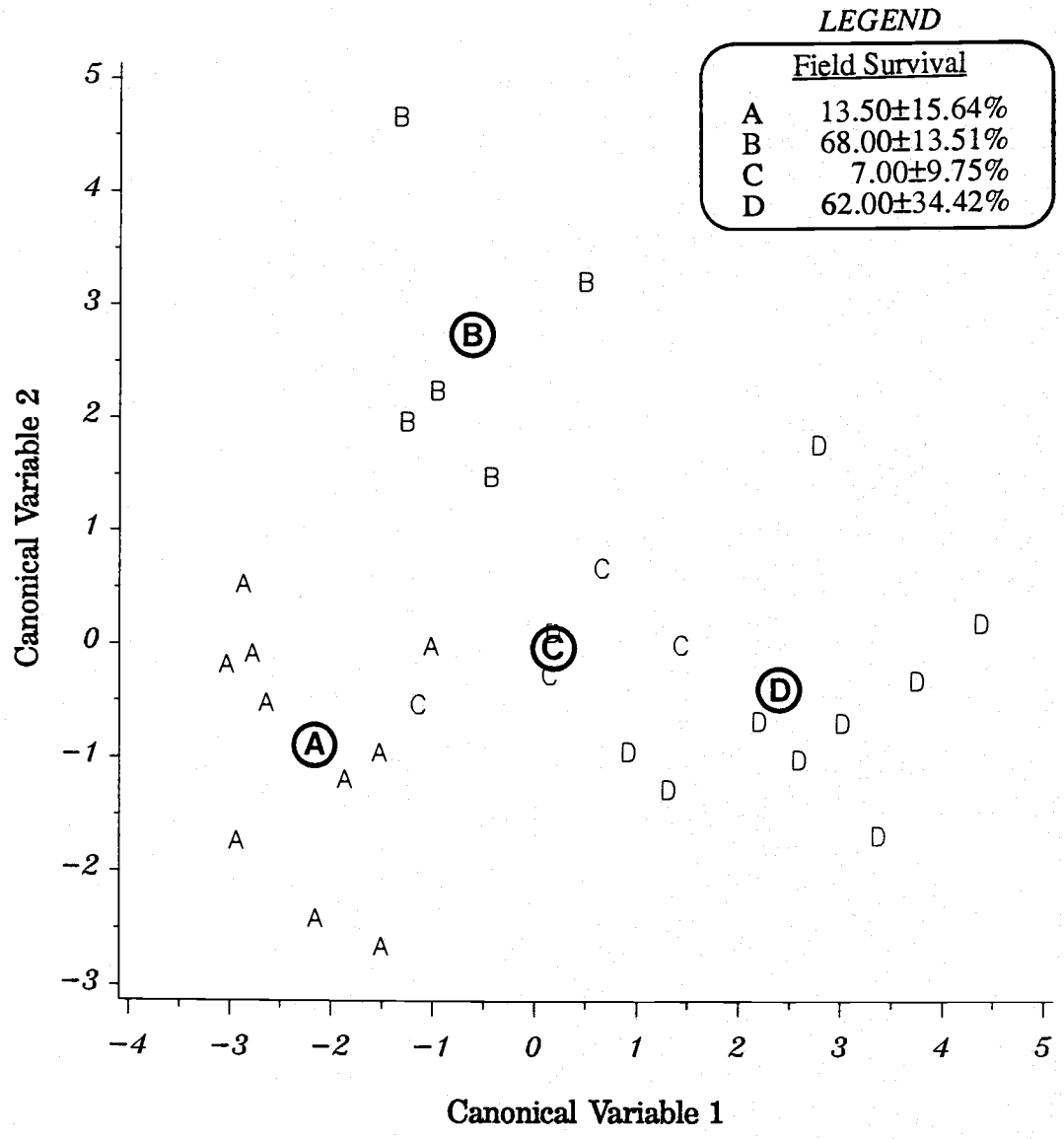

Figure 13. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak frequency data. Classes were determined from Figure 12. The first three variables accounted for 57.49% (0.0001), 24.19% (0.0001), and 18.32% (0.0001) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the four groups are indicated.

arranged the treatments into three groups with group C treatments being divided between groups A and D. The first canonical variable is responsible for explaining 57% of the variation.

Though there is a slight clustering of treatments based on survival as indicted by the dendrogram, standard deviation values are extremely large. Also, canonical plot groups are inconsistent with the dendrogram groups. There appears to be other factors, in addition to field survival, creating treatment similarity that leads to the observed clustering.

Markers of Growth Room Survival

Five groups were created based on growth room survival (Fig. 14). Group A contains every treatment with 100% survival except for treatment 13 (a January freezing treatment) which is classified as the only entry into group B. Though treatment 13, based on survival, should belong to group A, there apparently are other factors that make it dissimilar enough to be excluded. Standard deviation values for clusters are very large and there is a high degree of overlap between groups. Using only two groups, A+B+C (88.14 $\pm$ 22.13%) and D+E (38.93 $\pm$ 27.40%), the chance of growth room survival is 66 to 100% and 12 to 66%, respectively. The canonical plot shows a separation of the five groups (Fig. 15), with 'group' B varying greatly from other treatments on the second canonical variable. All four canonical variables are significant ($p<0.026$) while 64.29% of the variation is explained by the first canonical variable.

Markers of Root Growth Potential

Treatments based on the number of new roots clustered more distinctly than did treatments based on total length of new roots. Observations were clustered into three groups using number of new roots (Fig. 16). The canonical plot (Fig. 17) was consistant with the dendrogram indicating a well-defined separation of the three groups.

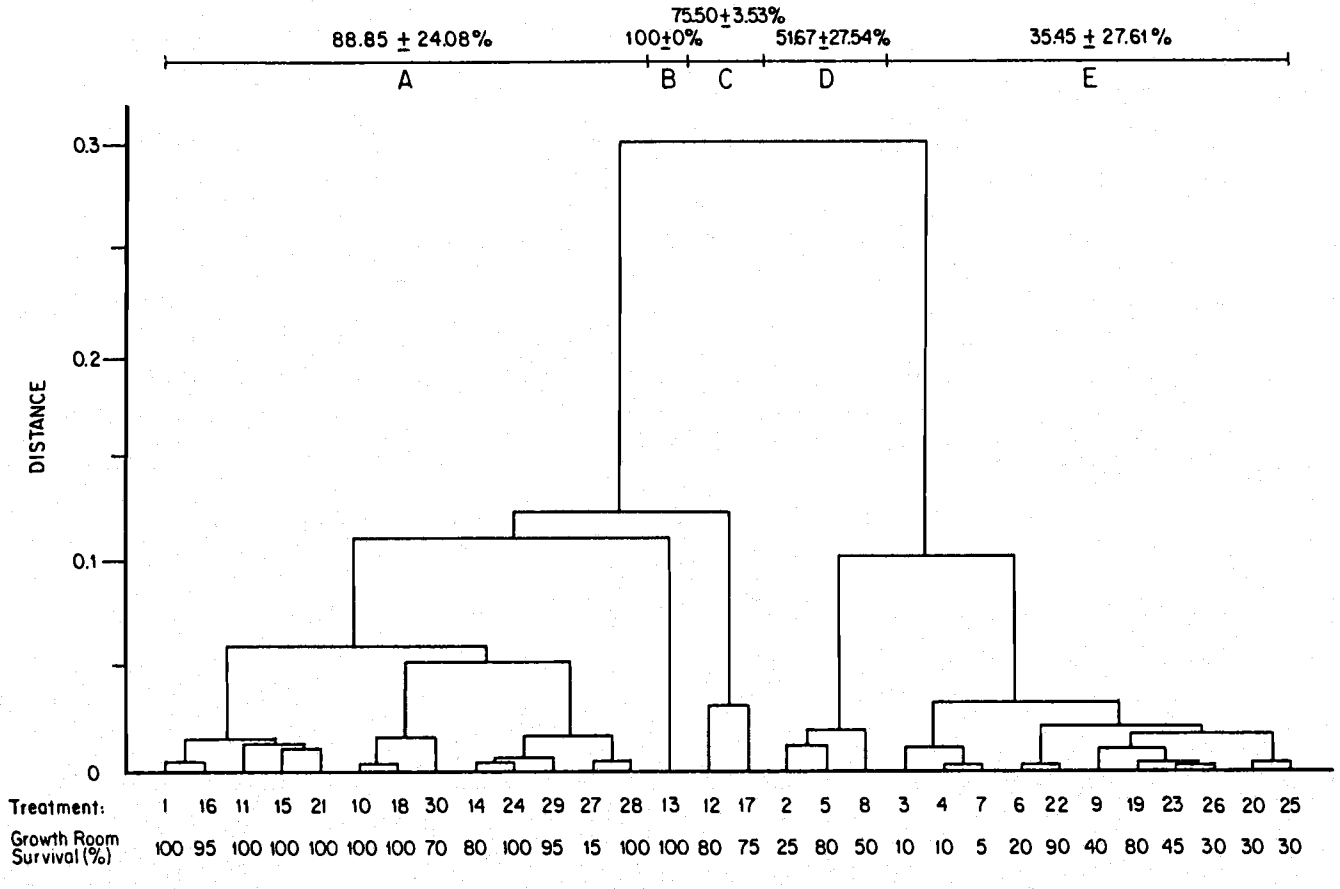

Figure 14. Dendrogram based on squared euclidean distance measures of thirty treatments of varying growth room survival. See Table 1 for explanation of treatments.

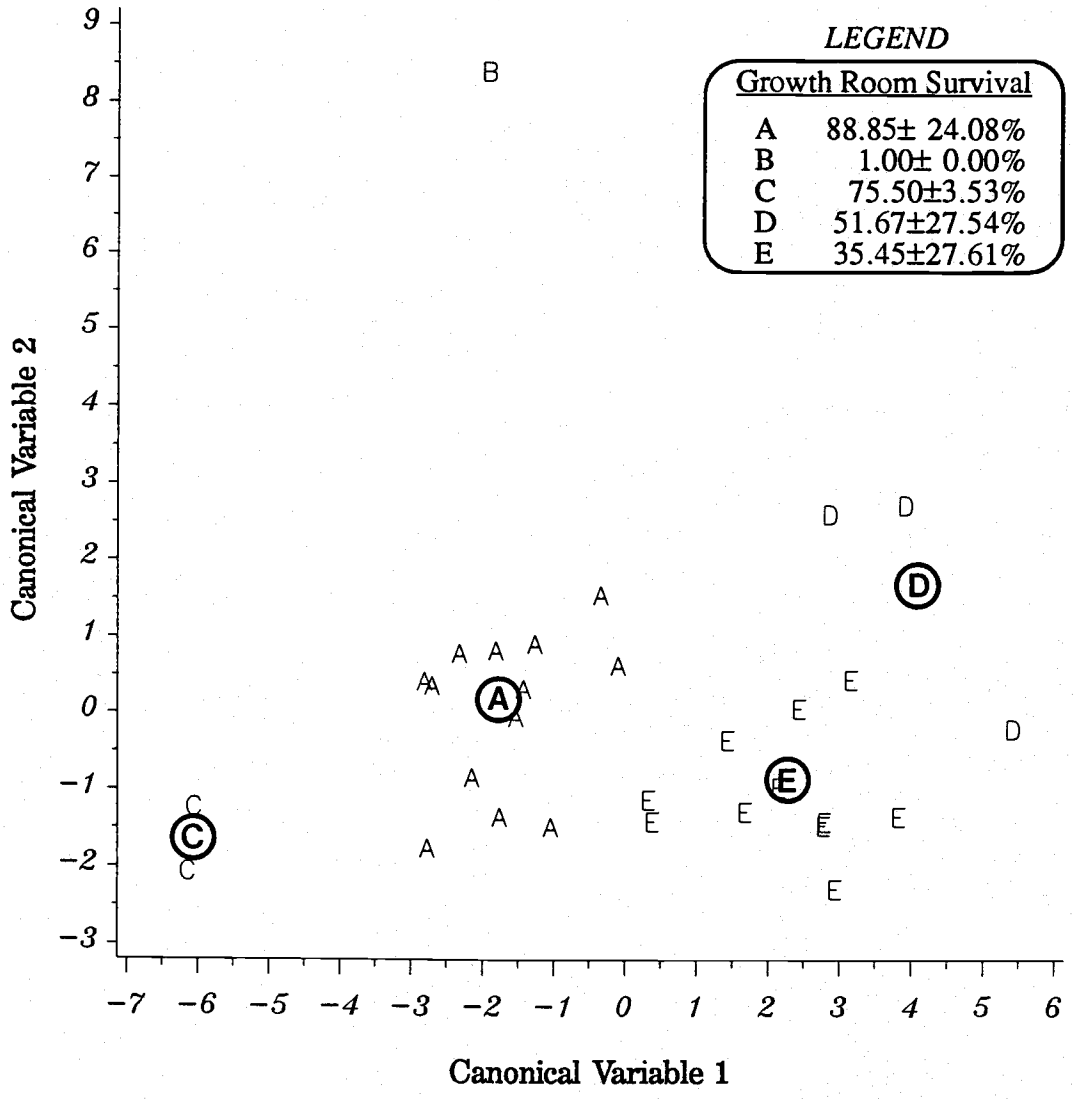

Figure 15. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak frequency data. Classes were determined from Figure 14. The first four variables accounted for 64.29% (0.0001), 28.43% (0.0001), 4.63% (0.0042), and 2.65% (0.0253) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the four groups are indicated. There is only one observation in group B.

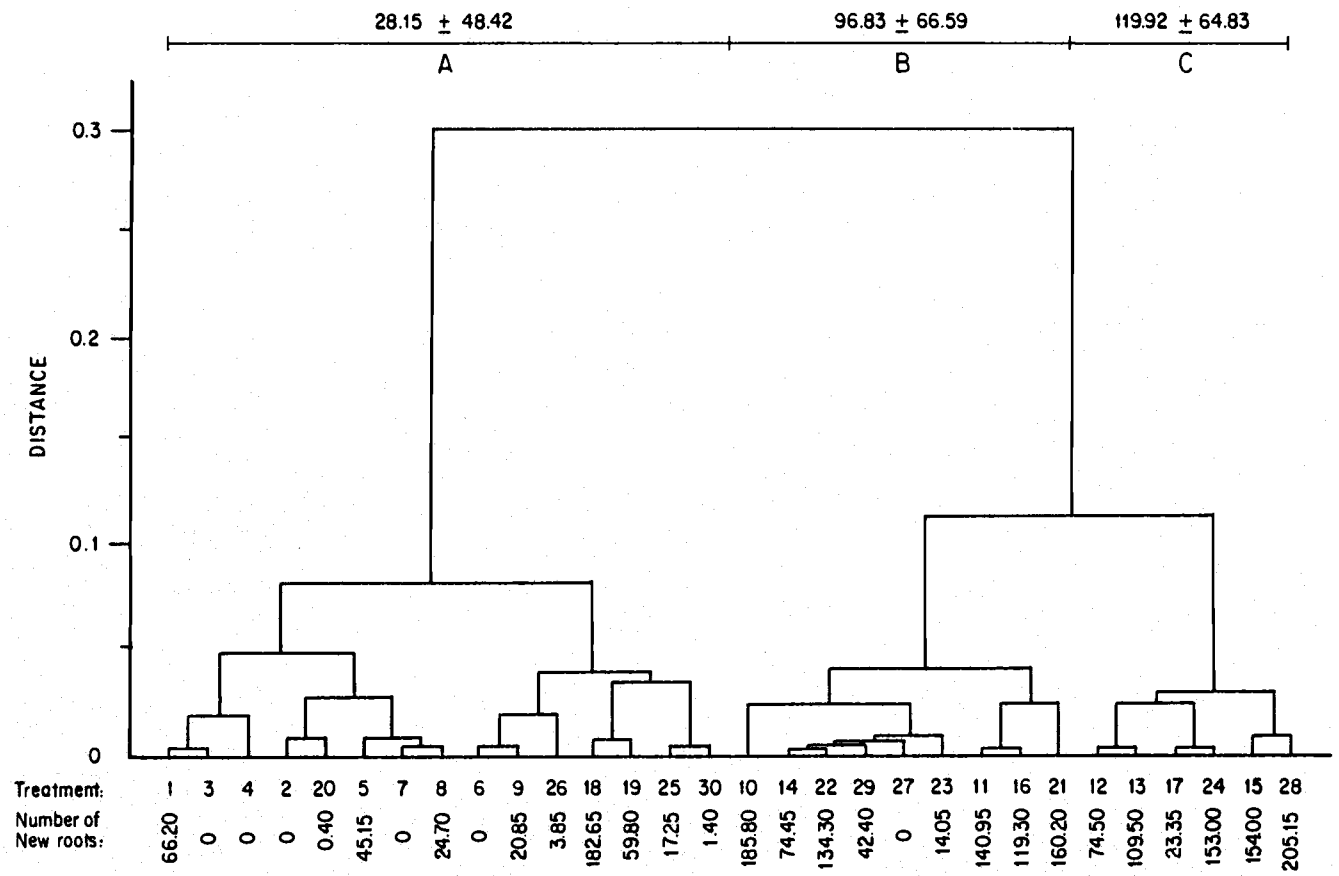

Figure 16. Dendrogram based on squared euclidean distance measures of thirty treatments with varying number of new roots. See Table 1 for explanation of treatments.

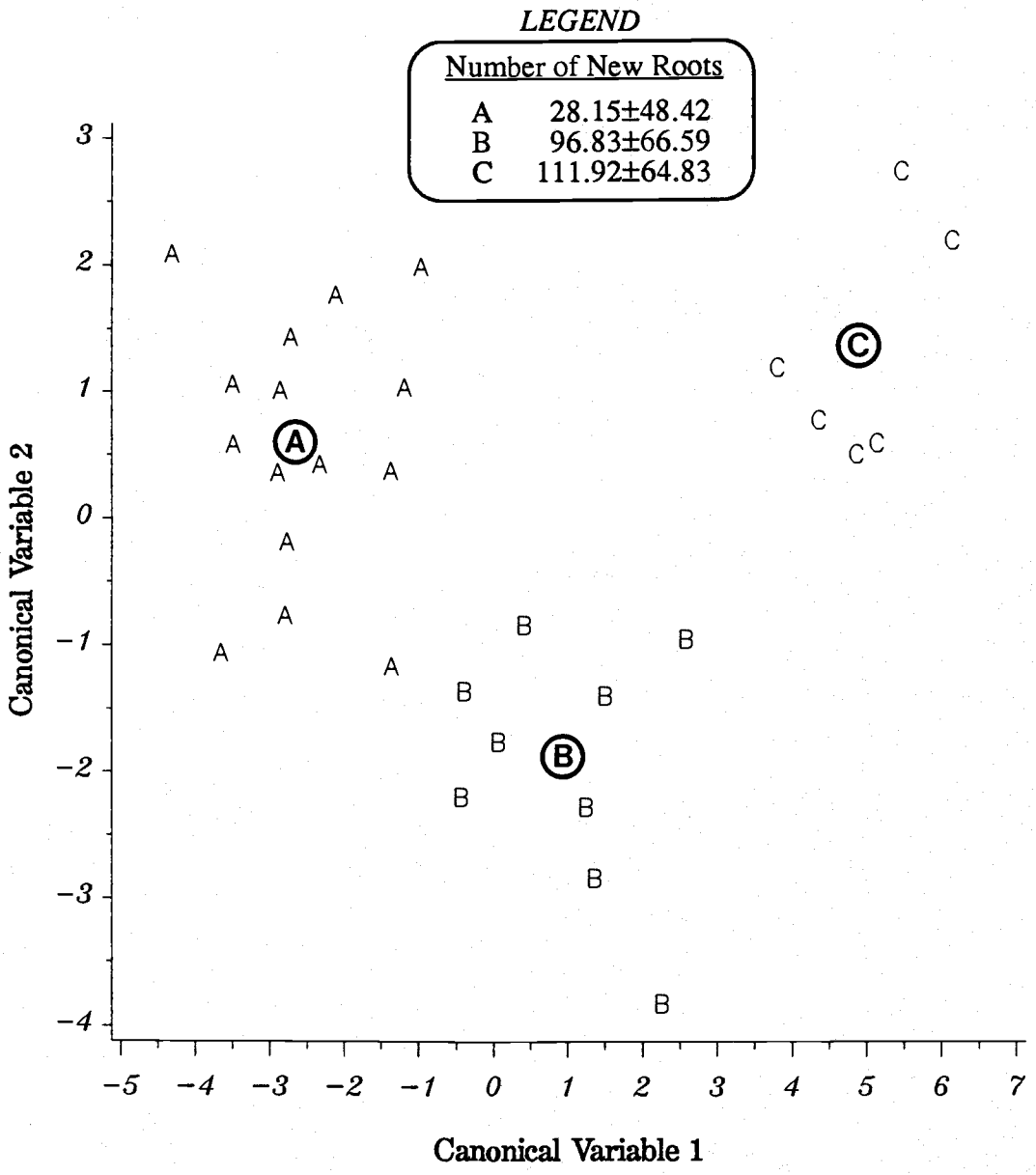

Figure 17. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak frequency data. Classes were determined from Figure 16. The first two variables accounted for 83.54% (0.0001), and 16.46% (0.0004) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the three groups are indicated.

Both canonical variables were significant ($p \leq 0.0004$) with an overwhelming 84% of the variation explained by the first canonical variable.

Six groups were indicated when clustering treatments based on total length of new roots (Fig. 18). However, standard deviation values for the groups were extremely large and group means differ slightly. Four groups were indicated when plotting the first two canonical variables (Fig. 19). Group C was included with group B observations and group D treatments were divided into groups A, B, and E. If treatment similarity was based only on total length of new roots, groups A (226.5 ± 453.0 mm) and B (246.5 ± 311.3 mm) should have been observed together on the canonical plot. Because they were indicated as separate groups in Figure 19, there were apparently other factors responsible for creating the similarity between treatments in the cluster analysis.

Markers of Freezing and Drying Damage

Well defined clusters for freezing and drying damage were not attainable using the cluster analysis approach. Freezing and drying treatments were scattered throughout each cluster (Fig. 20). The canonical plot (Fig. 21) indicated that treatments were being separated into three groups based on some unknown factors. Groups C and B, containing two freezing and four drying treatments, formed one group while groups A, D, and E, containing nine freezing and six drying treatments, formed another group. Separation of treatments was apparently not based on treatment exposure to freezing and drying conditions.

Markers of Lift Date

The ability to cluster treatments based on time of lift was very high. The only treatment lifted in December, treatment 10, was grouped with treatments lifted in October (Fig. 22). There were two March lifted groups, B and D, with the addition of one January-lifted treatment (treatment 12) into group B. Two treatments lifted in

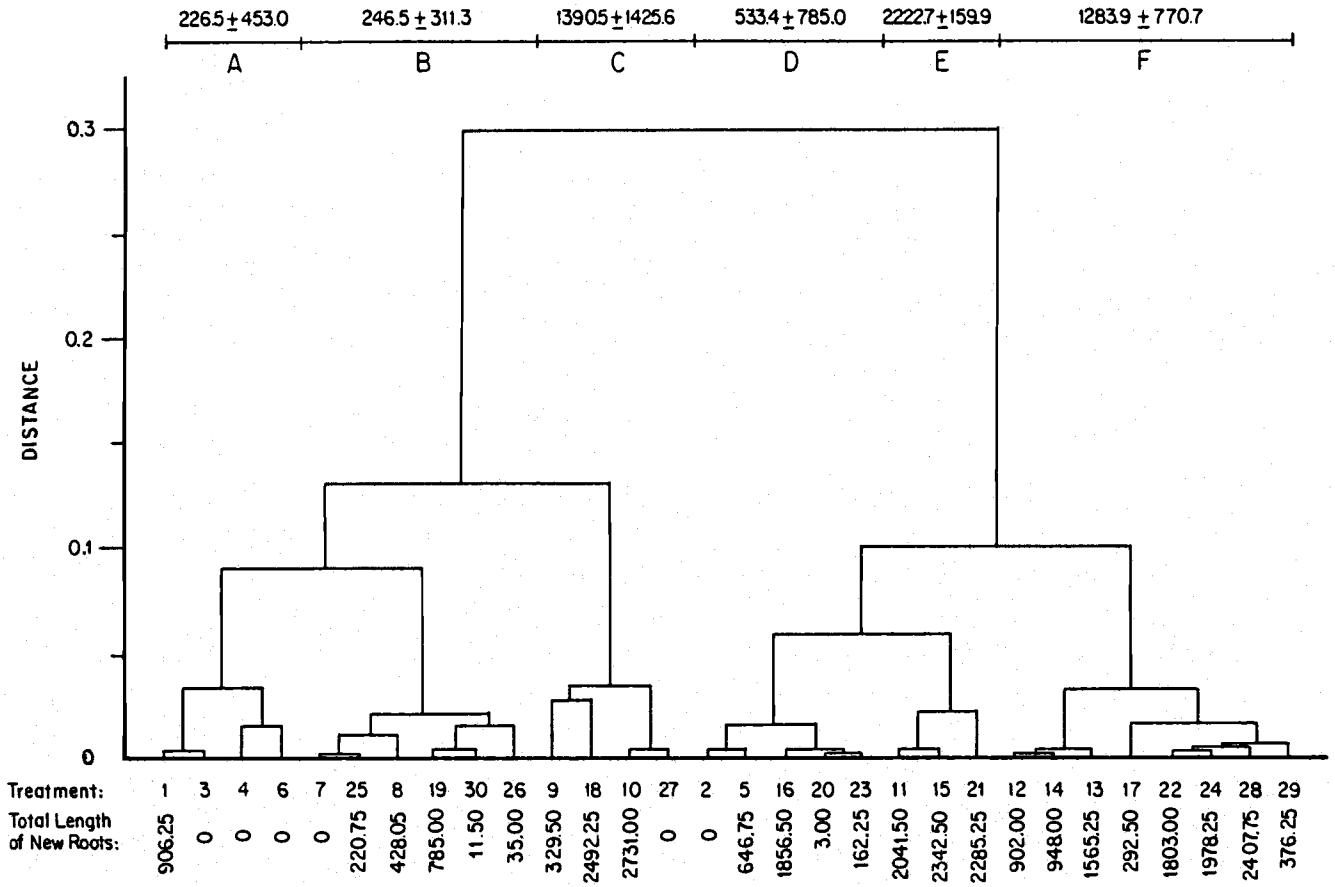

Figure 18. Dendrogram based on squared euclidean distance measures of thirty treatments with varying total length of new roots. See Table 1 for explanation of treatments.

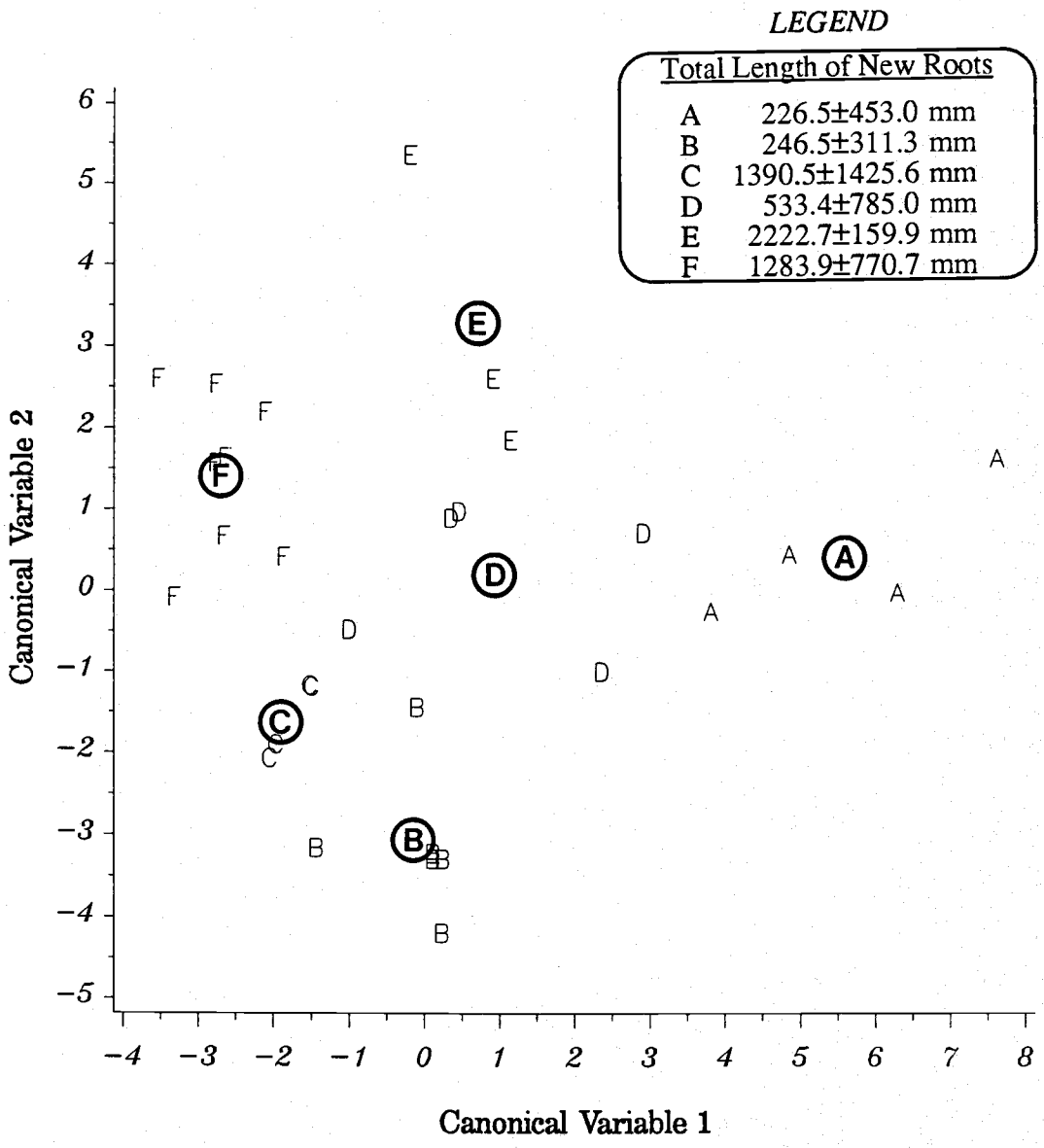

Figure 19. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak frequency data. Classes were determined from Figure 18. The first four variables accounted for 51.08% (0.0001), 29.37% (0.0001), 13.33% (0.0001), and 5.99% (0.0100) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the six groups are indicated.

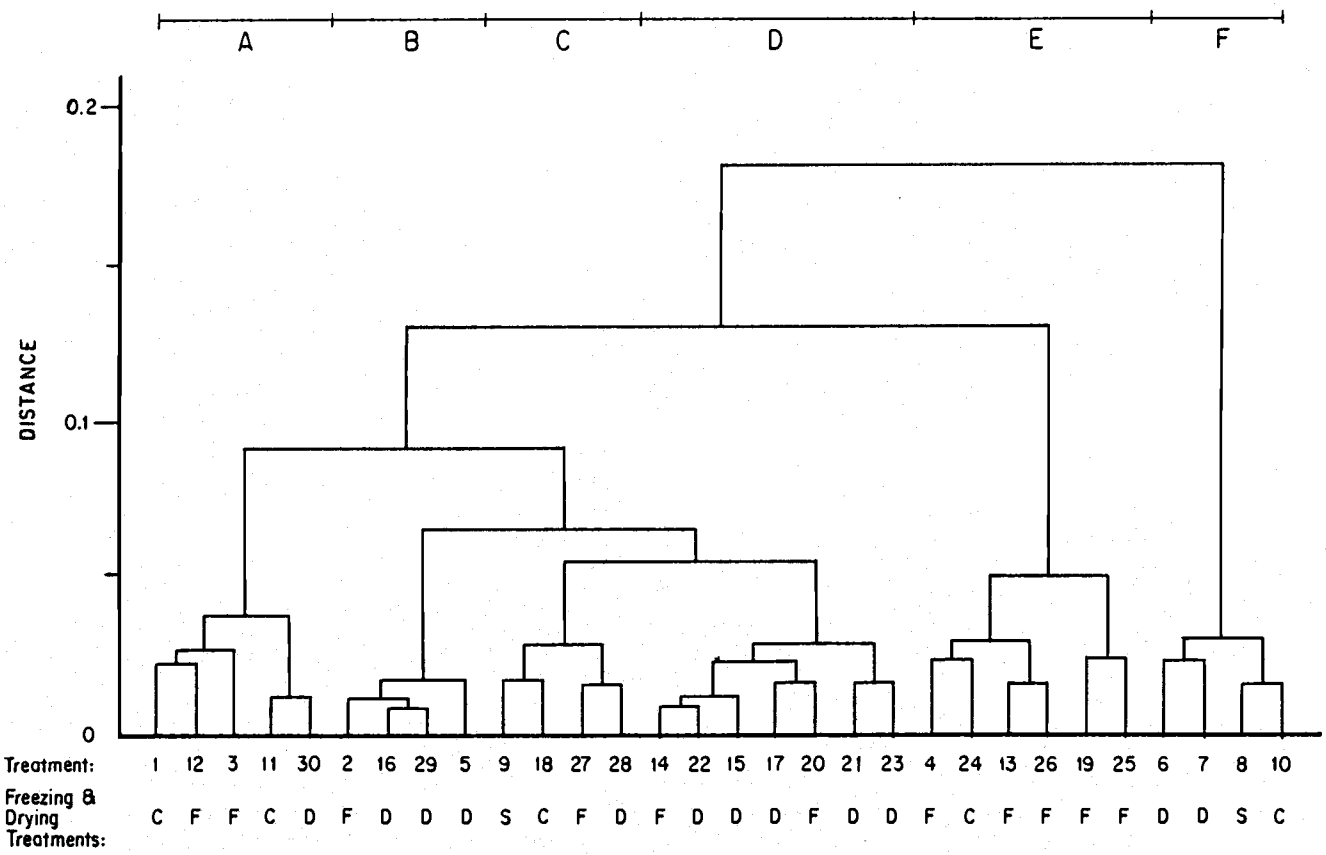

Figure 20. Dendrogram based on squared euclidean distance measures of thirty treatments exposed to freezing, drying, or cold storage conditions. See Table 1 for explanation of treatments.

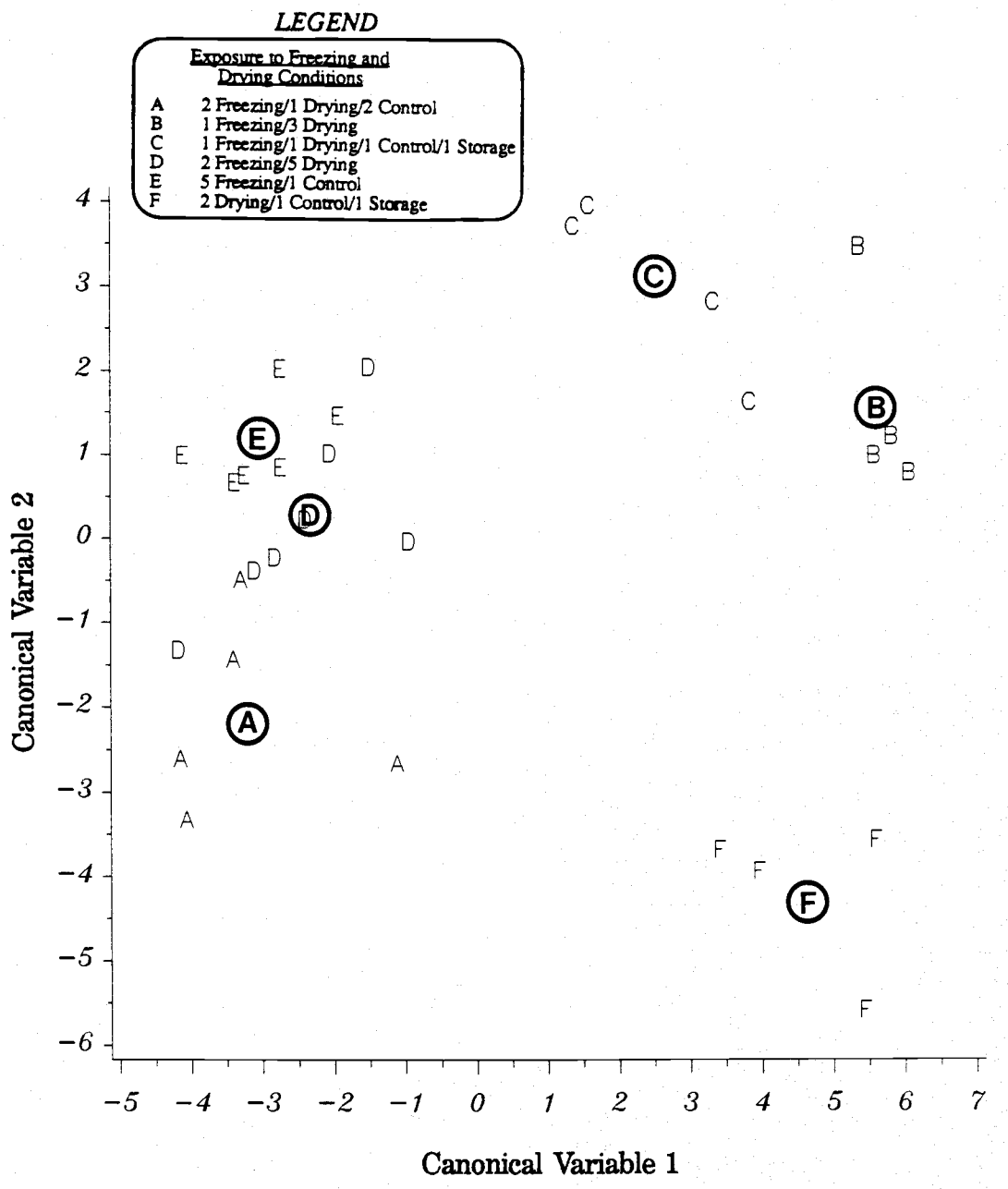

Figure 21. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak frequency data. Classes were determined from Figure 20. The first four variables accounted for 55.70% (0.0001), 21.18% (0.0001), 10.69% (0.0010), and 8.73% (0.0098) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the six groups are indicated.

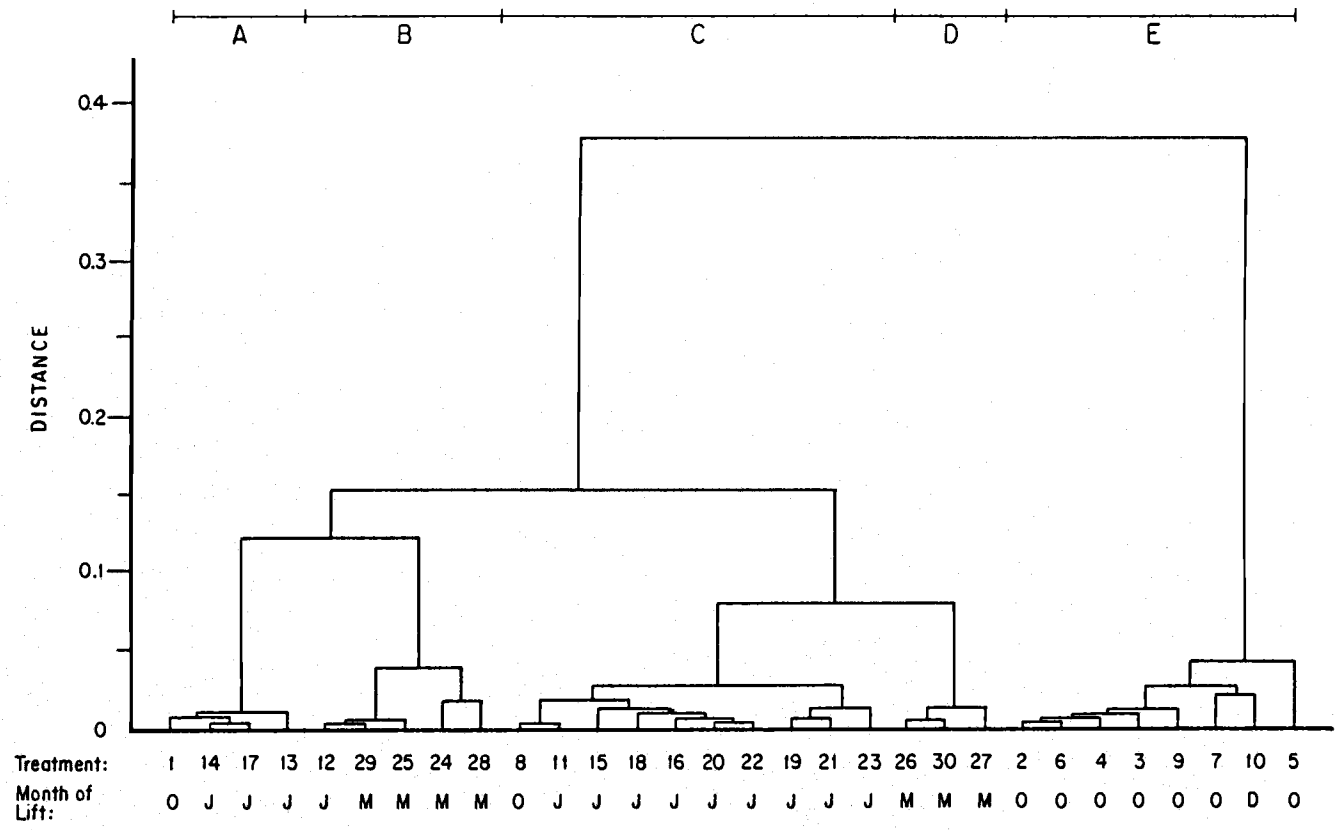

Figure 22. Dendrogram based on squared euclidean distance measures of thirty treatments lifted in varying months. See Table 1 for explanation of treatments.

October were mis-assigned. Treatment 1 was found in the January-lifted group A, and treatment 8 in the other January-lifted cluster, group C. The two groups of January lifted treatments (A and C) were combined on the canonical plot and were well separated from the other groups (Fig. 23). Treatments lifted in March were also found in two groups (B and D). The first canonical variable was critical in separating the treatments based on time of lift and explains 63% of the variation.

Summary

Overall, the cluster analysis approach was not effective in grouping seedlings based on a dependent variable. The exception is clustering treatments according to time of lift and less so according to growth room survival. Little overlap occurred between these clusters and the dendrogram results were consistent with the canonical plot groups. In many other cases, to varying degrees, dendrogram clusters were inconsistent with canonical plot groups. Well defined groups as shown in canonical plots indicate that 'true' groups are being formed (the Wilk's lambda statistic for all plots was highly significant ($p \leq 0.0001$)), but the accompanying large standard deviations and overlapping value ranges from the dendrograms suggest that the group similarity is based more on unknown factors that on the dependent variable. This is perhaps why R^2 values from stepwise multiple regression analyses are large even though clustering was so poor. The model equations are able to explain a majority of the variation between treatments yet the variation observed is due to more factors than only the respective dependent variables. One of the least successful clustering results was that predicting freezing and drying damage. However, the regression model had the largest R^2 value (0.94). This was possibly due to the inclusion of a large number of variables (13) relative to observations (30).

The number of peaks used to cluster treatments based on a dependent variable ranged from five to thirteen (Table 3). Several of the same peaks were found to be

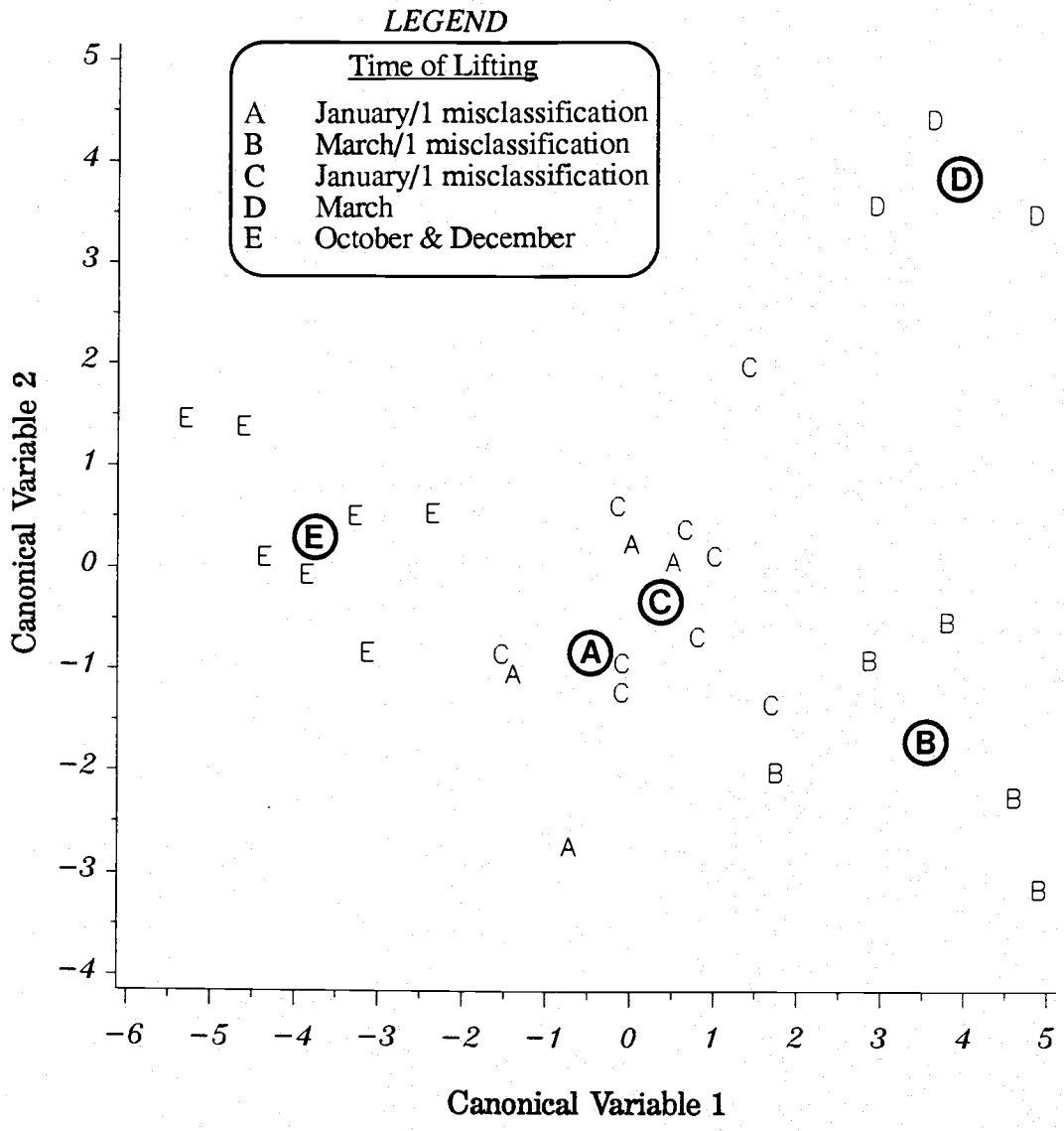

Figure 23. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak frequency data. Classes were determined from Figure 22. The first four variables accounted for 63.04% (0.0001), 17.73% (0.0001), 11.89% (0.0001), and 7.35% (0.0002) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the five groups are indicated.

involved with field survival, growth room survival, and RGP data. Peak 78 was included in all four clustering analyses, while peaks 46 and 79 also appeared important to both RGP and survival. Two peaks, 12 and 41, though important in RGP, apparently are not strongly involved in survival. Freezing and drying damage and date of lifting have no markers in common with any of the other dependent variables.

Because of the ineffectiveness of the cluster analysis approach to marker identification, a discriminant analysis approach, using *a priori* defined groups, was taken. This second approach was also used to confirm the selection of those biochemical markers identified using the cluster analysis approach.

DISCRIMINANT ANALYSIS APPROACH

Discriminant analysis techniques are used to describe or predict the behavior of a non-metric dependent variable (Hair et al. 1987; Afifi and Clark 1984). The common purpose of discriminant analysis procedures is to predict class membership on the basis of a set of independent variables (Afifi and Clark 1984; Verbyla 1986). However, these techniques are also used to identify the variables important to the classification process (Stiteler 1978). The objectives for applying discriminant analysis to marker detection were to determine if differences exist between *a priori* defined groups and to identify which variables contributed to making the classification.

The discriminant analysis approach to marker detection consisted of three techniques: stepwise discriminant analysis, the cross-validation option of discriminant analysis, and canonical discriminant analysis. Stepwise discriminant analysis selects the 'best' subset of independent variables to describe a discriminant function. Variables are entered into the function one at a time beginning with the single best discriminating variable. Each remaining variable is then paired with the initial variable. A second variable is chosen that, together with the first, most improves the discriminating

function. Subsequent variables are selected similarly. Previously selected variables may be removed from the function if other selected variables contain the same information about group differences (Hair et al. 1987).

A limitation of stepwise discriminant analysis is prediction bias. Verbyla (1986) states that bias "is likely to occur when a model contains many independent variables relative to sample size or when many different sets of independent variables are tested by a stepwise procedure". When there are many combinations of independent variables, it is likely that while one combination may fit the sample data well, it will predict poorly when applied to new data because of the prediction bias. Bias can be assessed by resampling procedures such as cross-validation. The cross-validation procedure excludes each observation singly from the model development and then tests the model on the excluded case. This is continued for each observation (Hair et al. 1987; Verbyla 1986). After observations have undergone cross-validation, canonical variables can be derived to visualize group dispersion. Canonical variables, as explained in the cluster approach section, are linear combinations of independent variables that, when plotted, illustrates the maximum possible separation among the groups (Hair et al. 1987).

The three multivariate statistical techniques were performed using the peak frequency and the mean peak area data sets for each dependent variable. Several *a priori* classifications of the dependent variables field survival, growth room survival, and RGP (number and total length of new roots) were also analyzed using both data sets. All stepwise discriminant analyses included between 25 and 30 independent variables into the function, always maximizing the R^2 value ($0.9998 \leq R^2 \leq 1.0000$). The variables included in the function in the first fifteen steps are presented for each analysis in Appendix D. This is approximately half the number of observations (n=30). It is likely that variables included in the function after the fifteenth step are accounting for group differences mainly because their number is approaching the number of observations.

The percent misclassification for each of the first fifteen steps was determined by cross-validation. The number of variables considered to contain adequate information about group differences was determined by F statistics (larger F-values indicate greater discriminating power), Wilk's lambda values (tests for equality of group means), and the number of misclassified observations in each step. For example, Table D1 in Appendix D is the summary results from stepwise discriminant analysis and cross-validation on the peak frequency data set using *a priori* field survival classifications of 0-40%, 41-80%, and 81-100% survival. The variables entered into the function in the first nine steps (those bracketed on the left side of the table) were determined to adequately explain group differences. The first nine variables all have a significant amount of discriminating power ($p \leq 0.0423$ for the F statistics). The Wilk's lambda value for the ninth step is also highly significant. Ten percent of the observations in the ninth step were misclassified. While the first nine variables included in the model are likely involved in explaining group differences, the variables included in steps 10 through 15 are probably not as involved in the discriminating power of the function.

It should be stressed that the number of variables chosen to explain group differences for each analysis is very conservative. In all analyses it is possible to get 'perfect' discrimination between groups, as observed when plotting canonical variables, if using a sufficient number of independent variables (in these analyses it was 21 or less variables). Because it is not possible to determine when prediction bias begins to play a major role in classification, the smallest number of variables that appeared to adequately explain group differences were chosen. Therefore, separation among groups as evidenced by plots of canonical variables depicts the poorest discrimination between groups that can be attained from the variables. Group separation is likely far better and class differences more legitimately explained with the inclusion of several additional variables.

Markers of Field Survival

Two *a priori* group classifications for field survival were analyzed using the mean peak area and peak frequency data sets. The first classification consisted of three groups (0-40%, 41-80%, and 81-100% field survival). Nine peak frequency variables (Table D1, Appendix D), that resulted in 10% misclassification of observations, were used to derive canonical variables which were subsequently plotted (Fig. 24). The first canonical variable explains almost all of the total dispersion (94%) and is highly significant ($p \leq 0.0001$).

When analyzing the classification with peak area data, 13% of the observations were misclassified using five peak area variables (Table D2, Appendix D). These variables were used in canonical discriminant analysis and resulting canonical variables plotted (Fig. 25). The first two canonical variables were significant ($p<0.013$) with 88% of the total variation explained by canonical variable 1. The better separation of groups seen in Figure 24 may be due to the fact that more variables were used to create Figure 24 than Figure 25 (9 and 5 variables, respectively).

Though good discrimination was obtained for this classification using both data sets, it may be advantageous for classes to contain smaller survival ranges for better predictive power. Therefore, the second classification analyzed consisted of five groups (0%, 1-20%, 21-50%, 51-80%, and 81-100% field survival). Using peak frequencies, six variables were determined to adequately explain group differences (Table D3, Appendix D). Dispersion explained by the first two derived canonical variables was 93% (Fig. 26). Groups D and B are separated from the other groups on the first canonical variable while groups E and C are separated from group A on the second canonical variable.

Five variables were determined to contain sufficient information to separate groups (Table D4, Appendix D) when performing the analysis with peak area data.

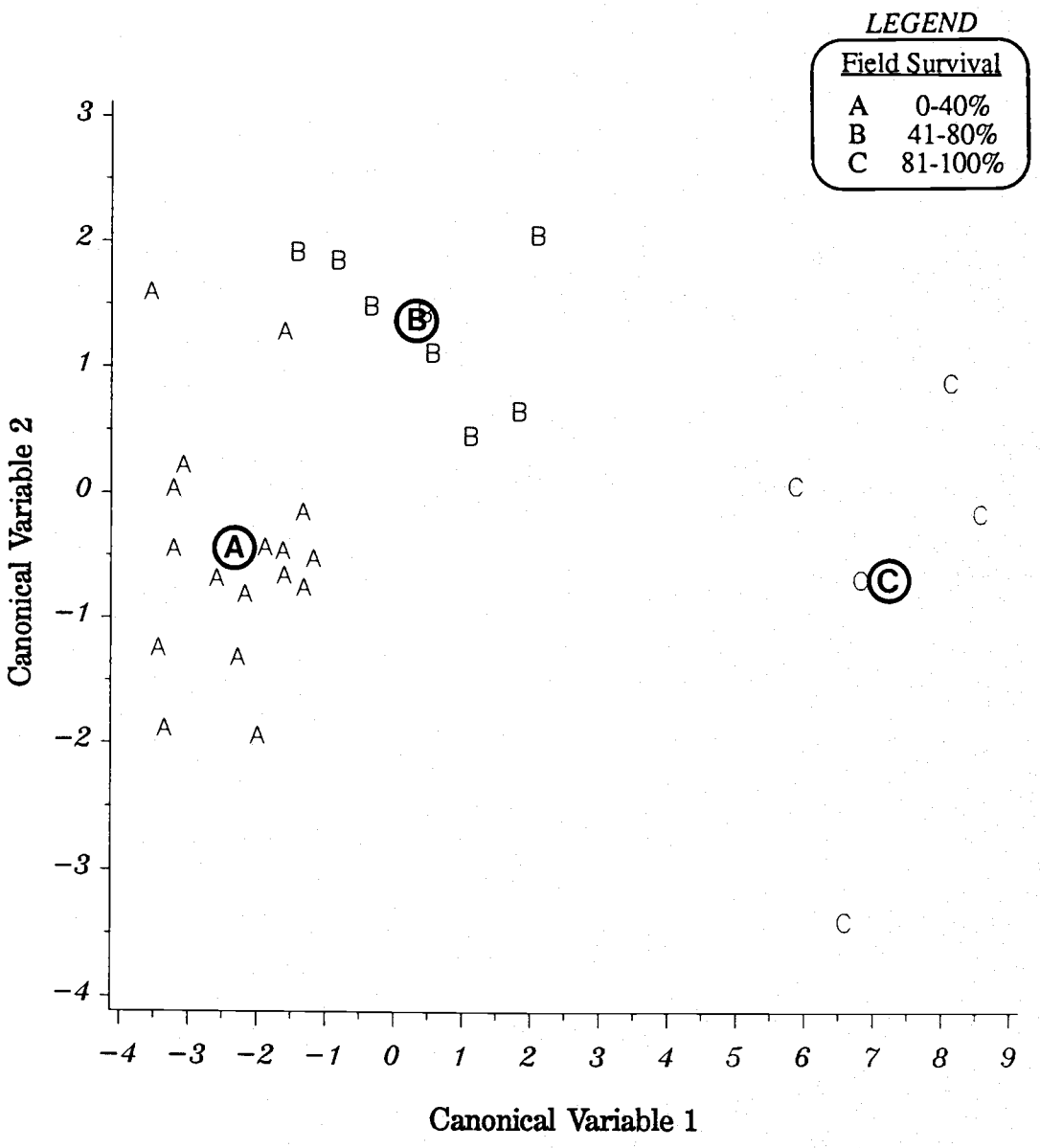

Figure 24. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak frequency data. Groups were based on one-year field survival. The first canonical variable accounted for 94.32% (0.0001) of the total dispersion. P value is given in parentheses. Position of the mean values of the canonical variables for each of the three groups are indicated.

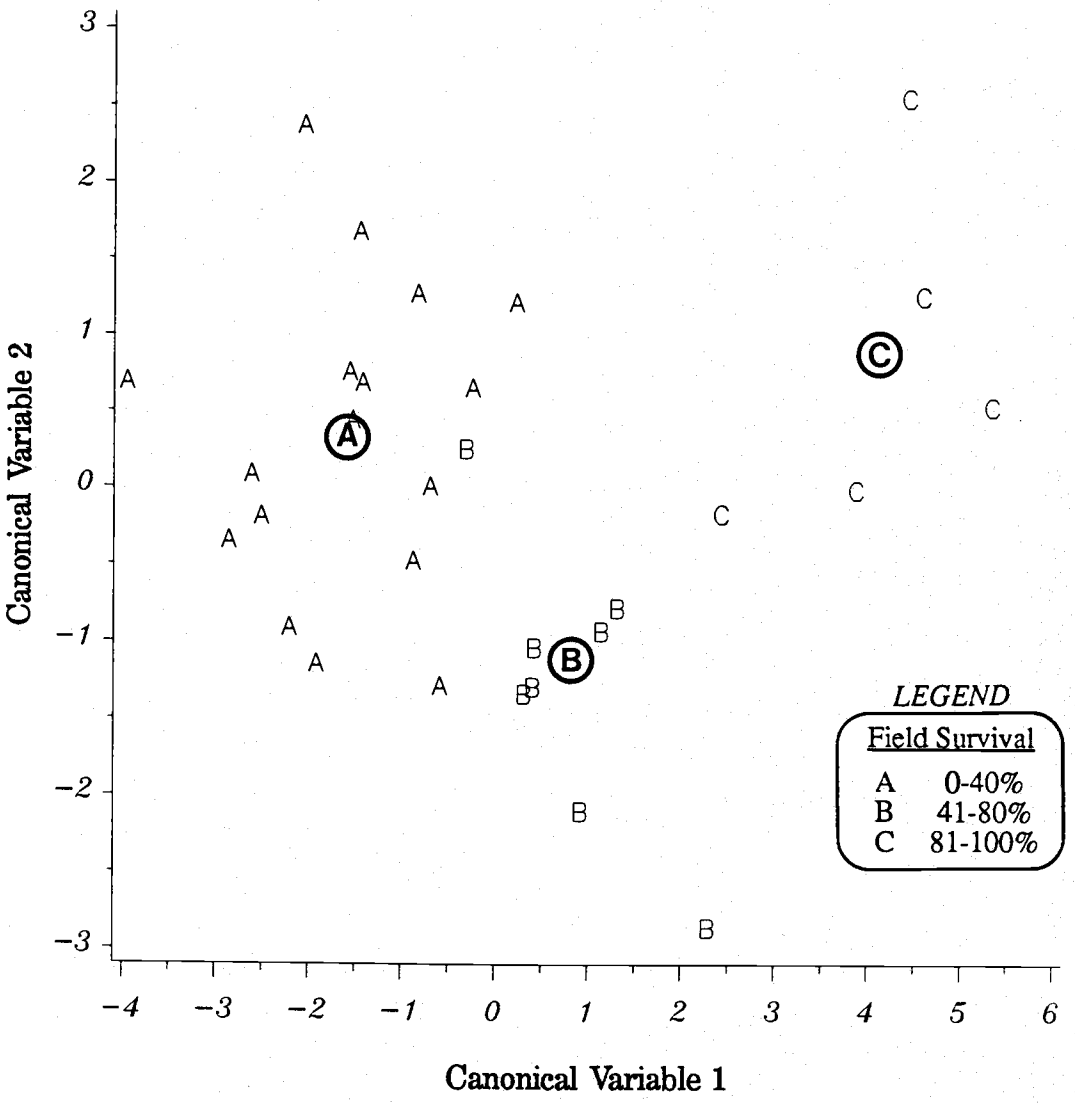

Figure 25. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak area data. Groups were based on one-year field survival. The first two canonical variables accounted for 88.14% (0.0001) and 11.86% (0.0126) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the three groups are indicated.

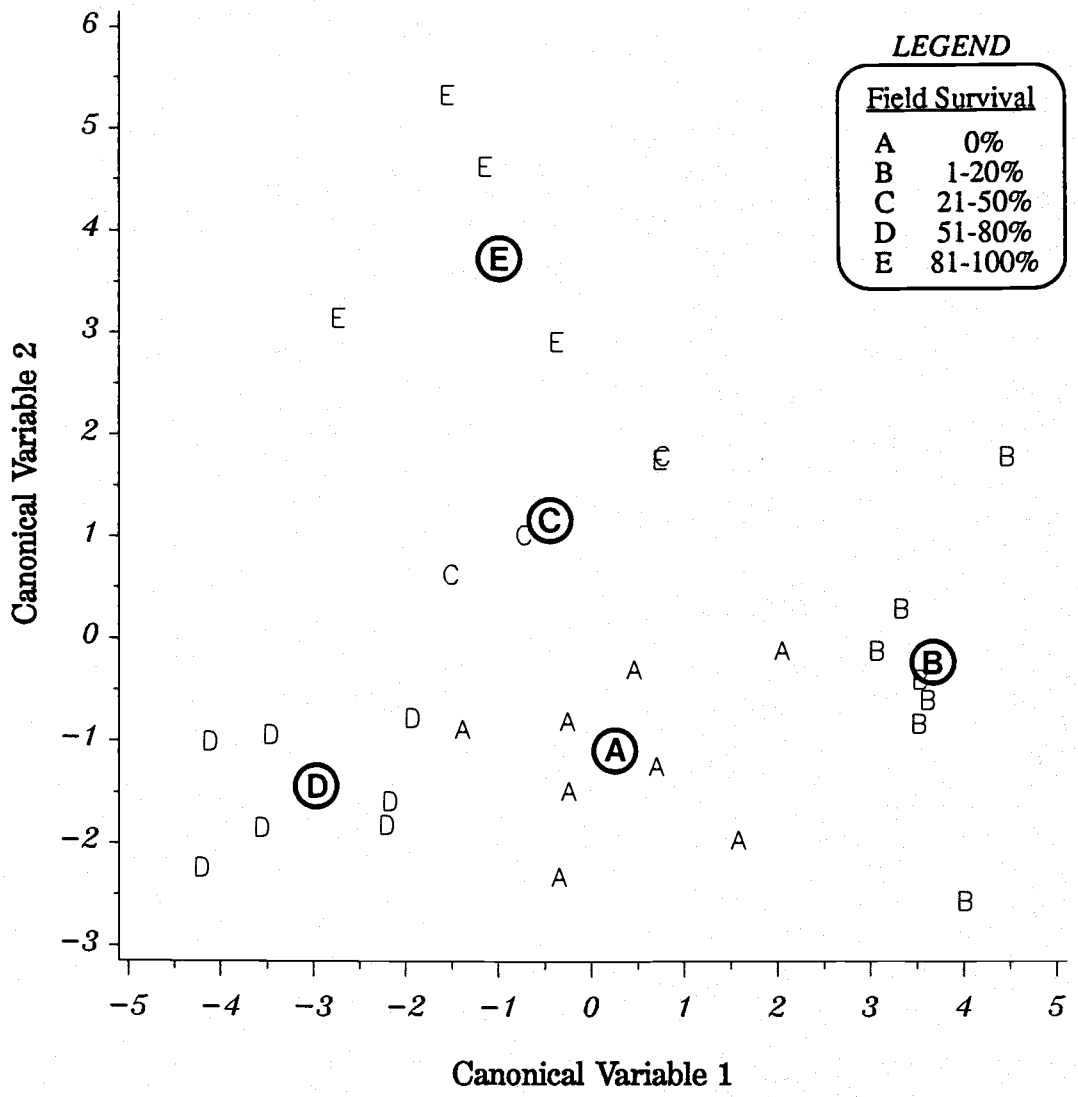

Figure 26. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak frequency data. Groups were based on one-year field survival. The first two canonical variables accounted for 59.35% (0.0001) and 33.17% (0.0004) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the five groups are indicated.

However, when plotting canonical variables, three groups (A, B, and D) remained unseparated (Fig. 27). Additional variables in the function are needed to better discriminate groups A, B, and D.

Markers of field survival were identified and adequate discriminating functions derived using peak frequency and peak area variables. Peak 58 is important to both *a priori* classifications using either data set. Other markers in common between the two data sets are peaks 20 and 228 for the three group classification and peaks 12, 48, 150, and 197 for the five group classification. When using the peak frequency data set, markers shared by both classifications are peaks 48, 58, and 197. Markers shared by both classifications when using peak area variables are peaks 58 and 113. Peaks 48, 58, and 197 appear to be the most informative of the field survival markers.

Markers of Growth Room Survival

The most successful *a priori* classification of growth room survival consisted of four groups (0-29%, 30-59%, 60-89%, and 90-100% growth room survival). Analysis with both data sets resulted in 'good' group separation. Seven variables were chosen from the stepwise discriminant analysis using peak frequencies (Table D5, Appendix D). Twenty-three percent of the observations were misclassified when samples were cross-validated. Canonical variables were plotted (Fig. 28) and 92% of the total dispersion explained. A similar plot (Fig. 29) was obtained when plotting the canonical variables derived from five peak area variables (Table D6, Appendix D). Twenty-seven percent of the observations were misclassified and the first two canonical variables explained 98% of the total dispersion. Over 90% of the variation is explained in the first canonical variable alone. There is some overlap between groups C (60-89% survival) and D (90-100% survival) on the second canonical variable. Additional variables included in the function would separate these two groups. It is also possible that the break between these groups does not belong between 89 and 90%.

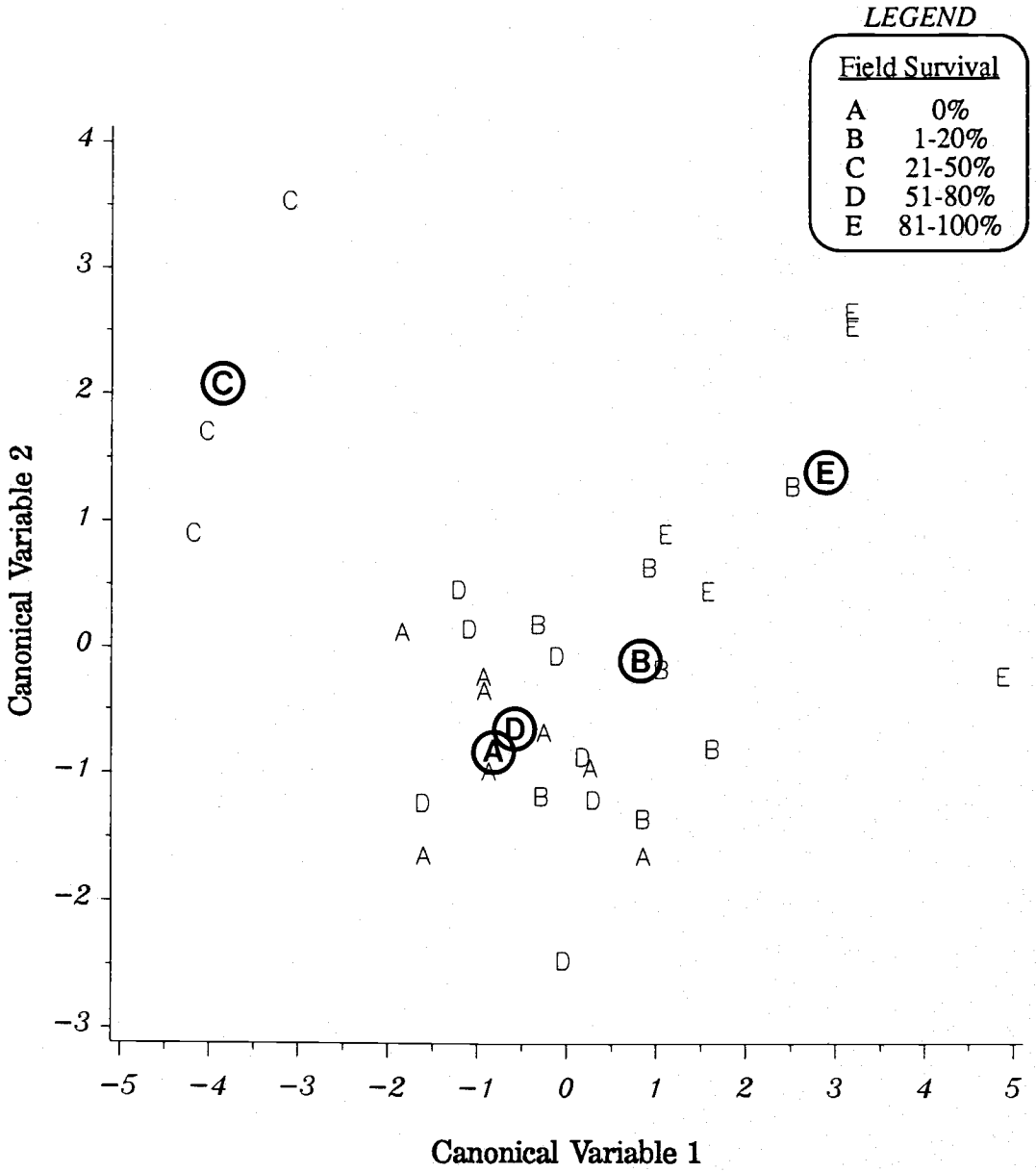

Figure 27. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak area data. Groups were based on one-year field survival. The first three canonical variables accounted for 66.53% (0.0001), 21.38% (0.0006), and 10.85% (0.0439) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the five groups are indicated.

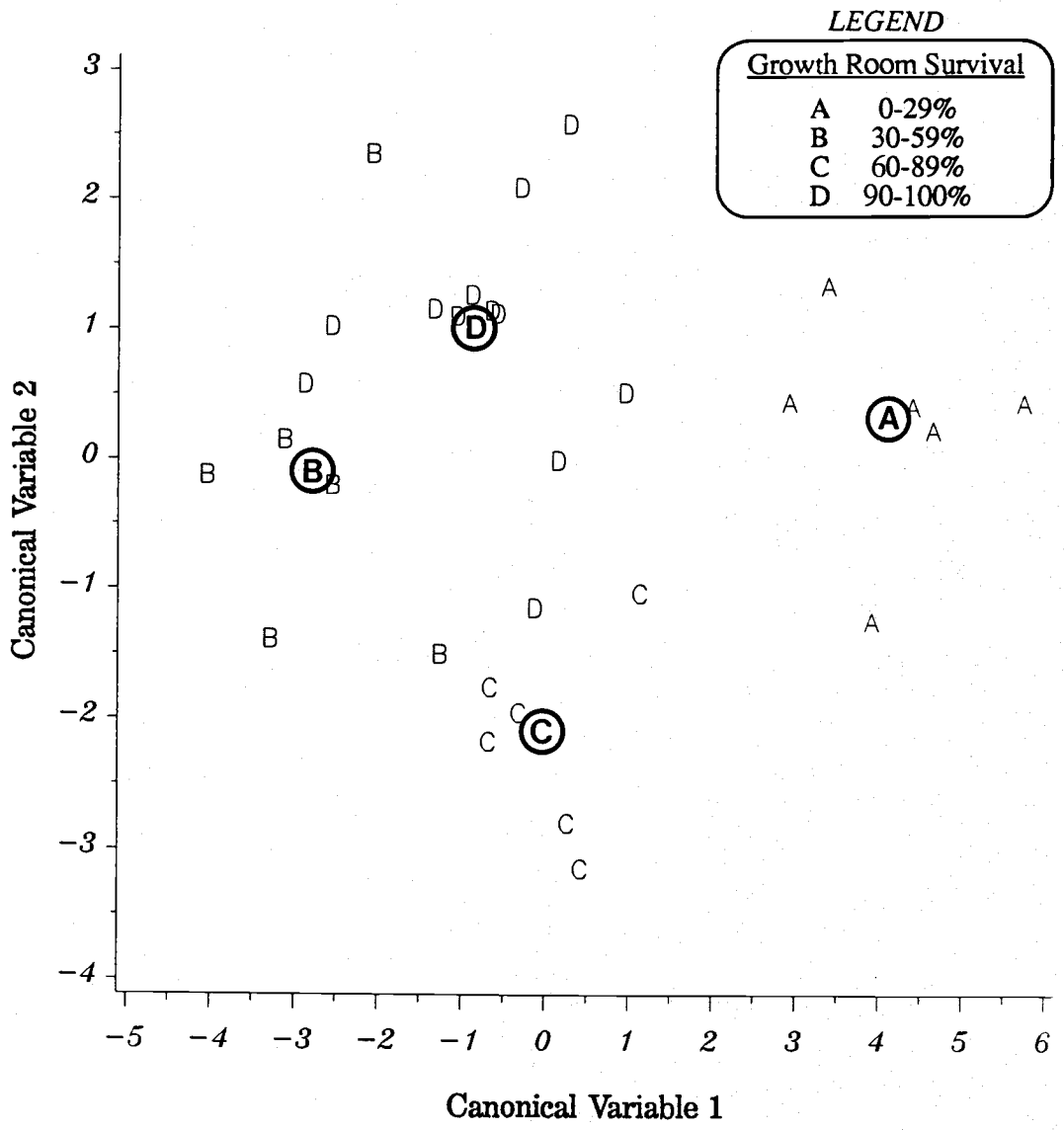

Figure 28. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak frequency data. Groups were based on six-week growth room survival. The first three canonical variables accounted for 73.56% (0.0001), 18.51% (0.0010), and 7.93% (0.0403) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the four groups are indicated.

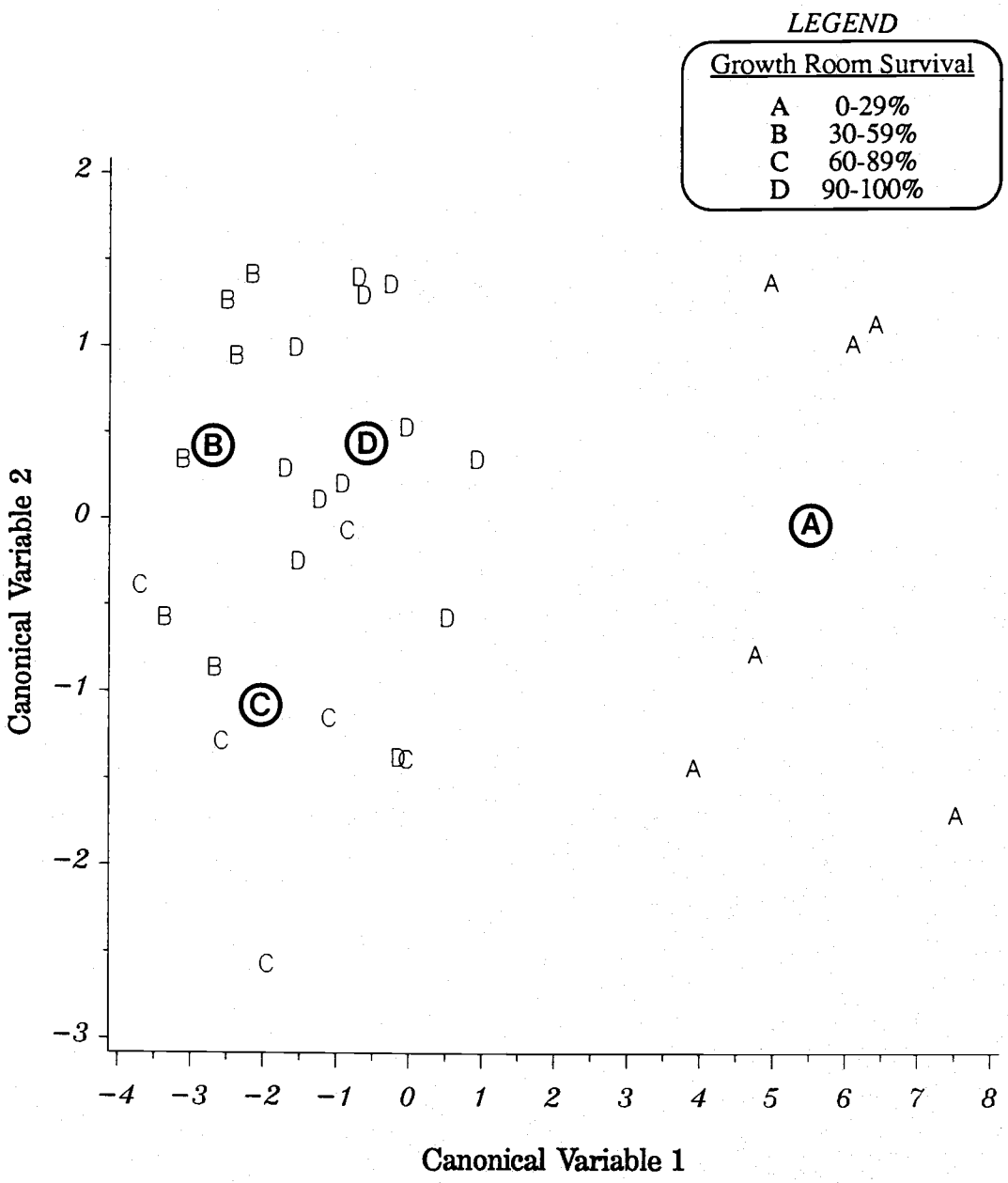

Figure 29. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak area data. Groups were based on six-week growth room survival. The first canonical variable accounted for 93.70% (0.0001) of the total dispersion, respectively. P value is given in parentheses. Position of the mean values of the canonical variables for each of the four groups are indicated.

Peak 133 was the only shared marker between the two data sets. It was also the first variable included in each function.

Markers of Root Growth Potential

Eight peak frequency variables were determined to adequately explain differences among *a priori* groups based on number of new roots (0-39, 40-99, 100-174, and 175-200 new roots) (Table D7, Appendix D). Five variables were identified using peak area data (Table D8, Appendix D). Peak 131 was found included in both functions. The first two canonical variables derived from peak frequency and peak area data explained 97% (Fig. 30) and 96% (Fig. 31) of the total dispersion, respectively. Canonical variable 1 explained the majority of the variation in each figure though group separation among A, B, and C observations was poor. In Figure 30, group A (0-39 new roots) slightly overlaps group B (40-99 new roots) and borders on group C (100-174 new roots). Groups A and C overlap in Figure 31. Again, additional variables will result in a better separation or perhaps classification boundaries are not appropriate.

A priori classification of total length of new roots consisted of five groups (0-50, 51-500, 501-1000, 1001-2000, and 2001-3000 mm of new root growth). Five variables were chosen from stepwise discriminant output using both peak frequencies and peak areas (Table D9 and Table D10, respectively, Appendix D). Misclassification of observations in each was high-- 37% using frequency data, 40% using peak area data. Plots of canonical variables show a slightly better separation of groups when using peak area data (94% of the total variation explained) (Fig. 32) than when using peak frequencies (88.86% of the total variation explained) (Fig. 33). The first canonical variable in Figure 32 separates groups C (0-50mm) and D (2001-3000mm). Group B (51-500mm), A (501-1000mm), and E (1001-2000mm) show some overlap. The same overlap is observed in Figure 33. Again, either additional variables need to be included

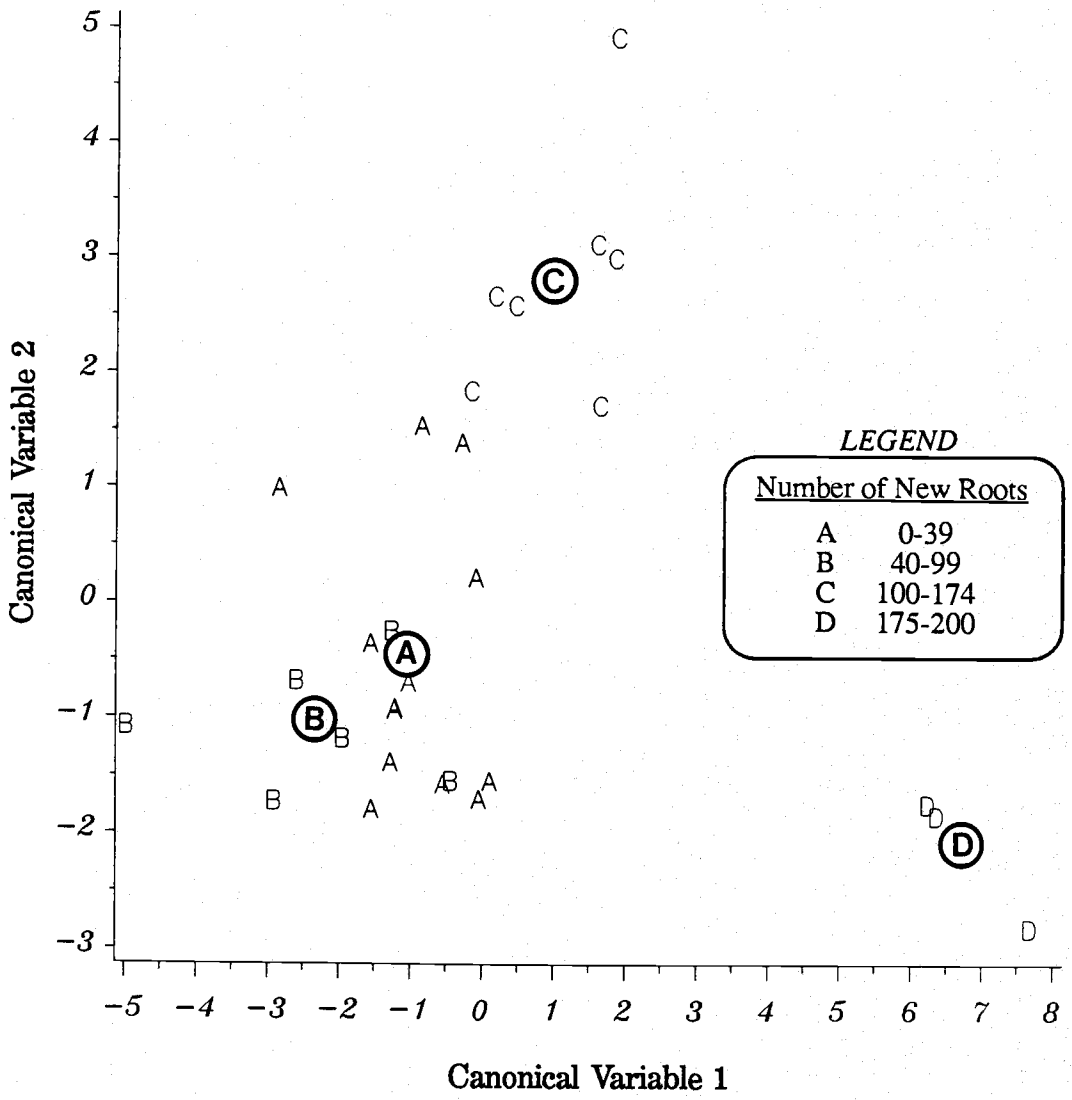

Figure 30. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak frequency data. Groups were based on the average number of new roots per treatment. The first two canonical variables accounted for 68.39% (0.0001) and 28.88% (0.0005) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the four groups are indicated.

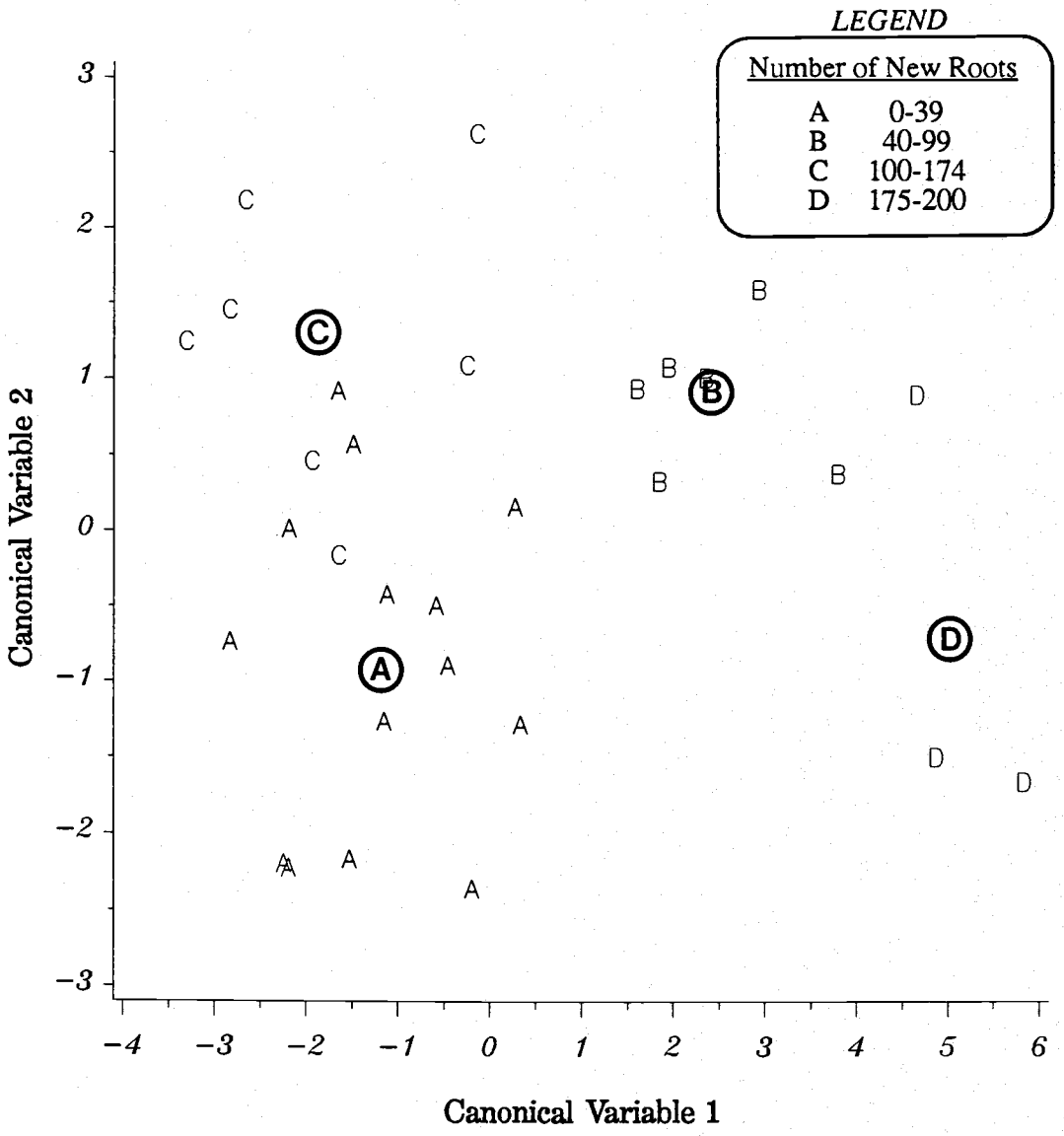

Figure 31. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak area data. Groups were based on the average number of new roots per treatment. The first two canonical variables accounted for 81.23% (0.0001) and 14.76% (0.0018) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the four groups are indicated.

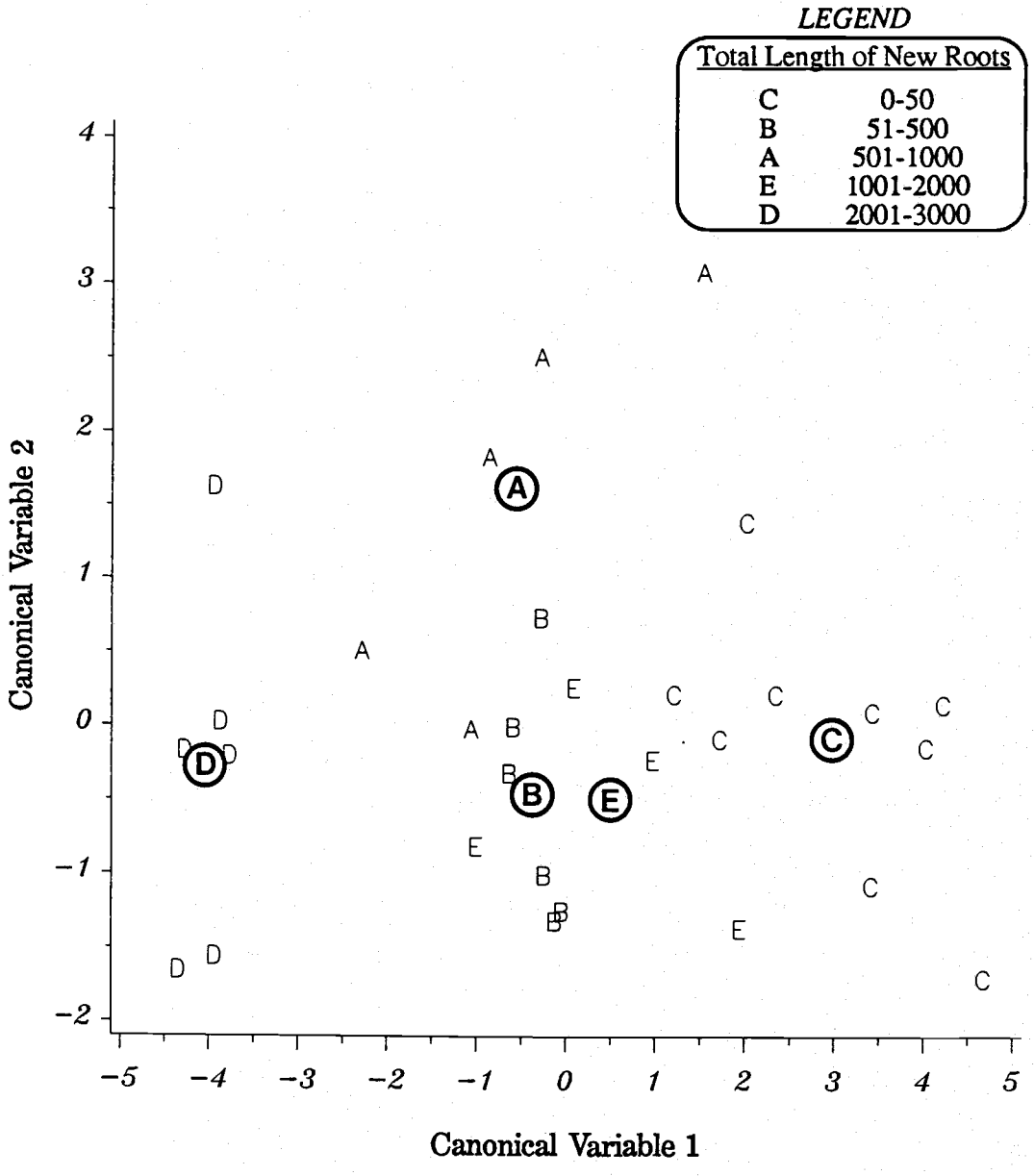

Figure 32. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak area data. Groups were based on the average total length of new roots per treatment. The first two canonical variables accounted for 86.28% (0.0001) and 7.67% (0.0325) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the five groups are indicated.

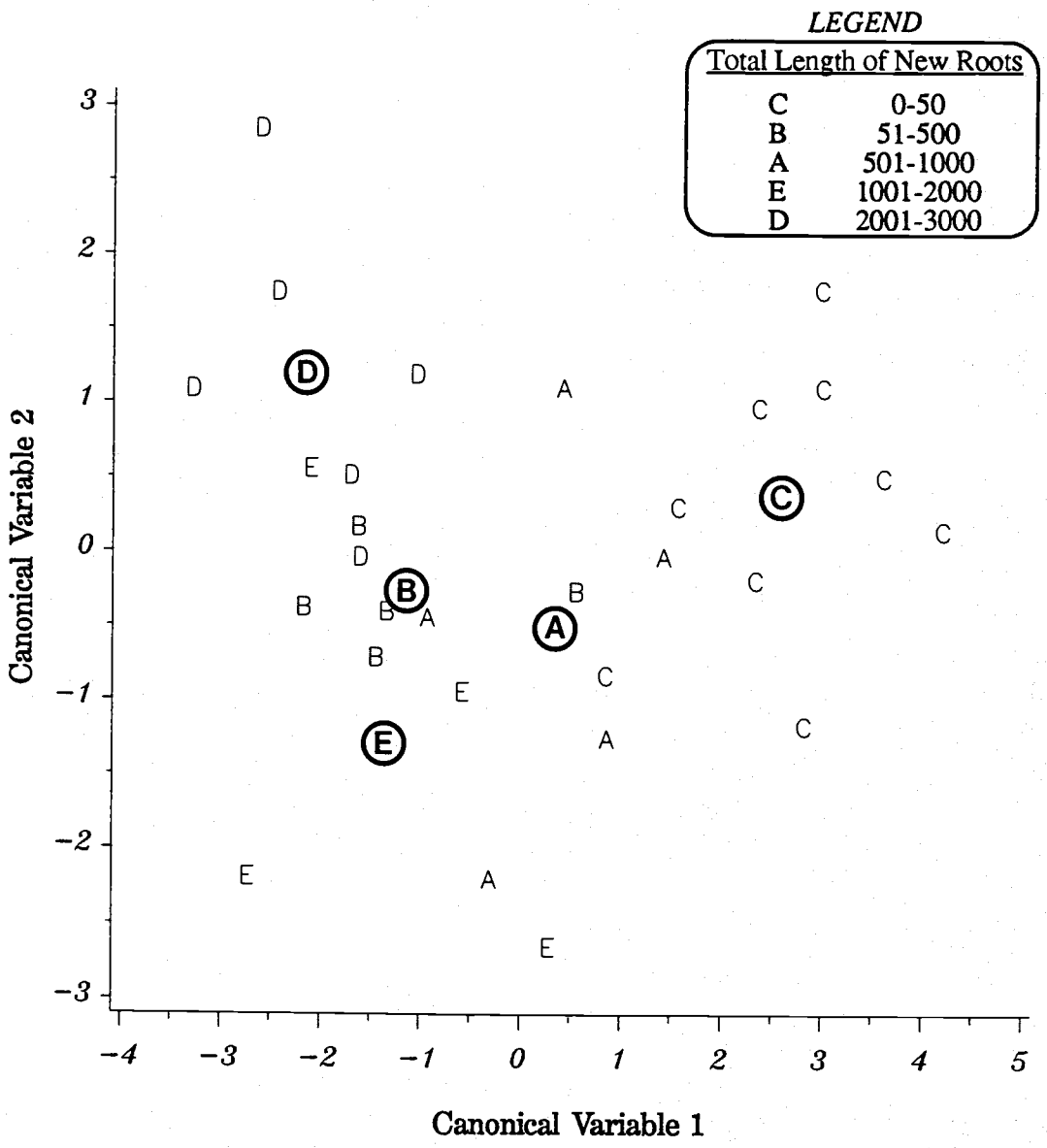

Figure 33. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak frequency data. Groups were based on the average total length of new roots per treatment. The first three canonical variables accounted for 75.23% (0.0001), 13.63% (0.0098), and 8.07% (0.0473) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the five groups are indicated.

in the function or current classification boundaries between groups A, B, and E are not appropriate.

Four markers of total length of new roots are shared between both data sets--peaks 78, 89, 133, and 205. Peak 78 was also a peak frequency marker for number of new roots while peaks 89 and 133 were peak area markers for number of new roots.

Markers of Freezing and Drying Damage

Treatments were classified into four groups: control, cold storage, exposure to freezing temperatures, and exposure to drying conditions. Six peak frequency variables were determined to adequately explain group differences (Table D11, Appendix D). The plot of the first two canonical variables derived from canonical discriminant analysis (Fig. 34) show a clear separation of the storage group (D) on the first canonical variable. Control treatments (C) are separated on the second canonical variable. The first two canonical variables were unable to separate the freezing and drying treatments from each other.

When analyzing the peak area data, six variables (Table D12, Appendix D) were used to derive canonical variables. When plotted (Fig. 35), the canonical variables separate groups in much the same pattern an in Figure 34. However, there is a slightly better separation of freezing and drying groups when using peak area data. Both data sets share peak 27 as a marker of freezing and drying exposure.

Markers of Lift Date

Treatments were classified into groups based on month of lift (October, December, January, and March). Using both peak frequency and peak area data sets, a minimum number of variables were needed to explain group differences well. Four peak frequency and peak area variables (Tables D13 and D14, respectively, Appendix D) were used to derive canonical variables that were subsequently plotted. Peak frequency data were analyzed, canonical variables derived, and the second and third

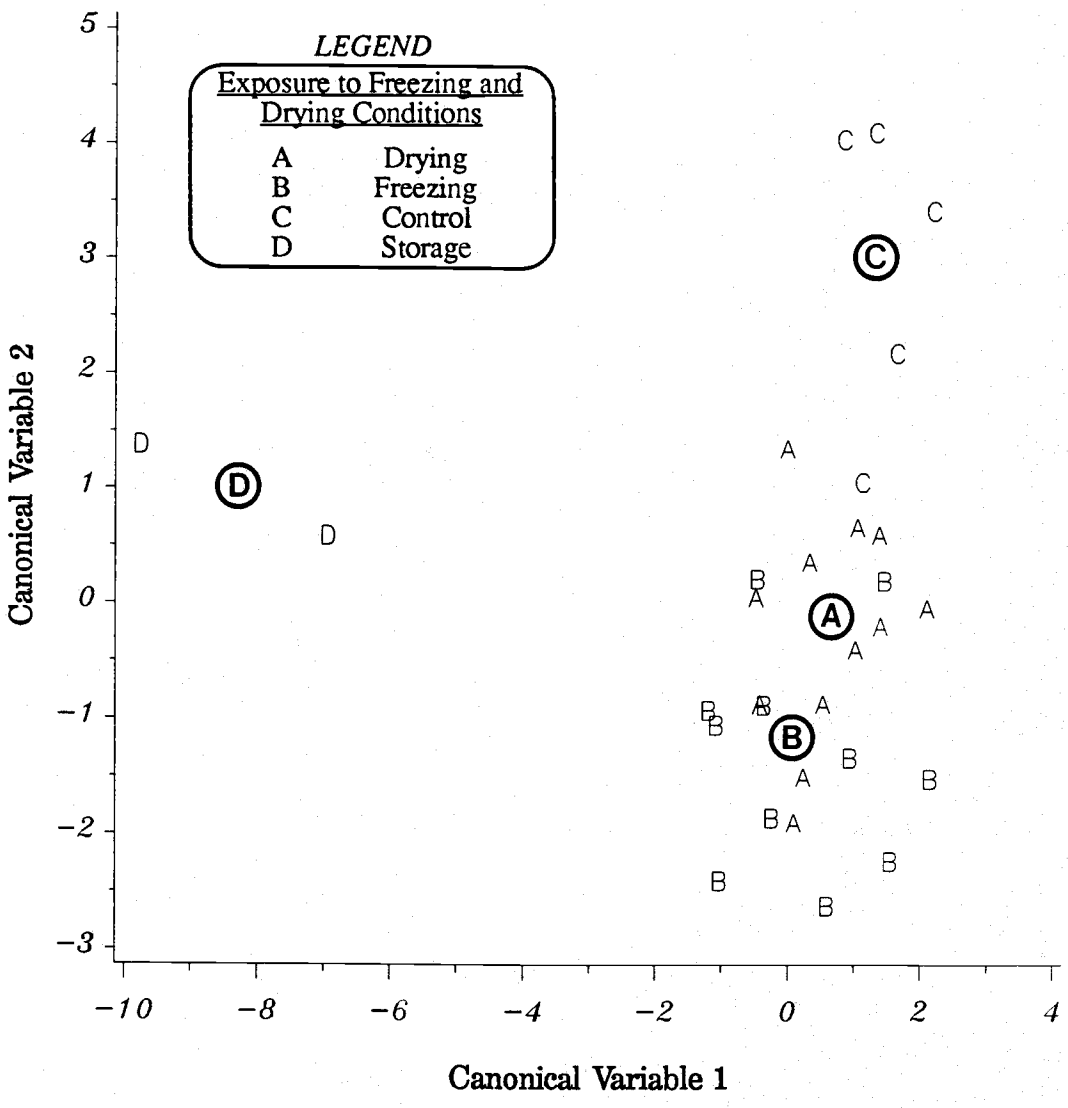

Figure 34. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak frequency data. Groups were based on treatment exposure to freezing and drying conditions. The first three canonical variables accounted for 66.07% (0.0001), 27.92% (0.0001), and 6.01% (0.0351) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the four groups are indicated.

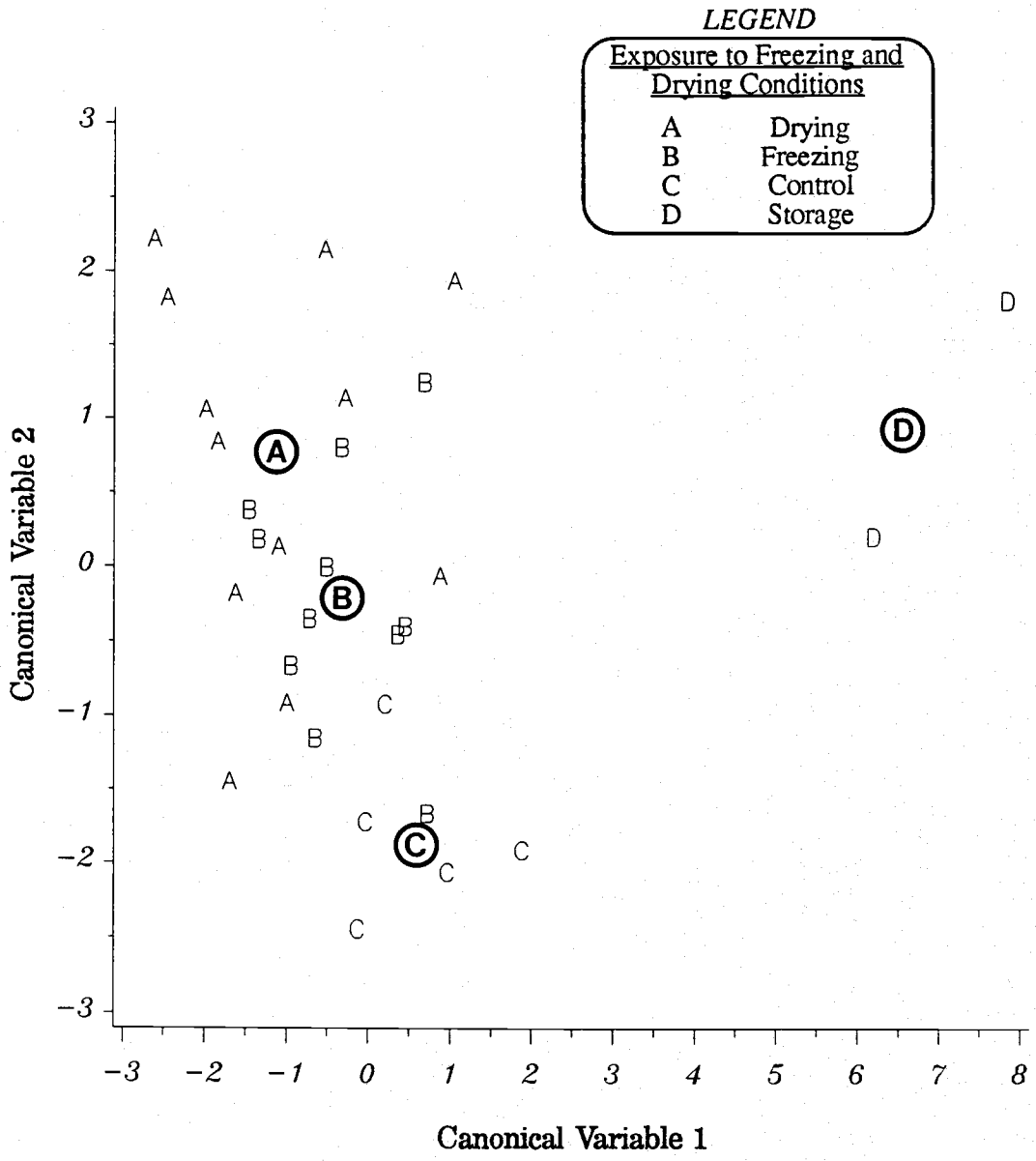

Figure 35. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak area data. Groups were based on treatment exposure to freezing and drying conditions. The first two canonical variables accounted for 76.50% (0.0001) and 16.62% (0.0071) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the four groups are indicated.

canonical variables were plotted (Fig. 36). The groups were well separated with the December-lifted treatment being included with treatments lifted in January.

The first two canonical variables derived using the peak area data explained 98% of the total dispersion and, when plotted, also showed excellent group separation (Fig. 37). The December-lifted treatment is positioned separately from the other groups.

The majority of group differences in both figures were explained by the first canonical variable. The data sets share peaks 20, 144, and 192 as markers of lift date.

Summary

The discriminant analysis approach was effective in determining if differences existed between *a priori* defined groups and in identifying variables important to the classifications. Markers for all dependent variables were identified using both the peak frequency and peak area data sets. Different combinations of markers were obtained for a given dependent variable for each data set used. It should be noted, however, that the data sets do share some peaks as markers for a given dependent variable. It is likely that different combinations of variables were selected when using peak frequency and area variables because the two data sets do not contain the same number of variables. In the peak frequency data set, peaks that appeared in all 300 samples were dropped from the analysis. It is possible that the area values of these peaks (which are still included in the peak area data set) contain some discriminating information and are, therefore, included as markers when using peak area data. When using peak frequency data this information is lost, and to effectively discriminate groups other peaks must be added to the function that explain the same information.

The number of peaks used to explain group differences for the dependent variables ranged from four to nine (see Appendix D). Several markers (peaks 77, 114, 133, 155, 177, 208, and 230) were found in common between field and growth room survival. RGP, both number of new roots and total length of new roots, appears to be

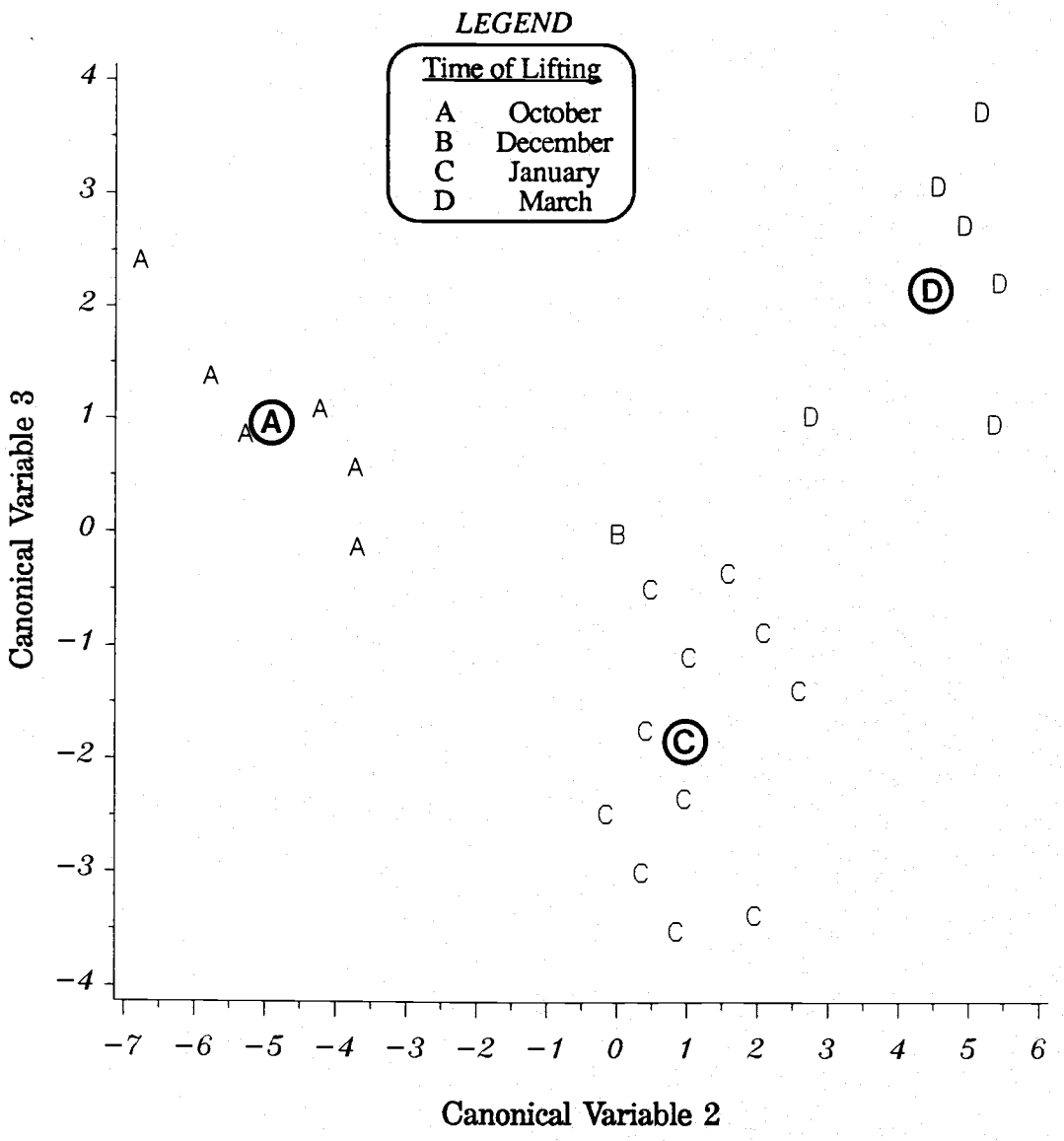

Figure 36. Plot of treatment means on the second and third canonical variables obtained by canonical discriminant analysis of peak frequency data. Groups were based on the time seedlings were lifted. The first two canonical variables accounted for 81.93% (0.0001) and 18.07% (0.0001) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the four groups are indicated.

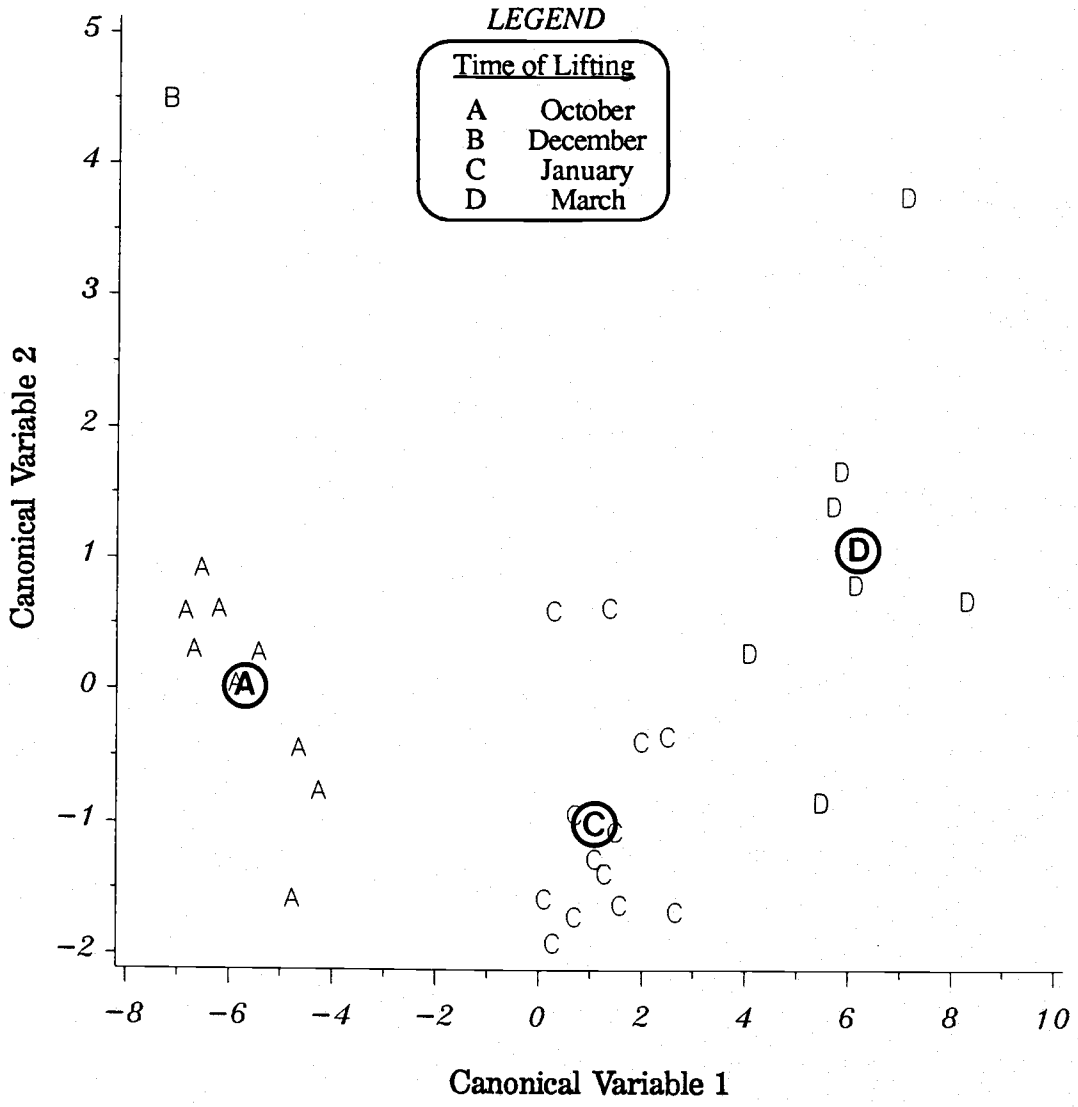

Figure 37. Plot of treatment means on the first two canonical variables obtained by canonical discriminant analysis of peak area data. Groups were based on the time seedlings were lifted. The first three canonical variables accounted for 92.01% (0.0001), 6.08% (0.0001), and 1.92% (0.0063) of the total dispersion, respectively. P values are given in parentheses. Position of the mean values of the canonical variables for each of the four groups are indicated.

strongly related to field and growth room survival. Eighteen markers are shared between RGP and survival while those most important to the relationship appear to be peaks 48, 77, 78, 133, and 197.

COMPARISON OF MULTIVARIATE STATISTICAL APPROACHES

The discriminant analysis approach was more effective at identifying biochemical markers than the cluster analysis approach. The major limiting factor of the cluster approach was that the similarity between treatments was caused by factors other than a single dependent variable such as field survival. Lift had an extremely strong influence on clustering, even when selected variables that were correlated to other factors (survival or RGP) were used in the analyses.

An example of an unknown factor that could account for similarity among treatments of varying field survival is time of outplanting. In the fall of 1985, during the field planting year, there was an extremely hard frost that likely damaged or even killed seedlings that had been recently outplanted. When treatments were clustered on field survival it is quite possible that the similarity observed was instead caused by the early planting. This could explain why the December-lifted treatment clustered with the treatments lifted in October since seedlings lifted in both months endured the same harsh frost conditions in the field.

Limitations of the discriminant analysis approach included the need for *a priori* classifications when group boundaries were unknown and the use of a large number of independent variables (approximately 230) relative to the number of observations (30). This resulted in a large amount of prediction bias.

Several markers identified with the cluster analysis approach were also identified using the discriminant analysis approach. Peak 85 was found to be a marker for field survival while peaks 35, 78, and 155 were identified as markers for growth room

survival using both approaches. Two peaks, 12 and 78, were identified using both approaches as markers for number of new roots and peak 78 was also found to be a marker of total length of new roots. Both procedures also identified peaks 27, 29, 45, and 192 as markers of freezing and drying exposure and peaks 20 and 111 as markers of time of lifting.

Peaks 78 and 133 appear to be particularly important to RGP (number and total length of new roots), growth room survival, and field survival. The importance of peak 78 to field survival was indicated strongly only by cluster analysis, while its contribution to the other three dependent variables (growth room survival and number and total length of new roots) was shown using both approaches. Peak 133 was repeatedly shown to be important to explaining treatment differences based on root growth potential and survival when using discriminant analysis but not cluster analysis.

The identification of the most important markers for each dependent variable was determined using results from the cluster analysis approach using peak frequency data and the discriminant analysis procedure using both data sets (Table 4). These are the peaks that should be chemically characterized to obtain information about the changes in physiological and biochemical processes in seedlings of varying quality.

Table 4. The most independently informative biochemical markers of seedling survival, RGP, exposure to freezing and drying conditions, and time of lifting.

DEPENDENT VARIABLES	HPLC PEAKS
Field Survival	78, 58, 85, 12, 197
Growth Room Survival	78, 133, 35, 155
Number of New Roots	78, 133, 89, 12
Total Length of New Roots	78, 133, 89, 205
Exposure To Freezing and Drying Conditions	27, 29, 45, 192
Time of Lifting	20, 111, 144, 192

CONCLUSION

Data from high performance liquid chromatography (HPLC) of methanol extracts of Douglas-fir apical shoots were able to distinguish known treatments of seedlings that varied in field survival, growth room survival, root growth potential (RGP) (number and total length of new roots), and time of lifting. Only a small number of HPLC peaks (or markers) were necessary to adequately explain treatment differences based on these factors. Two multivariate statistical approaches, discriminant and cluster analysis, were used to identify the biochemical markers. Stepwise discriminant analysis procedures were much more effective at selecting markers than was cluster analysis. Both peak frequency and peak area data were analyzed and found informative in their discriminating power, especially for explaining treatment differences based on time of lifting. The time seedlings were lifted influenced the HPLC peak profile in plants to such an extent that treatment differences based on survival, root growth potential, and exposure to freezing and drying conditions were sometimes overshadowed. This is especially true for groupings resulting from cluster analysis.

Treatments that varied in number of new roots were better separated by both approaches than those that varied in total length of new roots. Number of new roots was also found to be a better predictor of field survival that total length of new roots using a standard regression technique. Shared markers indicated a strong relationship between RGP measurements and survival in the field and growth room.

Seedling exposure to freezing and drying conditions could not be explained well by analysis with either type of data. This is not to imply that this approach cannot be used to identify markers of freezing and drying damage, but instead that this experimental design was unable to do so. However, it does appear that the presence

and quantity of compounds in treatments that were frozen or exposed to drying conditions differed depending on the time the seedlings were lifted.

Classes formed using growth room survival, field survival, RGP (number and total length of new roots), exposure to freezing and drying conditions, and time of lifting were found to be polythetic--class membership is founded on several markers, and no one marker is either necessary or sufficient to define the class. Markers that define classes may be used in a predictive capacity by classifying 'unknown' seedlings into appropriate survival, RGP, or time of lift groups. For predictive purposes markers cannot stand alone-- singly they are meaningless. Tables presented in Appendix D contain markers that define specific classes and may be useful in a predictive capacity.

For example, Table 3 contains fifteen peak frequency markers, the first six of which were adequate to explain group differences among treatments of known field survival. To predict the field survival of an 'unknown' lot of seedlings, peak frequencies of those fifteen markers in the plants must be determined and some, or all, of them used in a discriminant analysis procedure to classify the seedlings, with appropriate significance values, into the already defined groups-- in this example 0%, 1-20%, 21-50%, 51-80%, and 81-100% field survival. To correctly predict field survival only the first six markers may be needed, perhaps the first ten, or possibly all fifteen. However, markers from the tables must be kept in order; a cut-off point should be determined by trial-and-error classification of unknown seedlings. Tables are available for both peak frequency and peak area markers describing two field survival classifications, one growth room survival, exposure to freezing and drying conditions, and time of lift classifications, and two RGP group classifications. Since differences between classes defined by exposure to freezing and drying conditions could not be explained well by HPLC peaks, it is not likely that other seedlings can be accurately classified into these groups. Upon confirmation, this technique promises to be a rapid,

reliable, and quantitative means of identifying seedlings that have a low chance of surviving when planted in the field.

Although the informative capacity of markers, when used for predictive purposes, is dependent on the information carried by other peaks, several markers were consistently identified as important to explaining group differences for all factors. To illustrate, peak 78 was identified as a marker for field and growth room survival, number of new roots, and total length of new roots using multiple regression and cluster analysis of peak frequencies. It was also selected as a marker of growth room survival using stepwise discriminant analysis. Additionally, peak 58 was identified as a marker of field survival using stepwise discriminant analysis of either peak frequency or peak area data. These markers that do not rely heavily on the presence of others to be informative are identified in Table 4 and may, upon chemical characterization, aid in the investigation of the physiological and biochemical processes in seedlings of varying quality.

BIBLIOGRAPHY

Afifi, A. A. and V. Clark. 1984. Computer-Aided Multivariate Analysis. Lifetime Learning Publications, Belmont, CA.

Asen, S. and R. Griesbach. 1983. High performance liquid chromatographic analysis of flavonoids in geranium florets as an adjunct for cultivar identification. J. Amer. Soc. Hort. Sci. 108(5):845-850.

Burdett, A. N. 1979a. New methods for measuring root growth capacity: their value in assessing lodgepole pine stock quality. Can. J. For. Res. 9:63-67.

Burdett, A. N. 1979b. A non-destructive method for measuring the volume of intact plant parts. Can. J. For. Res. 9:120-122.

Burdett, A. N. 1987. Understanding root growth capacity: theoretical considerations in assessing planting stock quality by means of root growth tests. Can. J. For. Res. 17:768-775.

Burdett, A. N., D. G. Simpson, and C. F. Thompson. 1983. Root development and plantation establishment success. Plant and Soil 71:103-110.

Chapman, A. G. 1948. Survival and growth of various grades of shortleaf pine planting stock. Iowa State Col. J. Sci. 22:323-331.

Cleary, B. D., R. D. Greaves, and R. K. Hermann (eds.). 1982. Regenerating Oregon's Forests. Oregon State University Extension Service, Corvallis, OR. pp.3-4.

Dewald, L. E. and P. P. Feret. 1987. Changes in loblolly pine root growth potential from September to April. Can. J. For. Res. 17:635-643.

Duryea, M. L. 1984. Nursery cultural practices: impacts on seedling quality. *In* Forest nursery manual: Production of bareroot seedlings (M. L. Duryea and T. D. Landis, eds.). The Hague/Boston/Lancaster, for Forest Research Laboratory, Oregon State University, Corvallis, OR. p.143-164.

Duryea, M. L. 1985. Evaluating seedling quality: importance to reforestation. *In* Evaluating seedling quality: principles, procedures, and predictive abilities of major tests. (M. L. Duryea, ed.). Forest Research Laboratory, Oregon State University, Corvallis, OR. p.1-4.

Duryea, M. L. and T. D. Landis. 1984. Development of the Forest Nursery Manual: a synthesis of current practices and research. *In* Forest Nursery Manual: Production of Bareroot Seedlings. (M. L. Duryea and T. D. Landis, eds.). The Hague/Boston/ Lancaster, for Rorest Research Laboratory, Oregon State University, Corvallis, OR. p.3-5.

Duryea, M. L. and J. A. McClain. 1984. Altering seedling physiology to improve reforestation success. *In* Seedling Physiology and Reforestation Success. (M. L. Duryea and G. N. Brown, eds.). Martinus Nijhoff/Dr. W. Junk Publishers. p.77-114.

El-Kassaby, Y. A. and J. A. McLean 1985. Identification of the origins of lodgepole pine seeds by X-ray energy spectrometric determination of mineral profiles. Forest Sci. 31(3):539-551.

Everitt, B. S. 1977. Cluster Analysis. Heinemann Educational Books Ltd., London. pp.1-37.

Feret, P. P. and R. E. Kreh. 1985. Seedling root growth potential as an indicator of loblolly pine field performance. For. Sci. 31:1005-1011.

Feret, P. P., R. E. Kreh, and L. E. DeWald. 1985. Root growth potential of stored loblolly pine seedlings. *In* Proceedings, Third biennial southern silvicultural research conference (E. Shoulders, ed.). Atlanta, Georgia, Nov., 1984. pp.18-24.

Flury, B. and H. Riedwyl. 1988. Multivariate Statistics: A Practical Approach. Chapman and Hall, London.

Fowells, H. A. 1953. The effect of seed and stock sizes on survival and early growth of ponderosa and jeffrey pine. J. For. 51(7):504-507.

Hair, J. F., R. E. Anderson, and R. L. Tatham. 1987. Multivariate Data Analysis, 2nd Edition. Macmillan Publishing Company, New York.

Hermann, R. K. 1967. Seasonal variation in sensitivity of Douglas-fir seedlings to exposure of roots. For. Sci. 13:140-149.

Hermann, R. K. and D. P. Lavender. 1979. Testing the vigor of coniferous planting stock. Oregon State University Forest Research Laboratory, Corvallis, OR. Res. Note 63. 3 p.

Hermann, R. K., D. P. Lavender, and J. B. Zaerr. 1972. Lifting and storing western conifer seedlings. Oregon State University Forest Research Laboratory, Corvallis, OR. Res. Paper 17. 8 p.

Jay-Allemand, C., D. Cornu, and J. J. Macheix. 1987. Caracterisation du rajeunissement du noyer (*Juglans* sp) par une etude spectrophotometrique globale du contenu polyphenolique. Ann. Sci. For. 44(3):29-39.

Jenkinson, J. L. 1975. Seasonal patterns of root growth capacity in western yellow pines. *In* Proceedings of the 1975 convention, Society of American Foresters, Washington, D.C., September 28-October 2. pp. 445-453.

Jenkinson, J. L. 1984. Seed source lifting windows improve plantation establishment of Pacific slope Douglas-fir. *In* Seedling physiology and reforestation success (M. L. Duryea and G. N. Brown, eds.). Martinus Nijhoff/Dr. W. Junk Publishers. p. 115-141.

Jenkinson, J. L. and J. A. Nelson. 1978. Seed source lifting windows for Douglas-fir in the Humboldt Nursery. *In* Proceedings, Western forest nursery council and Intermountain nurseryman's association conference, Eureka, CA. p.1377-1395.

Johnsen, K. H., P. P. Feret, and J. R. Seiler. 1988. Root growth potential and shoot activity of northern and southern provenances of 1-0 eastern white pine seedlings grown in a Virginia nursery. Can. J. For. Res. 18:610-614.

Johnson-Flanagan, A. M. and J. N. Owens. 1986. Root respiration in white spruce (*Picea glauca* [Moench] Voss) seedlings in relation to morphology and environment. Plant Physiol. 81:21-25.

Kormanik, P. P. 1986. Lateral root morphology as an expression of sweetgum seedling quality. For. Sci. 32:595-604.

Krugman, S. L., E. C. Stone, and R. V. Bega. 1965. The effects of soil fumigation and lifting date on the root-regenerating potential of Monterey pine planting stock. J. For. 63(2):114-119.

Larsen, H. S., D. B. South, and J. M. Boyer. 1986. Root growth potential, seedling morphology and bud dormancy correlate with survival of loblolly pine seedlings planted in December in Alabama. Tree Physiol. 1:253-263.

Lavender, D. P. 1984. Plant physiology and nursery environment: interactions affecting seedling growth. *In* Forest nursery manual: Production of bareroot seedlings (M. L. Duryea and T. D. Landis, eds.). The Hague/Boston/Lancaster, for Forest Research Laboratory, Oregon State University, Corvallis, OR. p.133-141.

Lavender, D. P. 1985. Bud dormancy. *In* Evaluating seedling quality: principles, procedures, and predictive abilities of major tests. (M. L. Duryea, ed.). Forest Research Laboratory, Oregon State University, Corvallis, OR. p.7-15.

Lavender, D. P. and R. K. Hermann. 1976. Role of forest tree physiology in producing planting stock and establishing plantations. *In* XVI IUFRO World Congress Proceedings, Division II, Forest Plants and Forest Protection, Norway. April 1976. pp.34-35.

Lavender, D. P., R. K. Hermann, and P. Hinz. 1980. Controlled environment facilities for predicting vigor in outplanted Douglas-fir seedlings. *In* Proceedings, Fifth North American forest biology workshop, Gainesville, Florida. p. 242-244.

McCreary, D. D. and M. L. Duryea. 1985. OSU vigor test: principles, procedures, and predictive ability. *In* Evaluating seedling quality: principles, procedures, and predictive abilities of major tests. (M. L. Duryea, ed.). Forest Research Laboratory, Oregon State University, Corvallis, OR. p.85-92.

McCreary, D. D. and M. L. Duryea. 1987. Predicting field performance of Douglas-fir seedlings: comparison of root growth potential, vigor and plant moisture stress. New Forests 3:153-169.

McCreary, D. D. and J. B. Zaerr. 1987. Root respiration has limited value for assessing Douglas-fir seedling quality. Can. J. For. Res. 17:1144-1147.

Nambiar, E. K. S., G. D. Bowen, and R. Sands. 1979. Root regeneration and plant water status of *Pinus radiata* D. Don seedlings transplanted to different soil temperatures. J. Exp. Bot. 30:1119-1131.

Paton, R. R. 1929. The relation of size of seedling trees to their vigor. Ohio Expt. Sta., Bimonthly Bulletin. 141:191-194.

Pomeroy, K. B., F. R. Green, and L. B. Burkett. 1949. Importance of stock quality in survival and growth of planted trees. J. For. 47(9):706-707.

Puttonen, P. 1986. Carbohydrate reserves in *Pinus sylvestris* seedling needles as an attribute of seedling vigor. Scand. J. For. Res. 1:181-193.

Ritchie, G. A. 1982. Carbohydrate reserves and root growth potential in Douglas-fir seedlings before and after cold storage. Can. J. For. Res. 12:905-912.

Ritchie, G. A. 1984. Assessing seedling quality. *In* Forest nursery manual: Production of bareroot seedlings (M. L. Duryea and T. D. Landis, eds.). The Hague/Boston/Lancaster, for Forest Research Laboratory, Oregon State University, Corvallis, OR. p.243-259.

Ritchie, G. A. 1985. Root growth potential: principles, procedures, and predictive ability. *In* Evaluating seedling quality: principles, procedures, and predictive abilities of major tests. (M. L. Duryea, ed.). Forest Research Laboratory, Oregon State University, Corvallis, OR. p. 93-105.

Ritchie, G. A. and J. R. Dunlap. 1980. Root growth potential: its development and expression in forest tree seedlings. N. Z. J. For. Sci. 10(1):218-248.

Ritchie, G. A. and T. M. Hinckley. 1975. The pressure chamber as an instrument for ecological research. Adv. Ecol. Res. 9:165-254.

Schaefer, P. R. and J. W. Hanover. 1986. Taxonomic implications of monoterpene compounds of blue and Engelmann spruces. Forest Sci. 32(3):725-734.

Schopmeyer, C. S. 1940. Survival in forest plantations in the northern rocky mountain region. J. For. 38:16-24.

Schreiber, U., R. Fink, and W. Vidaver. 1977. Fluorescence induction in whole leaves: differentiation between the two leaf sides and adaptation to different light regimes. Planta 133:121-129.

Schreiber, U., L. Groberman, and W. Vidaver. 1975. Portable, solid-state fluorometer for the measurement of chlorophyll fluorescence induction in plants. Rev. Sci. Instrum. 46:538-542.

Shapiro, S. S. and M. B. Wilk. 1965. An analysis of variance test for normality. Biometrika 52:591-611.

Shiroya, T., G. R. Lister, V. Slankis, G. Krotkov, and C. D. Nelson. 1966. Seasonal changes in respiration, photosynthesis, and translocation of the ^{14}C labelled products of photosynthesis in young *Pinus strobus* L. plants. Ann. Bot. (London) 30:81-91.

Stiteler, W. M. 1978. Multivariate statistics with applications in statistical ecology. *In* The Second International Ecological Congress Satellite Program in Statistical Ecology, Parma, Italy. August 18-19, 1987.

Stone, E. C. 1955. Poor survival and the physiological condition of planting stock. For. Sci. 1(2):90-94.

Stone, E. C. and R. W. Benseler. 1962. Planting ponderosa pine in the California pine region. J. For. 60(7):462-466.

Stone, E. C. and J. L. Jenkinson. 1970. Influence of soil water on root growth capacity of ponderosa pine transplants. For. Sci. 16(2):230-239.

Stone, E. C. and J. L. Jenkinson. 1971. Physiological grading of ponderosa pine nursery stock. J. For. 69:31-33.

Stone, E. C., J. L. Jenkinson, and S. L. Krugman. 1962. Root regenerating potential of Douglas-fir seedlings lifted at different times of the year. For. Sci. 8(3):288-297.

Stone, E. C. and G. H. Schubert. 1959a. Root regeneration by ponderosa pine seedlings lifted at different times of the year. For. Sci. 5(4):322-332.

Stone, E. C. and G. H. Schubert. 1959b. The physiological condition of ponderosa pine (*Pinus ponderosa* Laws.) planting stock as it affects survival after cold storage. J. For. 57:837-841.

Sutton, R. F. 1979. Planting stock quality and grading. For. Ecol. and Manag. 2:123-132.

Sutton, R. F. 1980. Planting stock quality, root growth capacity, and field performance of three boreal conifers. N.Z. J. For. Sci. 10:54-71.

Sutton, R. F. 1983. Root growth capacity: relationship with field root growth and performance in outplanted jack pine and black spruce. Plant Soil 71:111-122.

Sutton, R. F. 1987. Root growth capacity and field performance of jack pine and black spruce in boreal stand establishment in Ontario. Can. J. For. Res. 17:794-804.

van den Driessche, R. 1978. Seasonal changes in root growth capacity and carbohydrates in red pine and white spruce nursery seedlings. *In* Proceedings, IUFRO symposium on root physiology and symbiosis. Nancy, France. (A. Riedacker and J. Gagnaire-Michard, eds.). p. 6-19.

van den Driessche, R. 1983. Growth, survival, and physiology of Douglas-fir seedlings following root wrenching and fertilization. Can. J. For. Res. 13:270-278.

van den Driessche, R. 1987. Importance of current photosynthate to root growth in planted conifer seedlings. Can. J. For. Res. 17:776-782.

Verbyla, D. 1986. Potential prediction bias in regression and discriminant analysis. Can. J. For. Res. 16:1255-1257.

Vidaver, W. and W. Binder. 1987. Variable chlorophyll fluorescence: predicting seedling quality. Solutions: the FRDA Research Newsletter 1(1):4-5.

Vidaver, W., K. Colbow, G. Hall, and S. Wessel. 1981. Chlorophyll fluorescence quenching by oxygen in plants. Can. J. Bot. 59:190-198.

von Rudloff, E. 1971. Chemosystematic studies in the genus *Psuedotsuga*. I. Leaf oil analysis of the coastal and Rocky Mountain varieties of the Douglas fir. Can. J. Bot. 50:1025-1040.

Wakeley, P. C. 1948. Physiological grades of southern pine nursery stock. Proc. Soc. Am. For. 43:311-322.

Wakeley, P. C. 1954. Planting the southern pines. USDA Agric. Mono. No. 18. 233p.

Weatherspoon, C. P. and R. J. Laacke. 1985. Infrared thermography for assessing seedling condition-- rationale and preliminary observations. *In* Evaluating seedling quality: principles, procedures, and predictive abilities of major tests. (M. L. Duryea, ed.). Forest Research Laboratory, Oregon State University, Corvallis, OR. p.127-135.

Weiser, C. J. 1970. Cold resistance and injury in woody plants. Science 169:1269-1278.

Winjum, J. K. 1963. Effects of lifting date and storage on 2+0 Douglas-fir and noble fir. J. For. 61:648-654.

Zaerr, J. B. 1967. Auxin and the root regenerating potential of ponderosa pine seedlings. For. Sci. 13(3):258-264.

Zaerr, J. B. 1985. The role of biochemical measurements in evaluating vigor. *In* Evaluating seedling quality: principles, procedures, and predictive abilities of major tests. (M. L. Duryea, ed.). Forest Research Laboratory, Oregon State University, Corvallis, OR. p.137-141.

Zaerr, J. B., B. D. Cleary, and J. L. Jenkinson. 1981. Scheduling irrigation to induce seedling dormancy. *In* Proceedings, Intermountain nurseryman's and Western forest nurseryman's 1980 meeting, Boise, Idaho. USDA For. Ser., Ogden, Utah. Int. For. and Range Exp. Sta. Gen. Tech. Rep. INT-109, p. 74-79.

Zaerr, J. B. and D. P. Lavender. 1976. Size and survival of 2-0 Douglas-fir seedlings. Forest Research Laboratory, Oregon State University, Corvallis, OR. Res. Pap. 32. 6 p.

Zavarin, E., K. Snajberk, and D. Bailey. 1976. Variability in the essential oils of wood and foliage of *Pinus aristata* and *Pinus longaeva*. Biochem. System. and Ecol. 4:81-92.

Zavarin, E., K. Snajberk, and P. Senter. 1979. Analysis of terpenoids from seedcoats as a means of identifying seed origin. Forest Sci. 25(1):20-24.

APPENDICES

APPENDIX A

Maxima table of custom results.

(The italicized number to the right of the component names are the final identifiers for each peak. They were determined after processing the data with the BASIC programs (Appendix B) and are not part of the normal Maxima custom report.)

The Maxima Chromatography Workstation (c)1985 Dynamic Solutions Corporation

MAXIMA CUSTOM REPORT

Printed: 2-APR-1987 16:58:01

SAMPLE: L1T1#2

#2 in Method: UV Method
Acquired: 4-SEP-1986 21:12
Rate: 2.941 points/sec
Duration: 41.001 minutes
Operator: VH

Type: UNKN
Instrument: Beckman-254
Filename: L1T1#2
Index: Disk

SAMPLE: L1T1#2 DETECTOR: UV-254

PK	Component Name		Retention Time (minutes)	Peak Area
1	P1D1	*(1)*	2.859	19691959
2	P2D1	*(2)*	7.171	1150143
3	P3D1	*(3)*	7.449	991218
4		*(6)*	7.947	1018736
5		*(7)*	8.112	1577729
6	P5D1	*(9)*	8.741	58312
7	P6D1	*(10)*	9.194	2629688
8	P9D1	*(13)*	10.543	2466608
9	P10D1	*(14)*	10.798	2442276
10	P12D1	*(15)*	11.342	3652765
11		*(17)*	11.908	6493952
12	P14D1	*(18)*	12.141	4293960
13	P15D1	*(21)*	12.464	8768146
14		*(23)*	12.656	79994
15	P19D1	*(27)*	13.251	3259710
16	P20D1	*(28)*	13.433	2555096
17	P21D1	*(29)*	13.795	10816919
18	P22D1	*(30)*	14.050	98182

19	P24D1	*(33)*	14.634	20095117
20	P26D1	*(35)*	15.093	24064
21	P27D1	*(37)*	15.354	105417
22	P28D1	*(39)*	15.626	241958
23	P29D1	*(40)*	15.847	176241
24	P30D1	*(41)*	16.000	486631
25		*(42)*	16.311	132489
26	P31D1	*(43)*	16.493	372392
27	P32D1	*(44)*	16.906	654717
28	P33D1	*(47)*	17.127	7631784
29	P34D1	*(49)*	17.592	325887
30	P35D1	*(51)*	17.915	19132
31	P36D1	*(53)*	18.363	15586246
32	P37D1	*(54)*	18.612	120526
33	P38D1	*(56)*	19.071	8163834
34	P41D1	*(62)*	20.057	42189831
35	P43D1	*(66)*	20.686	145237
36	P44D1	*(67)*	21.207	5771697
37	P45D1	*(68)*	21.683	32385380
38	P48D1	*(71)*	22.386	85190
39	P49D1	*(73)*	22.732	531152
40	P50D1	*(75)*	23.208	1009544
41	P51D1	*(76)*	23.514	201797
42		*(77)*	23.791	5753
43	P52D1	*(79)*	23.944	22704
44	P53D1	*(82)*	24.562	1171683
45	P54D1	*(83)*	24.789	1647356
46	P55D1	*(84)*	25.100	10230
47	P56D1	*(86)*	25.491	353682
48	P57D1	*(88)*	25.809	948895
49	P58D1	*(91)*	26.149	750432
50		*(94)*	26.574	346204
51	P59D1	*(95)*	26.829	4266198
52	P60D1	*(96)*	26.970	10758610
53	P62D1	*(99)*	27.486	20374
54	P63D1	*(100)*	27.741	257908
55	P64D1	*(101)*	28.149	3406960
56	P65D1	*(103)*	28.563	10262513
57	P66D1	*(104)*	28.693	278795
58	P67D1	*(106)*	29.197	4840493
59	P68D1	*(108)*	29.611	1549729
60	P69D1	*(109)*	29.838	2383274
61	P70D1	*(110)*	30.036	14723153
62	P71D1	*(113)*	30.518	240301
63	P72D1	*(114)*	30.784	1871411
64	P73D1	*(115)*	30.971	2082927
65	P74D1	*(116)*	31.118	4163038
66	P75D1	*(117)*	31.390	3974341
67	P76D1	*(118)*	31.532	4194059
68	P77D1	*(119)*	31.872	1519859

69		P78D1	*(120)*	32.082	20381153
70		P79D1	*(122)*	32.648	1758876
71		P80D1	*(123)*	32.790	7353647
72		P81D1	*(124)*	33.164	21446
73		P82D1	*(125)*	33.459	8296058
74		P83D1	*(126)*	33.640	78454
75		P84D1	*(127)*	33.974	12879755
76			*(removed)*	34.093	401031
77		P86D1	*(removed)*	34.870	36960
78		P87D1	*(removed)*	35.340	367012
79		P88D1	*(removed)*	36.944	10430
	TOTAL				336137357

APPENDIX B

Compiled BASIC programs for adjusting chromatography data.

```
1 REM **********************TEST PROGRAM*****************************
2 REM *  This program compares a reference with another file        *
3 REM *       By Jon B. Zaerr and Joe B. Zaerr                      *
4 REM *                    March, 1987                              *
5 REM *                        ***                                  *
6 REM ****************************************************************
10 DIM OUTFILE$(25)
20 DIM A(200),T(200),AREF(200),TREF(200)
30 DIM PK$(200),PKREF$(200)
40 DIM NUM(200),NUMREF(200),EXT(200),EXTREF(200)
45 DIM TEM$(200),TEM(200)
50 DEF SEG=0:POKE 1047,PEEK(1047)OR 64'ON CAPS LOCK
100 REM read reference file
110 RNAME$="L3T109FF.PRN"
120 OPEN RNAME$ FOR INPUT AS #1
130 INPUT #1, TEST$
140 IF TEST$ = "PROCESSED" THEN GOSUB 5000 'processed file
150 IF TEST$ <>"PROCESSED" THEN GOSUB 6000 'unprocessed file
160 REM main menu ++++++++++++++++++++++++++++++++++++
165 PRINT :PRINT
170 PRINT "Make choice..."
180 PRINT "1) Reconsider this whole adventure"
190 PRINT "2) Enter files to be processed"
200 PRINT "3) Exit"
210 PRINT :C=VAL(INPUT$(1))
220 ON C GOSUB 7000,300,290
230 GOTO 160
290 END
300 REM input files to be processed ******************************
310 PRINT "Enter .PRN files you wish to be processed: "
320 PRINT "Terminate entry by hitting <CR> with no file name.
330 'Begin file name input
340 NF = 1
350 PRINT
360 PRINT "file "NF": ";
370 INPUT "",OUTFILE$(NF)
380 IF NOT (OUTFILE$(NF) = "") THEN NF = NF + 1 : GOTO 360
390 NF = NF - 1
400 PRINT
410 FOR I = 1 TO NF
415 OUTFILE$(I)=OUTFILE$(I) + ".PRN"
420 PRINT OUTFILE$(I)
430 NEXT
440 PRINT : PRINT "ARE THESE FILES OK (Y/N)?"; : ANS$ = INPUT$(1)
```

```
450 IF ANS$ = "N" THEN PRINT "    RE-DO ALL ENTRIES..." : GOTO 300
460 IF ANS$ = "Y" THEN 480
470 GOTO 440
480 PRINT : PRINT
490 REM done with file input ---------------------
500 FOR FFF = 1 TO NF
505 LKJ=1
510 OPEN OUTFILE$(FFF) FOR INPUT AS #1
520 GOSUB 8000 ' read in file
530 GOSUB 955 'compute the differences
535 GOSUB 3000 'print the stuff out
540 NEXT FFF
550 GOTO 165
955  JR=1:J=1
960 LPRINT OUTFILE$(FFF)
965 PRINT OUTFILE$(FFF)
970 IF J>NPK THEN GOTO 1030
972 IF PK$(J)="" THEN J = J+1: GOTO 970
975 IF JR>NPKREF THEN GOTO 1030
980 IF PKREF$(JR)<>PK$(J) THEN JR=JR+1:GOTO 975
990 GOSUB 2000'difference subroutine
1000 JR=JR+1: J=J+1
1010 IF JR>NPKREF THEN GOTO 1030
1020 GOTO 970
1030 PRINT "Done"
1040 FOR I=LKJ TO LKJ+10
1050 TEM$(I)="" : TEM(I)=0
1060 NEXT I
1070 RETURN
2000 REM difference subroutine*************************
2010 DIF=TREF(JR)-T(J)
2020 DT=CINT(1000*DIF)
2030 PRINT PK$(J)"        "DT
2040 TEM$(LKJ)=PK$(J) : TEM(LKJ) = DT
2050 LKJ = LKJ + 1
2080 RETURN
3000 REM print the stuff subroutine ***********************
3010 HALF = CINT(LKJ/2) +1
3020 FOR I= 1 TO HALF
3030 LPRINT TEM$(I) TAB(15) TEM(I);
3040 LPRINT TAB(40) TEM$(HALF + I) TAB(55) TEM(HALF+I)
3050 NEXT I
3055 LPRINT CHR$(12) 'FORM FEED THE PRINTER
3060 RETURN
6000 REM input an unprocessed file *************************
6060 INPUT #1,JUNK$
6080 IF LEFT$(JUNK$,5)<>"-----"THEN GOTO 6060
6090 PRINT "READING REFERENCE FILE..."
6100 N=0
6110 B$=INPUT$(1,#1)
6120 IF B$=" "THEN :N=N+1:GOTO 6110
6130 L=N+10
6140 PKREF$(1)=B$+INPUT$(10,#1)
```

```
6150 PKREF$(1)=RIGHT$(PKREF$(1),11)
6160 INPUT #1,TREF(1),AREF(1)
6180 I=2
6190 PKREF$(I)=INPUT$(L,#1)
6200 PKREF$(I)=RIGHT$(PKREF$(I),10)
6210 IF PKREF$(I) <> "          "THEN GOTO 6220
6211 JJ$= INPUT$(55,#1)
6212 TREF(I)=VAL(JJ$)
6213 REM
6214 IF TREF(I)=0 THEN 6260
6216 INPUT #1,AREF(I)
6218 GOTO 6230
6220 INPUT #1,TREF(I),AREF(I)
6230 REM PRINT PKREF$(I),TREF(I),AREF(I)
6240 I=I+1
6250 GOTO 6190
6260 CLOSE #1
6270 NPKREF = I-1
6280 PRINT "# PEAKS = "NPKREF
6285 REM extract NUM & EXT ------------------
6286 PRINT "PROCESSING REFERENCE FILE..."
6290 FOR I=1 TO NPKREF
6300 T$=PKREF$(I)
6302 Q=INSTR(T$,"P")
6304 IF Q=0 THEN PRINT "Reference file has missing label at "I:GOTO 6345
6306 T$=RIGHT$(T$,LEN(T$)-Q+1)
6308 Q=INSTR(T$," ")
6310 IF Q<> 0 THEN T$=LEFT$(T$,Q-1)
6312 PKREF$(I)=T$
6350 NEXT I
6500 REM done with input ***************************
6600 RETURN
7000 REM save reference file
7010 PRINT "Enter your mothers maiden name"
7020 INPUT "Mach es fertig!";GARBAGE$
7030 RETURN
8000 REM input an unprocessed file to compare      ***********
8010 INPUT #1,JUNK$
8020 IF LEFT$(JUNK$,5)<>"-----"THEN GOTO 8010
8030 PRINT "READING FILE "OUTFILE$(FFF)
8110 I=1 : L=25
8130 PK$(I)=INPUT$(L,#1)
8150 IF INSTR(PK$(I),"P") <> 0 THEN 8220
8160 JJ$= INPUT$(55,#1)
8170 T(I)=VAL(JJ$)
8190 IF T(I)=0 THEN 8260
8200 INPUT #1,A(I)
8210 GOTO 8230
8220 INPUT #1,T(I),A(I)
8230 REM
8240 I=I+1
8250 GOTO 8130
8260 CLOSE #1
```

```
8270 NPK = I-1
8280 PRINT "# PEAKS = "NPK
8290 REM extract NUM & EXT ------------------
8295 PRINT "PROCESSING COMPARE FILE..."
8300 FOR I=1 TO NPK
8310 T$=PK$(I)
8312 Q=INSTR(T$,"P")
8314 IF Q=0 THEN EXT(I)=0 : NUM(I)=0 : PK$(I)="" : GOTO 8390
8316 T$=RIGHT$(T$,LEN(T$)-Q+1)
8318 Q=INSTR(T$," ")
8320 IF Q<> 0 THEN T$=LEFT$(T$,Q-1)
8322 PK$(I)=T$
8390 NEXT I
8400 REM done with input **************************
8410 RETURN
```

```
2 REM ********************FIX PROGRAM*****************************
4 REM * Inserts peaks from COMPARE files into REFERENCE file    *
5 REM * and saves COMPARE file as .REV                           *
6 REM * By Jon B. Zaerr and Joe B. Zaerr                        *
7 REM * Version 1.1   19 April, 1987                           *
8 REM ***********************************************************
10 DIM OUTFILE$(25)
20 DIM A(200),T(200),AREF(200),TREF(200)
30 DIM PK$(200),PKREF$(200)
40 DIM NUM(200),NUMREF(200),EXT(200),EXTREF(200)
50 DEF SEG=0:POKE 1047,PEEK(1047)OR 64'ON CAPS LOCK
70 DELTA = .1
100 REM read reference file
110 RNAME$="REFFILE.MAS"
120 OPEN RNAME$ FOR INPUT AS #1
130 INPUT #1, TEST$
140 IF TEST$ = "PROCESSED" THEN GOSUB 5500 'processed file
150 IF TEST$ <>"PROCESSED" THEN GOSUB 6000 'unprocessed file
160 REM main menu ++++++++++++++++++++++++++++++++++++
165 PRINT :PRINT
170 PRINT "Make choice..."
180 PRINT "1) Save reference file"
190 PRINT "2) Enter files to be processed"
200 PRINT "3) Exit"
210 PRINT :C=VAL(INPUT$(1))
220 ON C GOSUB 7000,300,290
230 GOTO 160
290 END
300 REM input files to be processed ******************************
310 PRINT "Enter .PRN files you wish to be processed: "
320 PRINT "Terminate entry by hitting <CR> with no file name.
330 'Begin file name input
340 NF = 1
350 PRINT
360 PRINT "file "NF": ";
370 INPUT "",OUTFILE$(NF)
380 IF NOT (OUTFILE$(NF) = "") THEN NF = NF + 1 : GOTO 360
390 NF = NF - 1
400 PRINT
410 FOR I = 1 TO NF
415 OUTFILE$(I)=OUTFILE$(I)+".PRN"
420 PRINT OUTFILE$(I)
430 NEXT
440 PRINT : PRINT "ARE THESE FILES OK (Y/N)?"; : ANS$ = INPUT$(1)
450 IF ANS$ = "N" THEN PRINT "    RE-DO ALL ENTRIES..." : GOTO 300
460 IF ANS$ = "Y" THEN 480
470 GOTO 440
480 PRINT : PRINT
490 REM done with file input --------------------
500 FOR FFF = 1 TO NF
505 LPRINT "FILE "OUTFILE$(FFF)
508 LPRINT "Delta = "DELTA" min"
510 OPEN OUTFILE$(FFF) FOR INPUT AS #1
```

```
520 GOSUB 8000 ' read in file
530 GOSUB 955 'compare the files
532  PRINT "Done with file ":LPRINT "Done with file"
534  GOSUB 5000 ' save the file
540 NEXT FFF
550 GOTO 165
955  JR=1:J=1
1000 PRINT PKREF$(JR)"+++++"PK$(J)"----"
1001 IF PKREF$(JR)=PK$(J) THEN 1200
1005 PRINT "JR= "JR"        J= "J
1010 IF PKREF$(JR)="" THEN GOSUB 3500:GOTO 1200 'insert after
1020 IF PK$(J)<>""THEN JR=JR+1 : GOTO 1000 'match peak names
1030 REM pk$ must be blank... find where to insert it
1040 JROLD=JR-1
1050 GOSUB 2000 'get fit value
1055 IF JR>NPKREF THEN GOSUB 3500 : GOTO 1200
1060 IF FIT >0 THEN JR=JR+1 : GOTO 1050
1065 JRB = JR-1 : JRA = JR ' bracketing peaks
1070 REM find next number in compare file
1075 NN=J
1080 NN=NN+1
1082 IF NN>NPK THEN GOTO 1100 'at end of compare file
1085 NXNUM = NUM(NN)
1090 IF NXNUM = 0 THEN GOTO 1080
1095 IF NXNUM < NUMREF(JRA) THEN GOSUB 4000 : GOTO 1210
1100 REM peak before
1105 IF EXTREF(JRB) = 1 THEN GOTO 1120
1110 JR=JRB : GOSUB 2000 ' get fit value
1115 IF ABS(FIT) < DELTA THEN GOTO 1150
1120 REM peak after
1125 JR = JRA : GOSUB 2000 ' get fit value
1130 IF EXTREF(JRA) = 1 THEN GOTO 1140
1135 IF ABS(FIT) < DELTA THEN GOTO 1150
1137 IF JR >NPKREF THEN GOSUB 3500:GOTO 1150
1140 GOSUB 3000 'insert a peak before index jr
1150 NUM(J) = NUMREF(JR)
1160 EXT(J) = EXTREF(JR)
1170 PK$(J) = PKREF$(JR)
1200 JR=JR+1
1205 J=J+1
1210 IF J>NPK OR JR>NPKREF THEN 1260
1220 IF NUMREF(JR) > NUM(J) AND(NUM(J)<>0) THEN JR=JR-1:PRINT "DEC JR": GOTO 1220
1250 GOTO 1000
1260 REM done with that file ------------------------
1280 RETURN
2000 REM fit subroutine ****************************************
2010 DT=T(J)-T(J-1)
2020 DTREF=TREF(JR)-TREF(JROLD)
2030 FIT=DT-DTREF
2040 RETURN
3000 REM insert blank peak subroutine ****************************
3010 FOR I=NPKREF TO JR STEP -1
3020 PKREF$(I+1)=PKREF$(I)
```

```
3030 NUMREF(I+1)=NUMREF(I)
3040 EXTREF(I+1)=EXTREF(I)
3050 AREF(I+1)=AREF(I)
3060 TREF(I+1)=TREF(I)
3070 NEXT I
3080 NUMREF(JR)=NUMREF(JR-1)
3090 EXTREF(JR)=EXTREF(JR-1)+1
3091 PRINT "NUM="NUMREF(JR)"   EXT="EXTREF(JR)
3100 NJR=JR
3110 GOSUB 3180 ' name pkref from num & ext
3120 AREF(JR)=A(J)
3130 TREF(JR) = TREF(JROLD)+(T(J)-T(J-1))
3140 NPKREF=NPKREF+1
3150 PRINT "PEAK INSERTED AT "NUMREF(JR)
3152 REM check if EXT is in order
3154 NJR=NJR+1
3156 NR = EXTREF(NJR)
3158 IF NR<>1 THEN EXTREF(NJR)=NR+1:GOSUB 3180:GOTO 3154
3160 FOR III=JR-2 TO JR+3
3162 IF III<1 THEN GOTO 3170
3170 NEXT III
3175 RETURN
3180 REM subroutine to name PKREF from NUMREF & EXTREF ***********
3192 NR$=STR$(NUMREF(NJR))
3193 NR =     NUMREF(NJR)
3194 T$="P"+RIGHT$(STR$(NR),LEN(NR$)-1)
3196 NR$=STR$(EXTREF(NJR))
3197 NR =     EXTREF(NJR)
3198 T$=T$+"D"+RIGHT$(STR$(NR),LEN(NR$)-1)
3199 PKREF$(NJR)=T$:PRINT T$
3200 RETURN
3500 REM add a peak at end subroutine ******************************
3510 PKREF$(JR)="P"+STR$(NUMREF(JR-1)+1)+"D1"
3520 AREF(JR)=A(J)
3530 TREF(JR)=T(J)
3540 NPKREF=NPKREF+1
3550 PRINT "Peak added at "NPKREF
3560 RETURN
4000 REM error in order of numbers subroutine ********************
4005 LPRINT "Error near "PKREF$(JR)" ***********"
4006  PRINT "Error near "PKREF$(JR)
4010 JR=JRB
4015 J = NN
4020 JR=JR+1
4030 IF EXTREF(JR)<>1 THEN GOTO 4020
4050 IF NUM(J) < NUMREF(JR) AND J<NPK THEN J= J+ 1 : GOTO 4050
4060 IF PKREF$(JR) <> PK$(J) THEN GOTO 4020
4070 RETURN
5000 REM save the compare file subroutine ************************
5005 SN$=LEFT$(OUTFILE$(FFF),LEN(OUTFILE$(FFF))-3)+"REV"
5006 PRINT SN$
5010 OPEN SN$ FOR OUTPUT AS #1
5030 PRINT #1,SN$
```

```
5035 PRINT #1,NPK
5040 FOR I=1 TO NPK
5050 WRITE #1, NUM(I),EXT(I),T(I),A(I)
5060 NEXT I
5065 WRITE #1,"END",0
5070 CLOSE #1
5080 PRINT "Compare file saved as "SN$
5090 RETURN
5500 REM read processed reference file ****************************
5510 INPUT #1,GARBAGE$
5515 INPUT #1, NPKREF
5520 FOR I=1 TO NPKREF
5530 INPUT #1,PKREF$(I),TREF(I)
5540 NEXT I
5550 CLOSE #1
5560 GOSUB 6280
5600 RETURN
6000 REM input unprocessed REFERENCE file ***************************
6060 INPUT #1,JUNK$
6080 IF LEFT$(JUNK$,5)<>"-----"THEN GOTO 6060
6090 PRINT "READING REFERENCE FILE..."
6100 N=0
6110 B$=INPUT$(1,#1)
6120 IF B$=" "THEN :N=N+1:GOTO 6110
6130 L=N+10
6140 PKREF$(1)=B$+INPUT$(10,#1)
6150 PKREF$(1)=RIGHT$(PKREF$(1),11)
6160 INPUT #1,TREF(1),AREF(1)
6180 I=2
6190 PKREF$(I)=INPUT$(L,#1)
6200 PKREF$(I)=RIGHT$(PKREF$(I),10)
6210 IF PKREF$(I) <> "          "THEN GOTO 6220
6211 JJ$= INPUT$(55,#1)
6212 TREF(I)=VAL(JJ$)
6213 REM
6214 IF TREF(I)=0 THEN 6260
6216 INPUT #1,AREF(I)
6218 GOTO 6230
6220 INPUT #1,TREF(I),AREF(I)
6230 REM PRINT PKREF$(I),TREF(I),AREF(I)
6240 I=I+1
6250 GOTO 6190
6260 CLOSE #1
6270 NPKREF = I-1
6280 PRINT "# PEAKS = "NPKREF
6285 REM extract NUM & EXT ------------------
6286 PRINT "PROCESSING REFERENCE FILE..."
6290 FOR I=1 TO NPKREF
6300 T$=PKREF$(I)
6302 Q=INSTR(T$,"P")
6304 IF Q=0 THEN PRINT "Reference file has missing label at "I:GOTO 6345
6306 T$=RIGHT$(T$,LEN(T$)-Q+1)
6308 Q=INSTR(T$," ")
```

```
6310 IF Q<> 0 THEN T$=LEFT$(T$,Q-1)
6312 PKREF$(I)=T$
6318 T$=RIGHT$(T$,LEN(T$)-1)
6320 Q=INSTR(T$,"D")
6330 NUMREF(I)=VAL(LEFT$(T$,Q-1))
6340 EXTREF(I)=VAL(RIGHT$(T$,LEN(T$)-Q))
6345 REM PRINT NUMREF(I)"    EXT"EXTREF(I)
6350 NEXT I
6500 REM done with input
6600 RETURN
7000 REM save reference file ************************************
7010 OPEN RNAME$ FOR OUTPUT AS #1
7020 PRINT #1, "PROCESSED"
7030 PRINT #1,"ref file"
7035 PRINT #1,NPKREF
7040 FOR I=1 TO NPKREF
7050 WRITE #1, PKREF$(I),TREF(I)
7060 NEXT I
7065 PRINT #1,"END",0
7070 CLOSE #1
7080 PRINT "master reference file saved as "RNAME$
7090 RETURN
8000 REM input an unprocessed file to compare      ***********
8010 INPUT #1,JUNK$
8020 IF LEFT$(JUNK$,5)<>"-----"THEN GOTO 8010
8030 PRINT "READING FILE "OUTFILE$(FFF)
8110 I=1 : L= 25
8130 PK$(I)=INPUT$(L,#1)
8150 IF INSTR(PK$(I),"P") <> 0 THEN 8220
8160 JJ$= INPUT$(55,#1)
8170 T(I)=VAL(JJ$)
8190 IF T(I)=0 THEN 8260
8200 INPUT #1,A(I)
8210 GOTO 8230
8220 INPUT #1,T(I),A(I)
8230 REM
8240 I=I+1
8250 GOTO 8130
8260 CLOSE #1
8270 NPK = I-1
8280 PRINT "# PEAKS = "NPK
8290 REM extract NUM & EXT ------------------
8295 PRINT "PROCESSING COMPARE FILE..."
8300 FOR I=1 TO NPK
8310 T$=PK$(I)
8312 Q=INSTR(T$,"P")
8314 IF Q=0 THEN EXT(I)=0 : NUM(I)=0 : PK$(I)="" : GOTO 8380
8316 T$=RIGHT$(T$,LEN(T$)-Q+1)
8318 Q=INSTR(T$," ")
8320 IF Q<> 0 THEN T$=LEFT$(T$,Q-1)
8322 PK$(I)=T$
8330 T$=RIGHT$(T$,LEN(T$)-1)
8340 Q=INSTR(T$,"D")
```

```
8360 NUM(I)=VAL(LEFT$(T$,Q-1))
8370 EXT(I)=VAL(RIGHT$(T$,LEN(T$)-Q))
8380 REM PRINT NUM(I)"    EXT"EXT(I)
8390 NEXT I
8395 RETURN : REM done with input
```

```
1 REM ******************FILL PROGRAM*****************************
4 REM * Inserts peaks from REFERENCE file into COMPARE files   *
5 REM * and resaves the COMPARE files with .FIN                *
6 REM * By Jon B. Zaerr and Joe B. Zaerr                       *
7 REM * Version 1.1    April, 1987                            *
8 REM ***********************************************************
10 DIM OUTFILE$(25)
20 DIM A(200),T(200),AREF(200),TREF(200)
30 DIM PK$(200),PKREF$(200)
40 DIM NUM(200),NUMREF(200),EXT(200),EXTREF(200)
50 DEF SEG=0:POKE 1047,PEEK(1047)OR 64'ON CAPS LOCK
70 DFROM$ = "A:" : DTO$ = ""
80 DELTA = .101
100 REM read reference file
110 RNAME$="REFFILE.MAS"
120 OPEN RNAME$ FOR INPUT AS #1
130 INPUT #1, TEST$
140 IF TEST$ = "PROCESSED" THEN GOSUB 5500 'processed file
150 IF TEST$ <>"PROCESSED" THEN PRINT "RUN FIX FIRST":END
160 REM main menu ++++++++++++++++++++++++++++++++++++
165 PRINT :PRINT
166 PRINT "From drive "DFROM$
168 PRINT "To drive "DTO$
169 PRINT :PRINT
170 PRINT "Make choice..."
180 PRINT "1) Change from- and to- drives"
190 PRINT "2) Enter files to be processed"
200 PRINT "3) Exit"
210 PRINT :C=VAL(INPUT$(1))
220 ON C GOSUB 7000,300,290
230 GOTO 160
290 END
300 REM input files to be processed ******************************
310 PRINT "Enter .REV files you wish to be processed: "
320 PRINT "Terminate entry by hitting <CR> with no file name.
330 'Begin file name input
340 NF = 1
350 PRINT
360 PRINT "file "NF": ";
370 INPUT "",OUTFILE$(NF)
380 IF NOT (OUTFILE$(NF) = "") THEN NF = NF + 1 : GOTO 360
390 NF = NF - 1
400 PRINT
410 FOR I = 1 TO NF
415 OUTFILE$(I)=OUTFILE$(I)+".REV"
420 PRINT DFROM$+OUTFILE$(I)
430 NEXT
440 PRINT : PRINT "ARE THESE FILES OK (Y/N)?"; : ANS$ = INPUT$(1)
450 IF ANS$ = "N" THEN PRINT "    RE-DO ALL ENTRIES..." : GOTO 300
460 IF ANS$ = "Y" THEN 480
470 GOTO 440
480 PRINT : PRINT
490 REM done with file input ---------------------
```

```
500 FOR FFF=1 TO NF
505 RF$= DROM$+OUTFILE$(FFF)
506 PRINT "FILE "RF$
508 PRINT "Delta = "DELTA" min"
510 OPEN RF$ FOR INPUT AS #1
520 GOSUB 8000 ' read in file
530 GOSUB 955 'compare the files
532  PRINT "Done with file "
534  GOSUB 5000 ' save the file
540 NEXT FFF
550 GOTO 165
955 J=1
1000 FOR JR = 1 TO NPKREF
1002 IF J > NPK THEN GOSUB 3500 : GOTO 1070
1005 PRINT NUMREF(JR)"D"EXTREF(JR)"---"NUM(J)"D"EXT(J)
1008 IF NUM(J) = 0 THEN GOSUB 4000 ' ERROR ROUTINE
1010 IF NUMREF(JR)<>NUM(J) THEN GOTO 1060
1020 IF EXTREF(JR)<>EXT(J) THEN GOTO 1060
1030 IF EXTREF(JR) = 1 THEN GOTO 1070
1040 GOSUB 2000
1045 PRINT "FIT= "FIT
1050 IF ABS(FIT) < DELTA THEN GOTO 1070
1060 IF J > NPK THEN GOSUB 3500 ELSE GOSUB 3000
1070 J=J+1
1080 NEXT JR
1260 REM done with that file ----------------------
1280 RETURN
2000 REM fit subroutine ****************************************
2010 DT=T(J)-T(J-1)
2020 DTREF=TREF(JR)-TREF(JR-1)
2030 FIT=DT-DTREF
2040 RETURN
3000 REM insert blank peak subroutine ****************************
3010 FOR I=NPK TO J STEP -1
3030 NUM(I+1)=NUM(I)
3040 EXT(I+1)=EXT(I)
3050 A(I+1)=A(I)
3060 T(I+1)=T(I)
3070 NEXT I
3080 NUM(J)=NUMREF(JR)
3090 EXT(J)=EXTREF(JR)
3091 PRINT "NUM="NUMREF(JR)"   EXT="EXTREF(JR)
3100 NJR=J
3120 A(J)=10
3130 T(J) = T(J-1)+ (TREF(JR)-TREF(JR-1))
3140 NPK=NPK+1
3150 PRINT "PEAK INSERTED AT "NUM(J)
3152 REM check if EXT is in order
3154 NJR=NJR+1
3155 IF NUM(J) <> NUM(NJR) THEN GOTO 3160
3156 NR = EXT(NJR)
3158 IF NR<=EXT(NJR-1) THEN EXT(NJR)=NR+1: GOTO 3154
3160 REM
```

```
3175 RETURN
3500 REM add a peak at end subroutine ********************************
3510 NUM(J)=NUMREF(JR) : EXT(J)=EXTREF(JR)
3520 A(J)=10
3530 T(J) = T(J-1)+ (TREF(JR)-TREF(JR-1))
3540 NPK=NPK+1
3550 PRINT "Peak added at "NPK
3560 RETURN
4000 REM error sub ***************************************
4010 PRINT "Big goof- check the "J"th entry in file "OUTFILES$(FFF)
4020 END
5000 REM save the compare file subroutine ************************
5005 SN$=LEFT$(OUTFILE$(FFF),LEN(OUTFILE$(FFF))-3)+"FIN"
5006 SN$ = DTO$+SN$ : PRINT SN$
5010 OPEN SN$ FOR OUTPUT AS #1
5030 WRITE #1,SN$
5035 WRITE #1,NPK
5040 FOR I=1 TO NPK
5045 DUM$="P"+STR$(NUM(I))+"D"+STR$(EXT(I))
5050 WRITE #1, DUM$,A(I)
5060 NEXT I
5070 CLOSE #1
5080 PRINT "Compare file saved as "SN$
5090 RETURN
5500 REM read processed reference file ***************************
5510 INPUT #1,GARBAGE$
5515 INPUT #1, NPKREF
5520 FOR I=1 TO NPKREF
5530 INPUT #1,PKREF$(I),TREF(I)
5540 NEXT I
5550 CLOSE #1
5560 GOSUB 6280
5600 RETURN
6280 PRINT "# PEAKS = "NPKREF
6285 REM extract NUM & EXT ------------------
6286 PRINT "PROCESSING REFERENCE FILE..."
6290 FOR I=1 TO NPKREF
6300 T$=PKREF$(I)
6302 Q=INSTR(T$,"P")
6304 IF Q=0 THEN PRINT "Reference file has missing label at "I:GOTO 6345
6306 T$=RIGHT$(T$,LEN(T$)-Q+1)
6308 Q=INSTR(T$," ")
6310 IF Q<> 0 THEN T$=LEFT$(T$,Q-1)
6312 PKREF$(I)=T$
6318 T$=RIGHT$(T$,LEN(T$)-1)
6320 Q=INSTR(T$,"D")
6330 NUMREF(I)=VAL(LEFT$(T$,Q-1))
6340 EXTREF(I)=VAL(RIGHT$(T$,LEN(T$)-Q))
6345 REM PRINT NUMREF(I)"    EXT"EXTREF(I)
6350 NEXT I
6500 REM done with input
6600 RETURN
7000 REM subroutine to change drive specs ************************
```

```
7010 PRINT :PRINT "From "DFROM$
7020 PRINT "To "DTO$
7030 PRINT :PRINT "Enter new drive specifications"
7040 INPUT "From ";DFROM$
7050 INPUT "To ";DTO$
7060 RETURN
8000 REM subroutine to read compare file ***************************
8010 INPUT #1,G$
8030 INPUT #1,NPK
8040 FOR J=1 TO NPK
8050 INPUT #1,NUM(J),EXT(J),T(J),A(J)
8060 NEXT J
8070 CLOSE #1
8080 RETURN
```

APPENDIX C

Master files of chromatography peaks and associated retention times.

Master file of compounds detected with 254 nm UV absorption.

Peak Number	Retention Time (min.)
1	2.90
2	8.17
3	8.33
4	8.55
5	8.67
6	8.85
7	9.09
8	9.57
9	9.72
10	10.13
11	10.72
12	11.18
13	11.52
14	11.73
15	12.26
16	12.60
17	12.74
18	12.90
19	13.12
20	13.30
21	13.37
22	13.52
23	13.62
24	14.03
25	14.13
26	14.23
27	14.40
28	14.64
29	14.92
30	15.18
31	15.39
32	15.65
33	15.85
34	16.10
35	16.36
36	16.52
37	16.63
38	16.76
39	16.91

Peak Number	Retention Time (min.)
40	17.14
41	17.27
42	17.55
43	17.73
44	17.87
45	17.99
46	18.17
47	18.30
48	18.53
49	18.81
50	18.97
51	19.16
52	19.33
53	19.59
54	19.74
55	20.00
56	20.19
57	20.31
58	20.48
59	20.60
60	20.70
61	20.95
62	21.37
63	21.61
64	21.75
65	21.97
66	22.16
67	22.55
68	22.96
69	23.23
70	23.42
71	23.72
72	24.00
73	24.15
74	24.32
75	24.52
76	24.73
77	25.00
78	25.19
79	25.30
80	25.45
81	25.61
82	25.71
83	25.89
84	26.12
85	26.28
86	26.46
87	26.68
88	26.78
89	26.90

Peak Number	Retention Time (min.)
90	27.07
91	27.23
92	27.33
93	27.47
94	27.60
95	27.70
96	27.81
97	27.97
98	28.17
99	28.38
100	28.51
101	28.77
102	28.95
103	29.28
104	29.43
105	29.60
106	29.79
107	29.99
108	30.30
109	30.47
110	30.65
111	30.81
112	30.93
113	31.05
114	31.26
115	31.40
116	31.57
117	31.80
118	31.92
119	32.24
120	32.44
121	32.69
122	32.99
123	33.12
124	33.53
125	33.82
126	34.07
127	34.40
128	34.45

Master file of compounds detected with a fluorescence detector.
(emmision: 360 nm, excitation: 290 nm)

Peak Number	Retention Time (min.)
129	2.79
130	3.82
131	4.15
132	4.61
133	5.09
134	5.92
135	6.84
136	7.66
137	8.75
138	9.31
139	9.66
140	10.92
141	11.16
142	11.41
143	11.79
144	12.01
145	12.32
146	12.64
147	12.83
148	13.22
149	13.57
150	13.76
151	13.87
152	14.07
153	14.32
154	14.47
155	14.62
156	14.80
157	14.97
158	15.08
159	15.24
160	15.42
161	15.86
162	16.08
163	16.29
164	16.45
165	16.73
166	16.96
167	17.11
168	17.31
169	17.53
170	17.68
171	17.90
172	18.23
173	18.43
174	18.67
175	19.03

Peak Number	Retention Time (min.)
176	19.42
177	19.62
178	19.78
179	20.08
180	20.41
181	20.68
182	21.00
183	21.19
184	21.52
185	21.72
186	21.91
187	22.16
188	22.34
189	22.62
190	22.96
191	23.29
192	23.63
193	24.05
194	24.38
195	24.50
196	24.69
197	25.04
198	25.23
199	25.48
200	25.78
201	26.18
202	26.46
203	26.65
204	26.85
205	27.02
206	27.28
207	27.42
208	27.92
209	28.43
210	28.92
211	29.30
212	29.60
213	30.01
214	30.60
215	30.85
216	31.08
217	31.18
218	31.30
219	31.52
220	31.71
221	31.83
222	31.92
223	32.10
224	32.28
225	32.45

Peak Number	Retention Time (min.)
226	32.69
227	32.89
228	33.58
229	34.40
230	36.36

APPENDIX D

Stepwise discriminant analysis and cross-validation results obtained using the discriminant analysis approach to marker detection.

Table D1. Stepwise discriminant analysis and cross-validation results obtained using the peak frequency data set. Groups are based on one-year field survival classes of 0-40%, 41-80%, and 81-100%. * Indicates a significance level to enter of 0.05 and to stay equal to 0.10. Otherwise, significance levels to enter and stay are equal to 0.15. † Indicates variables determined to adequately explain group differences and subsequently used to derive canonical variables that were plotted in Figure 24.

Step	Variable		Percent Misclass-ification	Partial R^2	Partial F Statistic	P>F	Wilk's Lambda	P<Lambda
	Entered	Removed						
* 1	† 58		57	0.3415	7.000	0.0036	0.65853	0.0036
* 2	† 79		47	0.2885	5.272	0.0120	0.46852	0.0005
* 3	† 85		27	0.2761	4.767	0.0176	0.33918	0.0001
* 4	† 136		23	0.2317	3.620	0.0423	0.26057	0.0001
* 5	† 20		27	0.2474	3.780	0.0381	0.19612	0.0001
* 6	† 111		20	0.3083	4.902	0.0174	0.13566	0.0001
* 7	† 114		20	0.2937	4.365	0.0260	0.09582	0.0001
* 8	† 230		17	0.3858	6.282	0.0076	0.05885	0.0001
* 9	† 48		10	0.3192	4.454	0.0259	0.04007	0.0001
10	77		3	0.2795	3.491	0.0523	0.02887	0.0001
11	37		10	0.3439	4.456	0.0278	0.01894	0.0001
12	210		13	0.3360	4.049	0.0378	0.01258	0.0001
13	108		6	0.3952	4.901	0.0230	0.00761	0.0001
14	197		3	0.5800	9.665	0.0023	0.00319	0.0001
15	228		0	0.5698	8.608	0.0042	0.00137	0.0001

Table D2. Stepwise discriminant analysis and cross-validation results obtained using the peak area data set. Groups are based on one-year field survival classes of 0-40%, 41-80%, and 81-100%. * Indicates a significance level to enter of 0.05 and to stay equal to 0.10. Otherwise, significance levels to enter and stay are equal to 0.15. † Indicates variables determined to adequately explain group differences and subsequently used to derive canonical variables that were plotted in Figure 25.

Step	Variable		Percent Misclass-ification	Partial R^2	Partial F Statistic	P>F	Wilk's Lambda	P<Lambda
	Entered	Removed						
* 1	† 58		60	0.3519	7.330	0.0029	0.64811	0.0029
* 2	† 155		33	0.2988	5.540	0.0099	0.45443	0.0003
* 3	† 20		37	0.4398	9.813	0.0007	0.25458	0.0001
* 4	† 205		30	0.3665	6.943	0.0042	0.16127	0.0001
* 5	† 168		13	0.3753	6.908	0.0045	0.10075	0.0001
* 6	113		7	0.4080	7.581	0.0031	0.05964	0.0001
* 7	228		7	0.4133	7.398	0.0037	0.03500	0.0001
* 8	97		7	0.3237	4.786	0.0200	0.02366	0.0001
* 9	148		7	0.6142	15.123	0.0001	0.00913	0.0001
*10	172		7	0.5409	10.602	0.0009	0.00419	0.0001
*11	64		10	0.5065	8.724	0.0025	0.00207	0.0001
*12	208		7	0.4045	5.435	0.0158	0.00123	0.0001
*13	200		7	0.5242	8.262	0.0038	0.00059	0.0001
*14	188		7	0.4404	5.509	0.0172	0.00033	0.0001
*15	50		7	0.4602	5.541	0.0182	0.00018	0.0001

Table D3. Stepwise discriminant analysis and cross-validation results obtained using the peak frequency data set. Groups are based on one-year field survival classes of 0%, 1-20%, 21-50%, 51-80%, and 81-100%. * Indicates a significance level to enter of 0.05 and to stay equal to 0.10. Otherwise, significance levels to enter and stay are equal to 0.15. † Indicates variables determined to adequately explain group differences and subsequently used to derive canonical variables that were plotted in Figure 26.

Step	Variable		Percent Misclass-ification	Partial R^2	Partial F Statistic	P>F	Wilk's Lambda	P<Lambda
	Entered	Removed						
* 1	† 12		73	0.4798	5.765	0.0020	0.52019	0.0020
* 2	† 58		47	0.4571	5.051	0.0043	0.28243	0.0001
* 3	† 41		47	0.3846	3.593	0.0204	0.17382	0.0001
* 4	† 197		43	0.3714	3.250	0.0308	0.10926	0.0001
* 5	† 219		50	0.3647	3.014	0.0412	0.06941	0.0001
* 6	† 218		30	0.3915	3.216	0.0342	0.04224	0.0001
* 7	144		30	0.4589	4.028	0.0157	0.02286	0.0001
8	48		17	0.3856	2.825	0.0557	0.01404	0.0001
9	150		20	0.4029	2.868	0.0552	0.00838	0.0001
10	59		23	0.5671	5.241	0.0068	0.00363	0.0001
11	188		23	0.6040	5.719	0.0053	0.00144	0.0001
12	106		23	0.6653	6.959	0.0027	0.00048	0.0001
13	92		27	0.6414	5.812	0.0066	0.00017	0.0001
14	195		13	0.6768	6.284	0.0058	0.00005	0.0001
15	222		17	0.8368	14.095	0.0003	0.00001	0.0001

Table D4. Stepwise discriminant analysis and cross-validation results obtained using the peak area data set. Groups are based on one-year field survival classes of 0%, 1-20%, 21-50%, 51-80%, and 81-100%. * Indicates a significance level to enter of 0.05 and to stay equal to 0.10. Otherwise, significance levels to enter and stay are equal to 0.15. † Indicates variables determined to adequately explain group differences and subsequently used to derive canonical variables that were plotted in Figure 27.

Step	Variable		Percent Misclass-ification	Partial R^2	Partial F Statistic	P>F	Wilk's Lambda	P<Lambda
	Entered	Removed						
* 1	† 12		77	0.4513	5.140	0.0037	0.54873	0.0037
* 2	† 4		70	0.3994	3.989	0.0128	0.32960	0.0004
* 3	† 58		50	0.5366	6.658	0.0010	0.15274	0.0001
* 4	† 141		50	0.4097	3.818	0.0167	0.09016	0.0001
* 5	† 197		50	0.3750	3.150	0.0355	0.05635	0.0001
6	217		47	0.3579	2.787	0.0545	0.03618	0.0001
7	150		33	0.3383	2.428	0.0834	0.02394	0.0001
8	113		27	0.3585	2.515	0.0778	0.01536	0.0001
9	177		33	0.4102	2.955	0.0505	0.00906	0.0001
10	112		27	0.5362	4.624	0.0113	0.00420	0.0001
11	48		20	0.6451	6.817	0.0025	0.00149	0.0001
12	120		13	0.4606	2.989	0.0561	0.00080	0.0001
13	128		13	0.5315	3.687	0.0323	0.00038	0.0001
14	179		10	0.5757	4.071	0.0260	0.00016	0.0001
15	133		3	0.6607	5.354	0.0121	0.00005	0.0001

Table D5. Stepwise discriminant analysis and cross-validation results obtained using the peak frequency data set. Groups are based on six-week growth room survival classes of 0-29%, 30-59%, 60-89%, and 90-100%. * Indicates a significance level to enter of 0.05 and to stay equal to 0.10. Otherwise, significance levels to enter and stay are equal to 0.15. † Indicates variables determined to adequately explain group differences and subsequently used to derive canonical variables that were plotted in Figure 28.

Step	Variable		Percent Misclass-ification	Partial R^2	Partial F Statistic	P>F	Wilk's Lambda	P<Lambda
	Entered	Removed						
* 1	† 133		67	0.5414	10.233	0.0001	0.45857	0.0001
* 2	† 117		50	0.3479	4.446	0.0123	0.29904	0.0004
* 3	† 22		43	0.3844	4.996	0.0078	0.18408	0.0003
* 4	† 98		37	0.3156	3.536	0.0306	0.12598	0.0004
* 5	† 78		30	0.3722	4.347	0.0150	0.07909	0.0001
* 6	† 13		27	0.3563	3.875	0.0238	0.05091	0.0001
7	† 230		23	0.3143	3.055	0.0521	0.03491	0.0001
8	155		27	0.3652	3.644	0.0314	0.02216	0.0001
9	35		13	0.4172	4.295	0.0188	0.01292	0.0001
10	54		17	0.4941	5.534	0.0077	0.00653	0.0001
11	207		13	0.4466	4.305	0.0209	0.00362	0.0001
12	209		23	0.4706	4.445	0.0200	0.00191	0.0001
13	26		20	0.5881	6.662	0.0051	0.00079	0.0001
14		98	20	0.2942	1.945	0.1688	0.00112	0.0001
15	51		10	0.6622	9.147	0.0013	0.00038	0.0001

Table D6. Stepwise discriminant analysis and cross-validation results obtained using the peak area data set. Groups are based on six-week growth room survival classes of 0-29%, 30-59%, 60-89%, and 90-100%. * Indicates a significance level to enter of 0.05 and to stay equal to 0.10. Otherwise, significance levels to enter and stay are equal to 0.15. † Indicates variables determined to adequately explain group differences and subsequently used to derive canonical variables that were plotted in Figure 29.

Step	Variable		Percent Misclass-ification	Partial R^2	Partial F Statistic	P>F	Wilk's Lambda	P<Lambda
	Entered	Removed						
* 1	† 133		63	0.5578	10.934	0.0001	0.44215	0.0001
* 2	† 165		43	0.4165	5.948	0.0033	0.25801	0.0001
* 3	† 8		47	0.4239	5.887	0.0037	0.14863	0.0001
* 4	† 147		33	0.4239	5.229	0.0067	0.08837	0.0001
* 5	† 208		27	0.4055	4.942	0.0090	0.05279	0.0001
* 6	198		30	0.4026	3.686	0.0282	0.03458	0.0001
* 7	135		20	0.3450	5.181	0.0082	0.01946	0.0001
* 8	77		23	0.4373	6.740	0.0028	0.00943	0.0001
* 9	114		17	0.5156	6.080	0.0048	0.00468	0.0001
*10	162		17	0.5033	5.376	0.0087	0.00240	0.0001
*11	69		17	0.4868	6.438	0.0046	0.00109	0.0001
*12	167		13	0.5469	4.016	0.0278	0.00060	0.0001
*13	214		20	0.4455	3.993	0.0301	0.00032	0.0001
*14	38		10	0.4611	4.151	0.0287	0.00017	0.0001
*15	177		13	0.4893	5.943	0.0101	0.00007	0.0001

Table D7. Stepwise discriminant analysis and cross-validation results obtained using the peak frequency data set. Groups are based on the average number of new roots per treatment. Classes consist of 0-39, 40-99, 100-174, and 175-200 new roots. * Indicates a significance level to enter of 0.05 and to stay equal to 0.10. Otherwise, significance levels to enter and stay are equal to 0.15. † Indicates variables determined to adequately explain group differences and subsequently used to derive canonical variables that were plotted in Figure 30.

Step	Variable		Percent Misclass-ification	Partial R^2	Partial F Statistic	P>F	Wilk's Lambda	P<Lambda
	Entered	Removed						
* 1	† 12		53	0.3222	4.119	0.0162	0.67784	0.0162
* 2	† 144		43	0.3146	3.825	0.0221	0.46461	0.0029
* 3	† 202		50	0.3622	4.543	0.0117	0.29632	0.0003
* 4	† 20		47	0.4369	5.948	0.0037	0.16687	0.0001
* 5	† 196		37	0.3799	4.493	0.0132	0.10347	0.0001
* 6	† 78		43	0.3885	4.446	0.0144	0.06328	0.0001
* 7	† 225		40	0.3665	3.857	0.0250	0.04009	0.0001
* 8	† 204		30	0.4388	4.953	0.0105	0.02249	0.0001
* 9	197		40	0.3618	3.401	0.0403	0.01436	0.0001
*10	52		23	0.4845	5.325	0.0090	0.00740	0.0001
*11	64		17	0.5247	5.887	0.0003	0.00352	0.0001
*12	108		27	0.6998	11.655	0.0018	0.00106	0.0001
*13	142		10	0.6465	8.535	0.0042	0.00037	0.0001
*14	39		10	0.6259	7.250	0.0137	0.00014	0.0001
*15	5		7	0.5754	5.421	0.0139	0.00006	0.0001

Table D8. Stepwise discriminant analysis and cross-validation results obtained using the peak area data set. Groups are based on the average number of new roots per treatment. Classes consist of 0-39, 40-99, 100-174, and 175-200 new roots. * Indicates a significance level to enter of 0.05 and to stay equal to 0.10. Otherwise, significance levels to enter and stay are equal to 0.15. † Indicates variables determined to adequately explain group differences and subsequently used to derive canonical variables that were plotted in Figure 31.

Step	Variable		Percent Misclass-ification	Partial R^2	Partial F Statistic	P>F	Wilk's Lambda	P<Lambda
	Entered	Removed						
* 1	† 56		70	0.3597	4.868	0.0081	0.64033	0.0081
* 2	† 141		60	0.3268	4.045	0.0179	0.43108	0.0013
* 3	† 133		53	0.3837	4.981	0.0079	0.26567	0.0001
* 4	† 48		37	0.6698	15.551	0.0001	0.08772	0.0001
* 5	† 130		23	0.4117	5.133	0.0077	0.05161	0.0001
* 6	142		27	0.3490	3.753	0.0265	0.03359	0.0001
* 7	89		17	0.3723	3.954	0.0230	0.02109	0.0001
* 8	15		17	0.4002	4.225	0.0190	0.01265	0.0001
* 9	117		17	0.4634	5.182	0.0093	0.00679	0.0001
*10	131		7	0.5764	7.711	0.0018	0.00287	0.0001
*11	189		7	0.4526	4.410	0.0193	0.00157	0.0001
*12	226		3	0.5276	5.585	0.0089	0.00074	0.0001
*13	17		0	0.5250	5.159	0.0131	0.00035	0.0001
*14	176		0	0.5456	5.203	0.0140	0.00016	0.0001
*15	4		0	0.5960	5.900	0.0103	0.00006	0.0001

Table D9. Stepwise discriminant analysis and cross-validation results obtained using the peak frequency data set. Groups are based on the average total length of new roots per treatment. Classes consist of 0-50, 51-500, 501-1000, 1001-2000, and 2001-3000 mm. * Indicates a significance level to enter of 0.05 and to stay equal to 0.10. Otherwise, significance levels to enter and stay are equal to 0.15. † Indicates variables determined to adequately explain group differences and subsequently used to derive canonical variables that were plotted in Figure 33.

Step	Variable		Percent Misclass-ification	Partial R^2	Partial F Statistic	P>F	Wilk's Lambda	P<Lambda
	Entered	Removed						
* 1	† 78		87	0.3863	3.933	0.0130	0.61375	0.0130
* 2	† 133		57	0.4302	4.529	0.0072	0.34974	0.0008
* 3	† 80		60	0.4075	3.955	0.0138	0.20721	0.0001
* 4	† 205		47	0.4072	3.777	0.0174	0.12284	0.0001
* 5	† 168		37	0.4804	4.853	0.0063	0.06383	0.0001
* 6	115		37	0.4061	3.419	0.0276	0.03791	0.0001
* 7	109		40	0.4573	4.003	0.0161	0.02057	0.0001
* 8	75		40	0.4185	3.238	0.0362	0.01196	0.0001
* 9	89		30	0.4455	3.414	0.0319	0.00663	0.0001
10	83		30	0.4145	2.832	0.0596	0.00388	0.0001
11	179		27	0.5189	4.045	0.0202	0.00187	0.0001
12	96		20	0.6377	6.161	0.0045	0.00068	0.0001
13	86		27	0.4875	3.092	0.0541	0.00035	0.0001
14	217		30	0.6507	5.587	0.0089	0.00012	0.0001
15	29		27	0.5320	3.127	0.0603	0.00006	0.0001

Table D10. Stepwise discriminant analysis and cross-validation results obtained using the peak area data set. Groups are based on the average total length of new roots per treatment. Classes consist of 0-50, 51-500, 501-1000, 1001-2000, and 2001-3000 mm. * Indicates a significance level to enter of 0.05 and to stay equal to 0.10. Otherwise, significance levels to enter and stay are equal to 0.15. † Indicates variables determined to adequately explain group differences and subsequently used to derive canonical variables that were plotted in Figure 32.

Step	Variable		Percent Misclass-ification	Partial R^2	Partial F Statistic	P>F	Wilk's Lambda	P<Lambda
	Entered	Removed						
* 1	† 78		83	0.4142	4.419	0.0077	0.58579	0.0077
* 2	† 205		40	0.5776	8.204	0.0003	0.24745	0.0001
* 3	† 133		40	0.4591	4.881	0.0054	0.13384	0.0001
* 4	† 141		40	0.4002	3.669	0.0195	0.08028	0.0001
* 5	† 188		40	0.4080	3.619	0.0215	0.04752	0.0001
* 6	187		40	0.3872	3.159	0.0364	0.02912	0.0001
* 7	77		40	0.4829	4.436	0.0106	0.01506	0.0001
* 8	219		30	0.4578	3.799	0.0207	0.00817	0.0001
9	165		33	0.3931	2.753	0.0622	0.00495	0.0001
10	167		13	0.5576	5.041	0.0080	0.00219	0.0001
11	225		7	0.5227	4.107	0.0192	0.00105	0.0001
12	125		7	0.6741	7.239	0.0022	0.00034	0.0001
13	227		7	0.7238	8.515	0.0013	0.00009	0.0001
14	89		10	0.7292	8.080	0.0021	0.00002	0.0001
15	43		3	0.7489	8.204	0.0026	0.00001	0.0001

Table D11. Stepwise discriminant analysis and cross-validation results obtained using the peak frequency data set. Groups are based on treatment exposure to freezing and drying conditions. * Indicates a significance level to enter of 0.05 and to stay equal to 0.10. Otherwise, significance levels to enter and stay are equal to 0.15. † Indicates variables determined to adequately explain group differences and subsequently used to derive canonical variables that were plotted in Figure 34.

Step	Variable		Percent Misclass-ification	Partial R^2	Partial F Statistic	P>F	Wilk's Lambda	P<Lambda
	Entered	Removed						
* 1	† 27		47	0.5460	10.423	0.0001	0.45400	0.0001
* 2	† 85		47	0.4984	8.280	0.0005	0.22774	0.0001
* 3	† 54		43	0.4203	5.801	0.0040	0.13201	0.0001
* 4	† 199		37	0.4211	5.577	0.0050	0.07642	0.0001
* 5	† 177		33	0.4230	5.376	0.0063	0.04409	0.0001
* 6	† 192		27	0.3946	4.563	0.0130	0.02669	0.0001
* 7	45		20	0.3760	4.017	0.0217	0.01666	0.0001
* 8	221		13	0.4667	5.542	0.0066	0.00888	0.0001
* 9	186		17	0.5138	6.340	0.0040	0.00432	0.0001
*10	70		17	0.4192	4.090	0.0234	0.00251	0.0001
*11		192	10	0.1849	1.285	0.3113	0.00308	0.0001
*12	229		10	0.3713	3.347	0.0438	0.00193	0.0001
*13	32		10	0.5106	5.563	0.0083	0.00095	0.0001
*14	103		13	0.6165	8.039	0.0020	0.00036	0.0001
*15	109		7	0.6852	10.158	0.0008	0.00011	0.0001

Table D12. Stepwise discriminant analysis and cross-validation results obtained using the peak area data set. Groups are based on treatment exposure to freezing and drying conditions. * Indicates a significance level to enter of 0.05 and to stay equal to 0.10. Otherwise, significance levels to enter and stay are equal to 0.15. † Indicates variables determined to adequately explain group differences and subsequently used to derive canonical variables that were plotted in Figure 35.

Step	Variable		Percent Misclass-ification	Partial R^2	Partial F Statistic	P>F	Wilk's Lambda	P<Lambda
	Entered	Removed						
* 1	† 210		47	0.4481	7.036	0.0013	0.55192	0.0013
* 2	† 69		43	0.3393	4.279	0.0144	0.36468	0.0002
* 3	† 114		47	0.3560	4.422	0.0130	0.23486	0.0001
* 4	† 78		47	0.2895	3.124	0.0455	0.16687	0.0001
* 5	† 121		43	0.3719	4.343	0.0151	0.10481	0.0001
* 6	† 27		40	0.3642	4.010	0.0211	0.06664	0.0001
* 7	65		40	0.4416	5.271	0.0077	0.03721	0.0001
* 8	184		50	0.4363	4.902	0.0109	0.02098	0.0001
* 9	43		47	0.4378	4.673	0.0139	0.01179	0.0001
*10	29		23	0.3815	3.495	0.0386	0.00729	0.0001
*11	135		23	0.4517	4.394	0.0195	0.00400	0.0001
*12	117		23	0.4481	4.060	0.0269	0.00221	0.0001
*13		210	40	2.4880	-	-	0.00330	0.0001
*14	218		20	3.5710	-	-	0.00190	0.0001
*15	20		13	6.9110	-	-	0.00780	0.0001

Table D13. Stepwise discriminant analysis and cross-validation results obtained using the peak frequency data set. Groups are based on the time seedlings were lifted. Classes consisted of seedlings lifted in October, December, January, and March. * Indicates a significance level to enter of 0.05 and to stay equal to 0.10. Otherwise, significance levels to enter and stay are equal to 0.15. † Indicates variables determined to adequately explain group differences and subsequently used to derive canonical variables that were plotted in Figure 36.

Step	Variable		Percent Misclass-ification	Partial R^2	Partial F Statistic	P>F	Wilk's Lambda	P<Lambda
	Entered	Removed						
* 1	† 144		100	0.9999	-	-	-	-
* 2	† 70		47	0.8068	54.303	0.0001	-	-
* 3	† 20		7	0.6553	23.764	0.0001	-	-
* 4	† 111		7	0.5219	13.100	0.0001	-	-
* 5	136		3	0.4977	11.397	0.0004	-	-
* 6	191		3	0.5196	11.897	0.0003	-	-
* 7	149		3	0.4613	8.991	0.0015	-	-
* 8	143		3	0.5105	10.427	0.0008	-	-
* 9	192		3	0.4441	7.590	0.0038	-	-
*10	155		3	0.4729	8.074	0.0031	-	-
*11	114		3	0.4394	6.663	0.0073	-	-
*12	79		3	0.4579	6.757	0.0075	-	-
*13	203		3	0.5915	10.859	0.0012	-	-
*14	24		3	0.4616	6.000	0.0131	-	-
*15	104		3	0.4287	4.879	0.0263	-	-

Table D14. Stepwise discriminant analysis and cross-validation results obtained using the peak area data set. Groups are based on the time seedlings were lifted. Classes consisted of seedlings lifted in October, December, January, and March. * Indicates a significance level to enter of 0.05 and to stay equal to 0.10. Otherwise, significance levels to enter and stay are equal to 0.15. † Indicates variables determined to adequately explain group differences and subsequently used to derive canonical variables that were plotted in Figure 37.

Step	Variable		Percent Misclass-ification	Partial R^2	Partial F Statistic	P>F	Wilk's Lambda	P<Lambda
	Entered	Removed						
* 1	† 64		20	0.8084	36.562	0.0001	0.19162	0.0001
* 2	† 33		13	0.6701	16.929	0.0001	0.06321	0.0001
* 3	† 57		7	0.6375	14.066	0.0001	0.02292	0.0001
* 4	† 20		3	0.5489	9.328	0.0003	0.01034	0.0001
* 5	18		3	0.4868	6.955	0.0018	0.00531	0.0001
* 6	74		3	0.5024	7.066	0.0018	0.00264	0.0001
* 7	216		3	0.4333	5.096	0.0088	0.00150	0.0001
* 8	175		3	0.5646	8.214	0.0010	0.00065	0.0001
* 9	163		3	0.6392	10.628	0.0003	0.00023	0.0001
*10	43		3	0.5787	7.784	0.0017	0.00010	0.0001
*11	127		3	0.6224	8.791	0.0011	0.00004	0.0001
*12	226		3	0.5858	7.070	0.0035	0.00002	0.0001
*13	144		3	0.6614	9.117	0.0013	0.000005	0.0001
*14	179		3	0.5532	5.365	0.0126	0.000002	0.0001
*15	192		3	0.5460	4.810	0.0201	0.000001	0.0001