1985

Crataegus Series Rotundifoliae In Southwestern Ontario: Patterns Of Morphometric Variation In An Agamic Complex

Paul Gardner Smith

CANADIAN THESES ON MICROFICHE

I.S.B.N.

THESES CANADIENNES SUR MICROFICHE

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

CRATAEGUS SERIES *ROTUNDIFOLIAE* IN SOUTHWESTERN ONTARIO:

Patterns of Morphometric Variation in an Agamic Complex

by

Paul Gardner Smith

Department of Plant Sciences

Submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Faculty of Graduate Studies

The University of Western Ontario

London, Ontario

March, 1985

ABSTRACT

The thesis defended in this dissertation asserts that individuals of *Crataegus* series *Rotundifoliae* (Rosaceae, Maloideae) in southwestern Ontario are members of an agamic complex and that the complex patterns of morphometric variation displayed by the series are a consequence of the reproductive behavior of agamic complexes. The specific problem addressed here arises out of a recognition that the longstanding taxonomic difficulty encountered in the series (and in the genus) can only be solved through intensive cytological, histlogical and numerical study of local populations.

Cluster Analysis, using a combination of resemblance measure-cluster method combinations, revealed the presence of four groups in a collection of 112 randomly sampled individuals evaluated for 29 morphological descriptors. These four groups and the populations into which they were subdivided display wide differences in levels of variability (univatiate descriptor variances and multivariate dispersion determinants) as well as in patterns of descriptor importance in Analysis of Dispersion Structure, Principal Components Analysis and Multi-group Discriminant Analysis. Members of the four groups are all apparently pseudogamous, facultative agamosperms with high pollen viability. All are apparently polyploid (mostly tetraploid with a few triploids) and self-fertile.

Taxonomic analysis, including the use of Cluster Analysis and Principal Components Analysis demonstrated that, in the study area, the series consists of six species, one of which (*C. ?lumaria*) is presumed to be of hybrid origin. The patterns of morphometric variation displayed by members of these species is related, in general, to a facultative apomictic mode of reproduction, and specifically to patterns of dispersal and to the postulated origin of the species within the agamic complex.

ACKNOWLEDGEMENTS

I wish to express sincere thanks to my supervisor, Dr. J. B. Phipps, not only for the material support he has provided me, through his research grants from the Natural Sciences and Engineering Research Council, but also for his understanding that research, and those who perform it, require time to develop. I am also grateful for financial support in the form of a Province of Ontario Graduate Scholarship and also for Special University Scholarships provided by the University of Western Ontario Faculty of Graduate Studies.

Throughout the time of development of my thesis, my way has been made smoother by the intellectual and emotional support of my friends, my colleagues and particularly by three people who are both of these to me. With Dr T. A. Dickinson it has been my pleasure to share a supervisor, a genus and many hours of discussion. Dr. J. Bowles has been a constant source of an objective point of view that is not too bound up in the mysteries of *Crataegus*. Above all, A. F. Stairs has given me the patience, understanding and support that has allowed me to complete this project. I cannot adequately express the depth of my gratitude to her.

I have also been assisted in many ways by past and present members of the Department of Plant Sciences. I would like to thank Ms. M. Mcleod for all of her help, specifically in performing the pollen analyses presented in Chapter 4. To Ms. M. Shaik I am extremely grateful for her work in the preparation and evaluation of embryological material, the results of which are also presented in Chapter 4. I would also like to thank the field and laboratory assistants with whom I have had the pleasure of working, particularly Ms. N. Power, Ms. J. Meinenger and Ms. D. Reimer. As well, the kindness and assistance of the secretarial staff, the support staff and the graduate students of this Department has been invaluable to me. Mr. J. Berry and Mr. R. Gardiner have been perticularly helpful in the final stages of preparation of this document.

Finally, I would like to express my utmost gratitude to my mother, Ms. H. Ailyn Smith and my father Mr. Gardner P. Smith, to whom I owe a debt that cannot be repaid. They have nourished and guided me throughout my life.

TABLE OF CONTENTS

LIST OF PHOTOGRAPHIC PLATES

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 The Thesis

The thesis defended in this work is as follows:

> Individuals of *Crataegus* series *Rotundifoliae* in Southwestern Ontario are members of an agamic complex the outer boundaries of which cannot be determined at present. The complex patterns of morphological variation exhibited by taxa in the series are specifically a consequence of the breeding patterns of agamic complexes.

An agamic heteroploid complex (Babcock and Stebbins 1938) comprises a grouping of specific and subspecific taxa that exhibit several levels of ploidy, often weak or nonexistent reproductive barriers between the polyploids and that reproduce by facultative or obligate agamospermy. Thus the first statement of the thesis is one concerning the reproductive behavior of members of the series. Very complex patterns of morphological variation are often associated with this sort of reproductive behavior. The second statement of the thesis, therefore, deals with how the particular patterns of variation exhibited by members of *Crataegus* series *Rotundifoliae* can be interpreted biologically.

The conclusions presented in this work represent a conservative synthesis of a long period of research and thought and are limited to a given set of taxa in a limited geographical area. However, the depth of the analysis, which is uncommon in studies of apomictic groups, permits one to approach questions of wider significance in the biology and morphometrics of apomictic situations and to speculate on their answers. Although a majority of this work will be directed towards establishing the verity of the thesis, some time will be spent considering these more speculative issues.

It should be clear from the statement of the thesis that the bounds of inquiry of the study are relatively restricted. No attempt will be made to provide a general taxonomic revision of *Crataegus* series *Rotundifoliae*. A true understanding of taxonomic relationships in the series can only come about if patterns of variation and reproductive behavior in local populations are first understood. There will be no attempt to explore the boundaries of the agamic complex of which it is maintained the series is a part. The scope of such a study is more on the order of a lifetime work than a doctoral thesis. Proximally, the thesis deals only with selected populations of *Crataegus* series *Rotundifoliae* in the vicinity of London, Ontario. However, evidence will be provided which justifies the extrapolation of the conclusions to a larger geographical range.

1.2 Assumptions

The thesis does not exist in isolation. It must be couched within the framework of our knowledge of biological systems. In order to do this,

certain assumptions have been made. These concern the sampling and taxonomy of *Crataegus* and also our understanding of apomictic modes of reproduction.

It is assumed that the genus *Crataegus* and the series *Rotundifoliae* are natural groups. That is to say, the name *Crataegus* series *Rotundifoliae* corresponds to some discrete entity in nature. It is also assumed that this entity has existed for some evolutionarily significant period of time, despite the fact that it was first recognized as a taxonomic group only within this century (Eggleston 1908). To a lesser extent, this is true of the groups that have been described as species and subspecies of the series in Ontario. These will be dealt with in more detail in a subsequent section. It is also assumed that the sample of individuals chosen for the research is representative of the series, at least within the geographical range specified.

Moreover, it must be assumed that the extant body of biological knowledge is generally true. Thus it is taken as given that the concept of an agamic complex is an appropriate interpretation of certain biological situations. That is, agamic complexes have objective existence.

1.3 The Study Organism

Crataegus L. is a member of the subfamily Maloideae of the Rosaceae. The subfamily includes about 25 genera worldwide (Hutchinson 1964). El Gazzar (1980) has attempted to divide *Crataegus* into two subgenera, the Eurasian subgenus *Crataegus* and the American subgenus *Americanae*, but

has done so on the basis of partially spurious evidence (Phipps 1983). Worldwide, the genus consists of about 35 series in 16 sections. In northeastern North America there are about 25 series according to Phipps (1983). One of these is series *Rotundifoliae* (Egglest. ex Egglest.) Rehd.

Crataegus species are woody shrubs or small trees reaching heights of up to 15 m. Most *Crataegus* individuals possess relatively long woody thorns. The branches bear short lateral shoots (short shoots) and long terminal shoots (long shoots), both of which bear leaves. The many-flowered inflorescences are only borne on short shoots. Flowers of *Crataegus* are pentamerous, shallow and dish shaped, with 10 or 20 (rarely 5 or 15) stamens. The anthers are ivory or pink to purple. Individuals undergo mass flowering in early to mid spring with anthesis being completed in as short a period as two or three days. The fruits are fleshy with stony nutlets (usually 2 to 4, occasionally 1 or 5) enclosing the seeds.

Crataegus species are mostly shade intolerant, usually occupying natural forest openings, erosion slopes and riverbanks, as well as numerous anthropogenic habitats such as abandoned fields. Colonization of newly opened areas appears at first sight to be random, with few to many different species occupying what is evidently the same habitat. Stands seldom persist for more than several generations of *Crataegus*, being replaced in the natural succession in this region by trees, particularly hardwoods.

The *Rotundifoliae* are among the smallest of the *Crataegus* in stature, seldom attaining a height of more than 3 metres as very thorny, irregularly branched shrubs. They are characterized by the possesion of small, more

or less round (obovate to isodiametric) leaves, ivory anthers and small red or orange fruit. Phipps and Muniyamma (1980) have indicated the presence of four fairly common to common species in Ontario: *C. aboriginum* Sarg., *C. chrysocarpa* Ashe, *C. dodgei* Ashe and *C. flavida* Sarg., and two rare ones: *C. irrasa* Sarg. and *C. margaretta* Ashe. Phipps considers all of them reasonably distinct from one another with the possible exception of *C. irrasa*, which is very similar to *C. chrysocarpa*. Further discussion of the taxonomy of series *Rotundifoliae* species will be reserved for chapter 5.

Phipps (1984) suggests that the most likely scenario for the origins of American *Crataegi* involves two main migration events from Asia across Beringia. In this scheme, the majority of those *Crataegi* currently occupying northeastern North America originated from the second wave of invasion.

1.4 The Crataegus Problem

The thesis has been couched within the framework of our understanding of biological phenomena. It is also important to couch it in an historical framework. By understanding how the '*Crataegus* Problem' (Palmer 1932, and others) arose, we can better evaluate the importance of the thesis and the significance of the conclusions reached.

The particular observations that gave rise to the formulation of the thesis were made by J. B. Phipps and co-workers during the course of investigations of the systematics of *Crataegus* (including a taxonomic revision) in Ontario. In the revision, Phipps and Muniyamma (1980) developed

a workable taxonomy for series *Rotundifoliae*, but noted that the series displayed complex patterns of morphological variation, which made identification very difficult in some cases. This observation is consonant with the views of Palmer (1946) who also noted a fair degree of difficulty in the taxonomy of the series. Muniyamma and Phipps (1979b) detected diploid, triploid and tetraploid individuals in the series. The same workers (1979a) demonstrated cytologically, for the first time, the presence of an apomictic mode of reproduction (apospory) in *C.* series *Pruinosae* Beadle and later (Muniyamma and Phipps 1984b) made the same observation for series *Rotundifoliae*.

Recognition of the problem has a much longer history, however, dating to the beginning of this century. It should be noted that the subject matter dealt with below has already been discussed extensively by Phipps and Muniyamma (1980), Dickinson (1983) and in references therein. This discussion, then, will simply attempt to highlight the most important aspects of the problem.

The history of the '*Crataegus* Problem' can be succinctly characterized by the number of new species described for the genus in each of three periods. These periods are: the early period (to about 1896), the period of expansion (1896 to about 1925) and the period of revision (1925 to present). During the early period, the majority of descriptions of North American *Crataegi* were based on material cultivated in European botanical gardens (Palmer 1925). Loudon (1838) included about 22 species in 15 sections (equivalent to series). Gray's Manual of Botany (1867) records only 10 species and 4 varieties. Subsequent to that publication, only a few new species were

described (eg. Chapman 1892) in the early period. The consensus of opinion by about 1890 was that North American *Crataegi* were relatively few but that some were polytypic (Palmer 1925).

Within 15 years the situation had changed drastically. In the period between 1896 and 1910, 866 species and 18 varieties of North American *Crataegus* were described (Brown 1910). Of these, 843 species and 9 varieties were attributable to four men: W. W. Ashe of the U.S.D.A. Forest service (165 species), C. D. Beadle, then director of the Biltmore Herbarium (144 species), W. W. Eggleston of the New York Botanical Gardens (10 species, 3 varieties) and C. S. Sargent, then Director of the Arnold Arboretum (524 species, 6 varieties). By the end of this period, over one thousand new descriptions had appeared (Palmer 1932). Sargent alone was eventually responsible for over 700 species and 22 varieties.

Since about 1925, taxonomic work on *Crataegus* has consisted primarily of revision with a considerable amount of reduction to synonymy of a large number of the taxa. W. W. Eggleston initiated much of this work, beginning as early as 1908. However, E. J. Palmer, of the Arnold Arboretum, is responsible for the vast majority of work on North American *Crataegus* in the middle part of this century. Palmer's contributions changed the state of *Crataegus* taxonomy considerably. In 1925 he published a nomenclator for the North American species of *Crataegus*, listing all of the vast number of species names present at the time and organizing them into series (Palmer 1925). By 1954, however, the number of species that he recognized had been reduced to 102 'valid' ones and another 60 or so 'mostly local species' for Northeastern and North Central U.S. and adjacent Canada (Palmer 1963).

E. P. Kruschke (1965) of the Milwaukee Public Museum published an annotated checklist (with synonomy) in which the treatment is broadly similar to Palmer's although somewhat more conservative (86 species for essentially the same geographical range). According to Phipps and Muniyamma (1980), this work represents the best treatment of northeastern North American *Crataegus* to date. Their revision, previously mentioned, lists 39 species, compared with the 91 species given in Sargent's (1908) account of the genus in Ontario. It is broadly comparable to the treatments of Palmer and Kruschke for the area in question, but certain refinements are made.

The sequence of early over-simplification followed by expansion and later by revision has probably occurred many times in the description of plant groups. What is notable about *Crataegus* is the amplitude of the changes. What remains to be explained are the underlying causes of this difference in perception of the genus. It is clear that both biological and purely human factors have contributed.

With respect to human factors, several explanations have been given. Brown (1910) surveyed the major *Crataegus* workers of the time on a number of questions, including the species concepts that they held. He concluded that the great increase in species descriptions was partly due to much narrower species concepts on the part of those workers as opposed to earlier workers. This situation is clearest in the case of C. S. Sargent, whose species descriptions were often based on the minutest of differences between forms. However, it is also true that Sargent had extensive field experience, whereas the earlier workers did not. Sargent's species, then, did reflect the wide

variety of types he observed in the field. Sargent also held the peculiar notion that new forms (equivalent to species in Sargent's view - see Sargent 1907a) would be found whenever a local area was investigated:

> "In every township in half a dozen states it is more than probable that forms exist which differ from those that have already been described..." Sargent (1907a, p. 291)

It is no wonder, then, that a large set of new species descriptions resulted from each of Sargent's collecting trips (Sargent 1901, 1903a,b,c, 1907a,b), and from the collections of his various correspondents.

C. D. Beadle (in Brown 1910), on the other hand, suggested that while there were probably a large number of *Crataegus* species, there had been a high degree of duplication of species descriptions. This was based on the opinion that those doing the describing were working too much in isolation. Palmer (1925) supported this observation (although he did not believe that the degree of duplication was as great as Beadle had suggested) and also asserted that the earliest descriptions of North American *Crataegi* lacked sufficient preciseness and were often unaccompanied by illustrations. These descriptions, he said, were variously interpreted by later workers, contributing even more to the confusion surrounding *Crataegus* taxonomy.

W.W. Ashe (in Brown 1910) suggested that earlier workers had based their (erroneous) understanding of the genus on herbarium vouchers (from which many important features are absent), rather than on direct field observation.

Narrow species concepts, more field observations and confusion over the correct application of binomials are not, in themselves, sufficient to explain the great burgeoning of species descriptions. The major factor militating against a simpler *Crataegus* taxonomy at that time was that the newly described species 'came true to seed'. That is, when seeds of the new forms were planted out (primarily by Sargent at the Arnold Arboretum, but also by C. D. Beadle), the minute differences between forms (which were the basis of the descriptions) were maintained in subsequent generations. Lacking an awareness of the possibility of asexual reproduction via seed (apomixis), these workers were convinced that each of the new forms was, in fact, a true species. The correct explanation of the phenomenon in *Crataegus* was not brought forward until the 1930's.

Meanwhile, evidence was brought forward that indicated that many *Crataegus* species were hybrids. Standish (1916) was the first to present experimental evidence for hybridization. She showed that a large number of *Crataegus* species displayed high degrees of pollen inviability. Longley (1924) provided further evidence through chromosome counts, showing a very high degree of triploidy in the genus. Unfortunately, Longley's counts are questionable (Palmer 1932), given that the material was hand-sectioned and the basic chromosome number was incorrectly reported as 16 (it is 17). Palmer (1932) gave qualified (and later (1946) wholehearted) support to the possibility of hybridization, on the basis of morphological evidence.

The experiments of Karl Sax of the Arnold Arboretum (reported by Palmer 1932) represent the first clear evidence that *Crataegus* species have the ability to undergo asexual reproduction via 'apogamy' (apomixis). Palmer

invoked apogamy (for triploid *Crataegi*) in order to explain the apparent conflict between evidence for hybridization and evidence for the constancy of progeny: If the triploid species were truly F_1 hybrids then their F_2 offspring would be expected to show wide variability unless the hybrids were reproducing apomictically. Rickett (1936) supports this hypothesis. Camp (1942a,b), relying heavily on the work of Babcock and Stebbins (1938) on the agamic complex in *Crepis*, was the first to take the modern view of *Crataegus* as a large agamic complex which:

> "...probably consists of a relatively few basic, sexual, diploid species, on which is superposed an unwieldy and highly complex population of triploids which, although they flower and produce fruit are really asexual apomicts. The comparatively few tetraploids may or may not be apomicts." Camp (1942b, p. 368)

This is essentially the current concept of *Crataegus*.

Thus, the 'period of expansion' of *Crataegus* descriptions can be understood in the following way. The workers of that period started with a poorly understood taxonomy inherited largely from European predecessors. When extensive field collections were made, a huge amount of previously unrecognized (but nonetheless real) variation was revealed. Lacking an awareness of alternate modes of reproduction and possessing rather narrow species concepts, these workers saw no alternative but to describe each of the multiplicity of forms as species. That these forms came true to seed was seen as vindication of their approach. As well, a peculiar sort of race developed to publish new species (Camp 1942a), which resulted in some degree of duplication and certainly in a large amount of confusion. As more and

more new species were described, it became more and more difficult to identify new collections. However, once additional evidence on the reproductive biology of *Crataegus* and other genera was incorporated, a more realistic understanding of the genus was obtained. Specifically, how such a multiplicity of forms could arise was better understood. Nonetheless, the fact that such a large number of forms do exist still remains. Thus, workers in the period of expansion were correct in their perception of wide variability in the genus if not in their taxonomic treatment of it.

In addition to the taxonomic revisions previously mentioned, more recent investigations of the genus have concentrated on lower taxonomic levels (mostly the population level) and have been aimed at illuminating aspects of reproductive biology and morphometric variation. The first of these was performed by Rickett (1936, 1937) who analysed univariate variation in populations of *C. pruinosa* (Wendl.) K. Koch and *C. crus-galli* L. in Missouri. Byatt (1975, 1976) examined hybridization in *Crataegus* populations in England, France and Belgium. Love and Feigen (1979) showed hybridization between the naturalized *C. monogyna* Jacq. and native *C. douglasii* var. *suksdorfii* Sarg. in Oregon.

Studies specifically examining *Crataegus* in the context of an agamic complex, however, have been confined to those performed by Phipps and co-workers. In addition to those studies previously mentioned, two other major analyses have been performed. Sinnott (1978, Sinnott and Phipps 1983) examined the systematics of series *Pruinosae* in Ontario. Dickinson (1983, Dickinson and Phipps 1984, 1985 in press) examined morphometric variation in relation to reproductive behavior in series *Crus-galli* L. Both of these

studies have been directed towards illuminating the particular form and extent of the *Crataegus* agamic complex by examining the behavior of local populations. This thesis can be seen in the same context, directed towards the same goal. Its specific aim is to examine the validity of the generalizations (about *Crataegus* and about series *Rotundifoliae*) heretofore made and to document the patterns of variation in the series and relate them to reproductive behavior. As well, a permanent contribution to the taxonomy of series *Rotundifoliae* will be made, by providing a taxonomic structure for series *Rotundifoliae* species in southwestern Ontario based on evidence from both morphometric and reproductive studies.

1.5 Plan of Analysis

It is clear from the statement of the thesis that two, mostly independent lines of research need to be pursued in order to establish its verity. One line of research involves establishing that the populations being examined are reproducing apomictically or at least have the ability to do so. Establishing this relies heavily upon histological studies of ovule development. However, additional evidence indicating ploidy levels, levels of pollen viability, degrees of outcrossing and flowering phenology can be used to evaluate the degree to which polyploidy and hybridization have contributed to formation of the agamic complex. In short, this line of research is an examination of elements of the reproductive biology of the series. Results from this line of investigation are presented in Chapter 4.

The other line of research is documentation of the patterns of morphological variation in the study populations. This is presented in

Chapter 3. There, both univariate and multivariate techniques will be used, that analyse variance, covariance and correlation and that involve cluster analysis. Synthesis of the two lines of research will consist of interpreting morphometric variability in the light of reproductive behavior displayed for each of the groups discovered.

CHAPTER 2

SAMPLING CONSIDERATIONS

2.1 INTRODUCTION

It is clear from the statement of the thesis that the geographical range of this study is limited. The choice of a limited range is a pragmatic one, given the detailed nature of the investigations to be performed and the time available. The local nature of the study presents a difficulty, however, in that a taxonomic investigation of the series, leading up to assignment of names and taxonomic rank for the entities studied here, should not be based on such a restricted range. In order to avoid this difficulty, a two-stage sampling strategy was devised. Four sampling sites, representative of the range of variation of the series in southwestern Ontario were non-randomly chosen for intensive investigation. These are all located within a 50 mile radius of London, Ontario. This particular region is extremely diverse in series *Rotundifoliae* species in that the distributional boundaries of both the more northern *C. chrysocarpa* and the more southern *C. dodgei* pass through this area. Samples collected from these sites were used in the analyses of morphometric variation (Chapter 3) and of reproductive behavior (Chapter 4). Eight additional sampling sites

from a much wider geographical range were also non-randomly chosen. These sites were chosen in order to include exemplars of series *Rotundifoliae* species, thought by Muniyamma and Phipps (1980) to occur in southwestern Ontario, in the taxonomic analysis presented in Chapter 5.

Descriptions of each of the four major and eight additional sites are presented in Appendix I. A map showing the location of each of the 12 sites is given in Figure 2.1

Major Collection sites

Choice of the four major collection sites was made with the assistance of my supervisor, Dr. J. B. Phipps (UWO), who has gained extensive field and herbarium experience of the range of variation of series *Rotundifoliae* in Ontario. Two of the sites (numbers 1 and 2) are abandoned fields (and adjacent forest edges) in the middle stages of succession. Site 3 is located along the banks of a small stream and an adjacent field edge. At site 4, series *Rotundifoliae* individuals occur along both sides of a path at the crest of a gorge of the Ausable river. On the side of the path opposite the gorge is a pine plantation.

Each of the major sampling sites was prepared for sampling in April, 1979. At sites 1 and 2, a 100 m. by 200 m. grid was laid out using compass and tape measure. Wooden stakes to mark grid points were laid out at 10 m. intervals. Thus, a total of 200 grid points were generated for sampling. From these, 75 grid points were randomly chosen. During sampling in May and June, an area of 10 m^2., centered around each randomly chosen grid

point, was assessed for the presence of series *Rotundifoliae* individuals. Only reproductively mature individuals, over 1 m. in height, were considered. If none occurred within that area, the grid point was declared empty. Otherwise, the individual nearest the grid point was sampled.

The same technique, but using a 100 m. by 130 m. grid was used at site 3. At site 4, a rectangular grid could not be established because the individuals occurred along the borders of a winding pathway. Here, grid points were located in a double column 5 m. distant from the centre of the pathway on both sides. 130 grid points were obtained and sampled in the same way as at the other sites.

Each sampled individual was permanently tagged, marked and numbered for later relocation. Individuals were sampled in the spring by clipping off two relatively exposed flowering branches from the middle level of the shrub. These branches, after pressing, were used for the morphometric analyses. At the same time, inflorescences of open or nearly open flowers were collected into preserving fluid. Preserved flowers were used for morphometric, pollen viability and histological studies. In late September, two branches were again collected and pressed and ripe fruit was collected into paper bags and cold stored. These collections were also used for morphometric analyses. Finally, in early spring, developing flower buds were collected, treated with colchicine for cytological studies and preserved. In this way 5 separate samples (spring vouchers, fall vouchers, flowers, fruit and buds) were obtained for each sampled individual.

The majority of sampling at the major sites was carried out in the spring and fall of 1979. However, complete samples were not obtained for all individuals during this period, either because of sampling error or because the individuals did not produce sufficient flowers or fruit for collection. Because of this, sampling at the major sites continued through 1982. Sometimes, individuals died (naturally or through human disturbance) before a complete sample could be obtained and these OTUs had to be disregarded. Ultimately, complete samples for 112 individuals were obtained.

2.3 Additional Sites

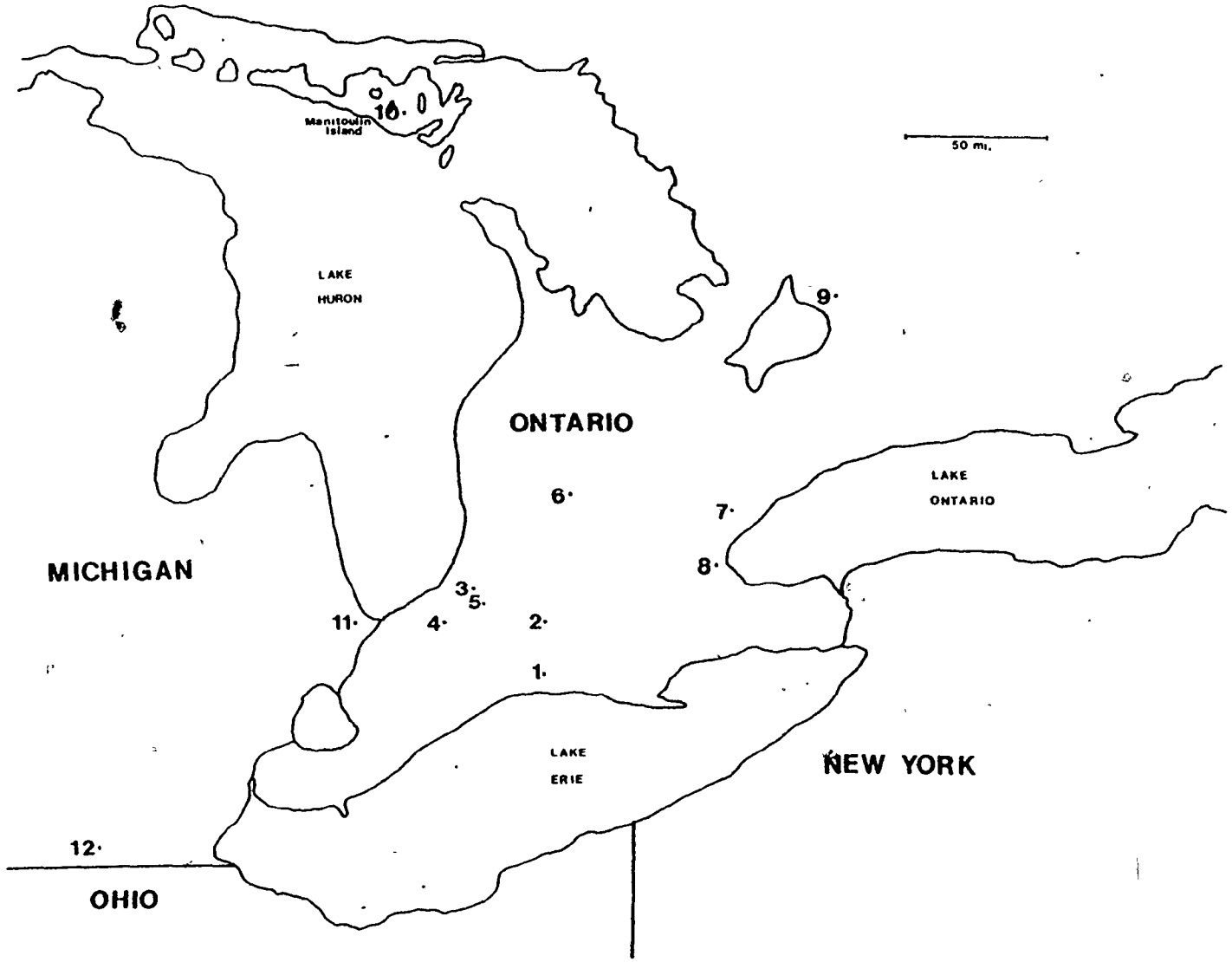
The eight additional sampling sites were also chosen with the assistance of J.B. Phipps. Individuals sampled from these sites were not randomly chosen. Instead, the range of variation at the site was assessed and individuals representative of that range were sampled. Individuals were tagged, numbered and marked in the same way as at the major sites. As well, similar sample material was taken from these individuals except for early flower buds, samples of which were taken from only a few individuals. The majority of sampling at these sites took place in 1982 and 1983.

One of the additional sites (site 10) represents a grouping of five individuals from three separate locations and seven individuals from one location all on Manitoulin Island. In addition, a few non-randomly sampled individuals from the major sampling sites were used in the investigations of reproductive behavior presented in Chapter 4. These individuals are included in the sample numerically analysed in Chapter 5.

For the numerical analyses reported in Chapters 3 and 5, Operational Taxonomic Units (OTUs) were defined as individual shrubs, as represented by the vouchers, flowers and fruit collected from them. OTUs are identified by the accession number given to them at the time of first collection (e.g. PS101).

In all, a total of 67 OTUs were non-randomly chosen and sampled at the additional sites and 5 individuals were non-randomly chosen from the major sites. Table 2.1 gives site number, location and number of individuals sampled for all 12 sampling sites.

Figure 2.1. Map showing the location of the 12 sampling sites used in morphometric and reproductive studies of *Crataegus* series *Rotundifoliae* individuals. Sites 1 to 4 represent the major collection sites (Chapter 3), sites 5 to 12 are the additional sites (Chapter 5).


Manitoulin Island
10.
50 mi.
LAKE
HURON
9.
ONTARIO
6.
7.
LAKE
ONTARIO
MICHIGAN
8.
3.
5.
11.
4.
2.
1.
LAKE
ERIE
NEW YORK
12.
OHIO

Site Number	Location	# of OTUs
1	Yarmouth Twp., Elgin Co. near St. Thomas, Ont.	20
2	London Twp., Middlesex Co. London, Ont.	43
3	McGillivray Twp., Middlesex Co. near Brinsley, Ont.	33
4	West Williams Twp., Middlesex Co. near Arkona, Ont.	21
5	London Twp., Middlesex Co. near Elginfield, Ont.	2
6	Elma Twp., Perth Co. near Listowel, Ont.	5
7	Nassagaweya Twp., Regional Mun. of Halton near Campbellville, Ont.	7
8	Flamborough Twp., Regional Mun. of Hamilton-Wentworth, Hamilton, Ont.	15
9	Carden Twp., Victoria Co. near Rohallion, Ont.	8
10	Allen, Gordon, Howland, Sandfield Twps. Manitoulin Island, Ont.	14
11	St. Clair Co. near Port Huron, Mich., USA	14
12	Lenawee Co. near Morenci, Mich., USA	3

Table 2.1 Location of the 12 sampling sites, and numbers of OTUs collected from each, used in analyses of morphometric variation and reproductive behavior of *C.* series *Rotundifoliae* individuals. Detailed site descriptions are given in Appendix 1.

CHAPTER 3

PATTERNS OF MORPHOMETRIC VARIATION

3.1 INTRODUCTION

This chapter will be devoted to documenting the patterns of morphometric variation displayed by the collection of randomly sampled individuals. First, however, it will be necessary to define exactly what is meant by patterns of morphometric variation.

Webster's New Collegiate Dictionary defines pattern as "a reliable sample of traits, acts or other observable features characterizing an individual". Sneath and Sokal (1973), referring to numerical taxonomic studies, define pattern as "any describable properties of the distribution of OTUs [individuals in this case] and groups of OTUs in A-space [the space defined by the descriptors characterizing the OTUs]". Thus, pattern can be thought of as a function of the descriptors (synonymous with characters or variables) used to characterize an individual or individuals. Moreover, pattern implies a recurrence of events (e.g. a bolt of cloth displays pattern in that the same sequence of colors is repeated over and over again). Thus,

it is repeatedness in aspects of these descriptors that gives rise to pattern. The specific aspect of the descriptors that is of interest is the behavior of these descriptors in terms of how they vary within and between groups in the sample.

Descriptor variation can be examined by both univariate and multivariate methods. Univariate analysis includes calculation of measures of central tendency, of variability and of the eveness of distribution of any one descriptor. Multivariate analysis examines the interrelationships between descriptors (including centroid, covariance and correlation). The multivariate approach is a more powerful one to the extent that the interrelationships between many descriptors are perceived to be important.

Patterns of variation occur at any level of organization. In this study, the entire sample of randomly collected individuals might be divided into groups representing taxa at or below the species level. Furthermore, these groups could be broken down into populations (e.g. according to the sampling site at which particular members of the group occur). The patterns of variation that may be exibited at different levels are not necessarily the same. Thus, documenting patterns of variation exhibited by the sample will require separate analyses at the level of the entire sample, at the level of groups within the sample, and at the level of populations within the groups.

3.2 MATERIALS AND METHODS

3.2.1 Descriptor Choice and Data Transformations

Twenty-nine descriptors were chosen for morphometric analysis of the random sample. The descriptors were chosen either because they had been found useful in describing variation in previous taxonomic studies of *Crataegus* in Ontario (Sinnott 1978, Phipps and Muniyamma 1980, Dickinson 1983), or because they had been found useful in an initial pilot study of series *Rotundifoliae*. Three ratio descriptors were included (despite the fact that ratios sometimes display non-normal distributions, Sokal and Rohlf 1973) because they provide useful measures of leaf shape (important in *Crataegus* taxonomy) that are independent of size. A list of the descriptors, their type and possible states that could be assumed (for multi-state characters) is presented in Table 3.1 (all tables and figures are presented at the end of the chapter). The descriptors were either continuous (9 descriptors), ratio (3 descriptors), meristic (6) or ordered multi-state (11). While this represents a statistically mixed descriptor set, all of the descriptors are subject to logical ordering and thus measures of distance between individuals can be meaningfully derived from descriptor scores.

Scores for the leaf characters were based on averages of measurements of 5 short-shoot leaves (replicates). Figure 3.1a illustrates selected leaf descriptors. Thorn measurements were based on 3 to 5 replicates. Measurments of fruits and nutlets were based on 20 replicates when sufficient fruit was available. Petal length and stamen number were based on averages of 5

replicates. Figure 3.1b illustrates selected thorn, fruit and flower descriptors. The decision as to the number of replicates to be used in generating descriptor scores was essentially a pragmatic one. The recognition that within-OTU descriptor variability exists and should be corrected for (in this case by the use of replicates) was balanced by the effort involved in obtaining and scoring replicate material. For all descriptors, it was subjectively determined that the number of replicates was sufficient to provide a reasonable estimate of the mean score for that descriptor.

In their raw form the descriptors were not commensurate, due to both size and scale differences. While there exist a number of data transformations that adjust descriptor scores to the same scale (Orlóci 1978), many of these seriously constrain the descriptor variances. Gower's (1971) ranging was used in the analyses presented here because it provides descriptor scores in the same range (0 to 1) without overly modifying the variances. The transformation from raw (X) to ranged (X') scores is made via the following formula.

$$X' = (X - X_{min}) / (X_{max} - X_{min})$$

where X_{max} is the largest score for that descriptor

and X_{min} is the smallest score.

Unless otherwise stated, results of the the morphometric analyses were all based on ranged, rather than raw, descriptor scores.

3.2.2 Univariate Analyses

The univariate analyses consisted of determining the mean, standard deviation, coefficient of variation (CV), frequency histograms and probability plots for each descriptor. All of the determinations were made with the use of the MINITAB statistical package (Ryan, Joiner and Ryan 1981) run on the DECsystem 1090 computer operated by the UWO Computing Centre.

Probability plots (and associated correlation coefficients) are used to determine the normality of descriptor distributions. In MINITAB, this is performed by generating a set of n (sample size) normal scores. Normal scores are the scores that could be obtained if a set of n individuals were randomly sampled from a normally distributed population with 0 mean and unit variance. The n normal scores are ranked according to magnitude and associated with actual descriptor scores (also ranked) on a one-to-one basis. Plots of actual versus normal scores can then be searched for departures from normality, indicated by departures from a straight line. Departures from normality can also be statistically tested based on the correlation coefficient between normal and actual scores (Ryan, Joiner and Ryan 1981). Significant lack of correlation indicates non-normality.

3.2.3 Multivariate Analyses

Four methods were used to examine multivariate patterns of variation. These are Cluster Analysis, Analysis of Dispersion Structure (ADS), Principal Components Analysis (PCA) and Multi-group Discriminant Analysis (MDA). The

first represents a family of classification techniques, seeking to find groups within a larger sample. The next two techniques perform manipulations of the dispersion (covariance) or correlation matrix. In ADS, sets of intercorrelated descriptors are identified. In PCA, linear compounds of descriptors are identified, that sequentially account for the greatest amount of variation remaining in the sample. The final technique (MDA) manipulates a matrix product related to within and between groups variability, so as to find linear compounds of descriptors that define greatest separation between groups.

3.2.3.1 Cluster Analysis

Cluster analysis is a general term covering a wide variety of techniques directed towards discovering group structure in a sample. The most widely used class of clustering techniques are those that Sneath and Sokal (1973) term SAHN (Sequential, Agglomerative, Hierarchical and Non-overlapping) techniques. These allow depiction of group structure in the form of an hierarchically stratified tree with OTUs at the branch tips, discrete clusters at internal nodes and resemblance between OTUs decreasing towards the root. This type of tree, when derived from a matrix of phenetic resemblances, is commonly known as a phenogram. SAHN techniques were used exclusively in this part of the analysis.

There are usually three major steps in Cluster Analysis, once OTUs have been chosen and scored. First, some pairwise measure of resemblance between OTUs is calculated and a matrix of OTU-OTU resemblance formed. Next, an algorithm for clustering OTUs on the basis of their resemblances

is applied, and a phenogram produced. Finally, some method of verifying the clusters discovered in the phenogram is required.

The user of clustering techniques must also make a choice between performing a single clustering of OTUs based on one preferred resemblance measure-clustering method (RM-CM) combination or performing several clusterings based on a variety of different combinations. In this study the latter approach was taken. The rationale for choice of this sort of approach is that each RM-CM combination, by the way in which it forms clusters, places certain constraints upon how the data (in the phenogram) is viewed. By applying a number of different RM-CM combinations, the user gains more information about group structure in the data set.

Two resemblance measures were chosen in the analysis. These are Euclidean Distances and Centered Cosines (correlation coefficients between OTUs). The former was chosen because of its extremely wide use in cluster analysis and because of its conceptual simplicity. The latter was chosen for its ease of calculation and because it represents a conceptually different approach to resemblance.

Three different clustering algorithms were used. These are Unweighted Pairs Grouped by Mathematical Averages (UPGMA, Sokal and Sneath 1963), Sums of Squares Agglomeration (SSA, Orlóci 1967, Ward 1963) and Single Linkage clustering (SL, Jardine and Sibson 1971). These three algorithms represent quite different approaches to clustering, specifically with respect to the type of clusters formed. SSA forms tight, well separated clusters (space dilating), SL forms long straggling clusters (space contracting) and

UPGMA represents the median between the other two (space indifferent). Thus, the different clustering methods themselves impose part of the group structure upon the data.

When combinations of conceptually different resemblance measures and conceptually different clustering strategies are applied to the same set of OTUs, a number of topologically different phenograms are usually produced. Each of these can be thought of as illuminating a different aspect of the group structure present in the sample. In this study, Euclidean Distances were combined with each of the clustering algorithms and Centered Cosines were combined with UPGMA. This gave a total of four phenograms.

All cluster analyses were performed using the CLUSTAN 1C computer package (Wishart 1979) run on a CYBER 170 computer operated by the UWO Computing Centre.

Clusters identified in the phenograms were verified in three ways. Voucher collections for each OTU in a given cluster were examined to subjectively determine whether the cluster was internally cohesive and distinct from other clusters, on morphological grounds. As well, the Cophenetic Correlation Coefficient (CPCC- Sokal and Rohlf 1962) was computed for resemblance matrix-phenogram and phenogram-phenogram comparisons. This coefficient provides a measure of how different each phenogram is from the resemblance matrix used to derive it and a measure of how different the phenograms are from each other. Finally, a durchschnitt Consensus Tree (Neumann 1983) was computed for the phenograms. Consensus trees (Adams 1972, Rohlf 1982, McMorris and Neumann

1983) pictorially summarize the information about group structure shared by a number of phenograms. Smith and Phipps (1984) discuss their use in cluster analysis. Consensus trees used in this analysis were computed using a BASIC computer program written for the purpose.

3.2.3.2 Manipulations of Dispersion Matrices

Before proceeding with descriptions of the remaining three multivariate techniques used in the analysis, the significance of manipulations of the dispersion matrix (and related matrices) in revealing patterns of variation will be discussed in more detail.

The dispersion or covariance between two descriptors is an expression of the degree to which, among different individuals, large values in one descriptor are associated with large values in another (e.g. whether or not those two descriptors are behaving in the same way). Thus dispersion is an expression of a particular relationship between the two descriptors. However, in a sample of more than two descriptors, there are many relationships between many descriptors. For instance, descriptor 1 may share high covariance with descriptor 2, descriptor 2 might show high covariance with descriptor 3 but descriptors 1 and 3 may have a lower covariance. So many relationships can often become highly confusing. In order to sort out these relationships, methods are required which manipulate dispersion matrices in order to reveal important relationships contained within those matrices (importance being defined by the method used).

Analysis of Dispersion Structure

One of the aspects of the dispersion matrix that can be examined is the degree to which subsets of descriptors covary together. This aspect will be termed dispersion structure. This method (originally derived under a different name and used for a different purpose) has been developed by Orlóci (1975) and Rohlf (1977). In the method, the total variance (S_{hh}) of each descriptor is partitioned into that portion which is unique to that descriptor (S_{hhs}) and that portion which is in common with the other descriptors in the sample (S_{hhc}). Thus;

$$S_{hh} = S_{hhs} + S_{hhc}$$

This partition is effected by inversion of the dispersion matrix (S to S^{-1}). The following relationships then apply;

$$S_{hhs} = 1/(S^{-1})_{hh}$$

and

$$S_{hhc} = S_{hh} - S_{hhs}$$

$(S^{-1})_{hh}$ is the hh (diagonal) element of the inverse of the dispersion matrix.

The multiple correlation coefficient (R_h^2) of descriptor h with the other descriptors describes the degree to which that descriptor shares variation with the other descriptors in the sample. It is defined as

$$R_h^2 = S_{hhc} / S_{hh}$$

e.g. the common component of variation relative to total variation.

The result of dispersion structure analysis is a table giving total, specific and common variances for each descriptor as well as multiple correlation coefficients.

Rohlf (1977) has suggested a further refinement of the technique whereby partial correlation coefficients ($R_{hi.}$) between any two descriptors can be calculated. These estimate the correlaton that would exist if the effects of the other descriptors were held constant. These are determined by

$$R_{hi.} = -(S^{-1})_{hi} / [(S^{-1})_{hh} (S^{-1})_{ii})]^{1/2}$$

These partial correlations can then be compared with Product Moment Correlation Coefficients (PMCCs) between each descriptor pair. If the partial correlation is very much larger than the PMCC then covariation between the two descriptors in the sample is being masked by other, stronger covariation patterns. If partial correlation is much smaller, then the observed covariation between the two descriptors is a result of their joint covariation with other descriptors.

Three important aspects of the dispersion matrix can be brought out by analysis of dispersion structure. These are

1. the number of descriptors with high multiple correlation coefficients, indicating the degree to which strong dispersion structure is present.

2. PMCCs, indicating the contribution of particular descriptor pairs to dispersion structure.

3. partial correlation coefficients, indicating the

degree to which descriptor covariances in the sample are affected by the presence of joint covariation with other descriptors.

Principal Components Analysis

Another method of analysing relationships contained in the dispersion matrix is Principal Components Analysis (PCA). The conceptual background and derivation of PCA has been dealt with in detail in many sources (Pimental 1979 gives a particularly clear account). What is important for this analysis is what descriptor relationships are brought out by PCA.

PCA seeks to partition variance in the sample into successive components of variation, rather than (as in the previous method) partitioning the variance of each descriptor separately. It does so by finding linear compounds of the original descriptors which successively account for the greatest amount of variation remaining in the sample. Each compound (component) is uncorrelated with the other components. Operationally, the method uses eigenanalysis of the dispersion matrix to extract the uncorrelated components.

Four important aspects of descriptor relationships are illuminated by PCA. These are

1. eigenvalues (roots), specifying the amount of the total variation in the sample that is accounted for by each component.

2. eigenvectors (component weights) specifying the degree to which each descriptor (in the linear

compound) is contributing to variation (loading) on each component.

3. component-descriptor correlations, indicating the correlation between each descriptor and each component.

4. percentage descriptor variance accounted for, specifying the amount of the total variance of each descriptor that is accounted for by each component.

The statistical significance of the components can be tested using Bartlett's sphericity test (Pimentel 1979). This compares the roots of the *n* smallest components (excluding the first component), to determine if they are statistically different. Equality of roots indicates that the variation being described by those component is random and thus not of interest (Pimental 1979).

Some difficulties arise, however, in the interpretation of component significance tests (Green 1978). The first is that the significance of the first component is assumed in tests of significance of subsequent components. Harris (1975, 1976) provides criticism of this assumption. Furthermore, the number of statistically significant components may not equal the number of components that can be reasonably retained on the basis of interpretability (Green 1978). In order to retain consistency in this analysis, the first two components in each PCA were given the greatest weight in interpreting results, regardless of statistical significance. Components after the first two that displayed statistical significance via Bartlett's test were given secondary consideration.

An additional use of PCA is that the position of each OTU in the space of any chosen two or three components can readily be plotted. This reduced dimension scatter plot can be used as an adjunct to Cluster Analysis, in that groups within the sample often appear separate when the sample is subjected to PCA and the results plotted.

Orlóci's (1978) R-algorithm Principal Components Analysis program (PCAR) was used in this part of the analysis. Program TABLE3 (written in BASIC by T. A. Dickinson) was used to compute component-descriptor correlations, percentage descriptor variance accounted for and to perform Bartlett's sphericity test. Both programs were run on the DECsystem 1090 computer of the U.W.O. Computing Centre.

Multi-Group Discriminant Analysis

The final method used in the analysis of pattern is Multi-Group Discriminant Analysis (MDA). MDA combines the notion of group structure with manipulations of the dispersion matrix to reveal descriptor relationships with respect to separation of groups in the sample. In fact it is a type of sums of squares and cross products (SSCP) matrix which is used in the analysis.

In MDA, total variation in the sample (SSCP matrix T) is partitioned into variation within groups (SSCP matrix W) and variation among groups (SSCP matrix A). Next, the asymetric matrix product $W^{-1}A$ is analysed in a way similar to PCA. Here, however, it is the ratio of among groups to within groups variability that is maximized on successive axes rather than

total variability in the sample. As in PCA, specific aspects of the relationships of the variables can be brought out. These are

1. eigenvalues (roots), representing the amount of among to within groups variability accounted for by each discriminant function (axis).

2. eigenvectors (discriminant weights), giving the contribution (loading) of each descriptor to discrimination among groups on each axis.

3. factor structure, indicating the degree of correlation between original descriptor scores and discriminant scores for each descriptor on each axis.

4. total communalities, indicating the overall contribution of each descriptor to discrimination among groups.

Plots of the position of each OTU in the space of the first two discriminant axes can also be produced to aid the interpretation of group differentiation.

The correct application of MDA to a particular data set relies on two assumptions; the equality of group dispersions and the inequality of group centroids. These assumptions can be tested; prior to performing MDA, with the use of Multivariate Analysis of Variance (MANOVA). MANOVA tests the null hypothesis of equal group dispersions first, via the F approximation to Box's M (1949). Equality of centroids is then tested using Rao's (1952) F approximation to Wilk's lamda (Λ). A further consideration in the application of MDA is that the number of descriptors can not exceed the

sample size of the smallest group.

Bartlett's test can also be used to test the significance of the discriminant axes.

In this analysis, MDA was performed using programs MANOVA and DSCRM (Cooley and Lohnes 1971), written in FORTRAN and run on the CYBER 170 computer of the UWO Computing Centre.

3.3 RESULTS

3.3.1 Patterns of Variation in the Random Sample

3.3.1.1 Univariate Results

As a first step in the analysis, means, standard deviations and CVs were determined for each descriptor based on raw descriptor scores (e.g. prior to ranging). These results are presented in Table 3.2, with descriptors grouped according to type. Frequency histograms and probability plots (not shown) were also produced for each descriptor, to find departures from normal distributions.

The first notable conclusion to be reached is that the magnitude of the coefficient of variation appears to be a function of descriptor type. The ordered-multistate descriptors generally possess the largest CV's and the ratio descriptors the smallest. This is clearly related to the observation (from frequency histograms) that the ordered multi-state descriptors are the least continuously distributed. Otherwise, results from both the

histograms and the probability plots indicated that most descriptors exhibit normal distributions with a few departures from normality in the multistate descriptors.

After applying Gower's ranging, the same univariate statistics were again computed. These results are presented in Table 3.3. The same general trends can be observed for the ranged data set. The calculation of CV's based on ranged (commensurate) data is unnecessary except for purposes of comparison with the raw data set. Although ranging has caused the CV's to have higher magnitudes, the multistate descriptors still exhibit the largest values and ratios the smallest. The same results apply for the standard deviations of the ranged data.

3.3.1.2 Multivariate Results

Dispersion Structure

Results from the Analysis of Dispersion Structure are presented in Table 3.4. In this table, multiple correlation coefficients (R^2_h, ranked by size) are given for each descriptor as well as total, common and specific variances. Below this, descriptor pairs with high PMCC values (|PMCC| > .7) are given along with an indication of whether or not the PMCC values that apply should be considered acceptable measures of pairwise contributions to multiple correlation (OK), overestimating contributions to correlation (HIGH) or underestimating contributions to correlation (LOW), on the basis of comparison with partial correlation coefficients.

It is clear from Table 3.4 that extremely strong dispersion structure is being displayed by the first 9 descriptors. Evidence for strong structure is supported by the high pairwise correlations for these descriptors. This is interpreted as evidence for a subset of descriptors all, in a sense, moving in the same direction and dominating variance in the sample. The three descriptors with the highest R^2_h values are all multistate and deal with pubescence on various parts of the plant (PDPUB, STPUB and HYPUB). The fourth and fifth largest multiple correlation coefficients are associated with two leaf descriptors (LFLEN and NSERR). The remaining four, in order, are LFPUB, LFWID, NTLEN and PELEN.

PCA

Principal Components Analysis was performed on the dispersion matrix of all 29 ranged descriptors. Results for the first 8 components are presented in Table 3.5. Table 3.5a gives eigenvalues (component roots), component weights (eigenvector elements) and individual and cumulative percentage variance accounted for by each component. Table 3.5b gives component-descriptor correlations and Table 3.5c percentage descriptor variance accounted for by each component. Results from Bartlett's test of equality of the last n (= 5 to 2) components are given in Table 3.6.

Approximately eighty percent of the variation in the sample was accounted for by the first 8 components. The first component accounts for nearly half of the variation in the sample. The rest of the variation is spread rather evenly over the remaining components. Component weights indicate a major contribution for the multistate pubescence descriptors

LFPUB, STPUB, PDPUB, HYPUB on the first component. Those for the second component show the strongest effects for the multistate gland descriptors PEGLN and CXGLN and for fruit colour FRCOA. The continuous and ratio descriptors do not seem to be loading heavily on any of the first 8 components.

Examination of tables 3.5b and 3.5c gives somewhat different results. Although the pubescence descriptors are very highly correlated with the first component, other descriptors (e.g. NSERR, TSERR, FRCOB and NTLEN) are also correlated. Similarly, on the second component LFLEN, LFWID and NNUTL as well as the gland descriptors are correlated. The same pattern is exhibited in Table 3.5c. Most descriptors have 30% or more of their variation accounted for on the first 3 (of 29) components. This reflects the high R^2_h values exhibited in the previous section. Bartlett's test (Table 3.6) indicates that only the last three components cannot be considered significantly different. Thus, 'significant' variation is also occurring on components 4 through 26.

For comparative purposes, a second PCA was performed using the correlation matrix. In this matrix, elements of the prime diagonal (equivalent to total variances for all descriptors in the dispersion matrix) are equal (unity). Thus, descriptors with relatively large variances are not allowed to dominate the analysis. Given that some of the multistate descriptors strongly influenced the PCA-dispersion results and that those same descriptors showed high variances, it was perceived that modifying the effect of those variances might prove worthwhile. Results of this analysis are presented in Table 3.7.

The results do not appear to strongly differ from those of the previous analysis with one exception; loading for all descriptors is more evenly spread over the components (Table 3.7a). Otherwise, both analyses are similar with respect to component descriptor correlations (Table 3.7b) and percentage variance accounted for (Table 3.7c). Bartlett's test was not performed.

Cluster Analysis

A total of four phenograms were produced in this analysis. As previously stated, these correspond to the four different resemblance measure - cluster method (RM-CM) combinations used. The phenograms are presented in Figures 3.2 to 3.5. In these figures, OTUs are represented by symbols identifying the four groups into which the sample was ultimately divided. Visual examination of these phenograms gives the impression of a high degree of congruence among them, indicating very strong group structure. This impression is substantiated by high Cophenetic Correlation Coefficients between resemblance matrices and phenograms and between phenograms (Table 3.8). It is clear from these that relatively little distortion has occurred in moving from resemblance measure to phenogram and that the phenograms are all very similar.

The durchschnitt consensus tree, summarizing the information shared by all the phenograms, is presented in Figure 3.6 (the actual sequence of OTU numbers in this figure is given in Appendix 3). In addition, a scatter diagram, showing the position of the OTU's on the first two components of the PCA-correlation, is presented in Figure 3.7. Subjective verification of group structure in the sample was made on the basis of these two figures.

Voucher specimens for each OTU were sorted into groups and evaluated for visible differences between groups. The result of this 'gestalt' of group structure was the division of the sample into four groups. A list of OTU's for each group is given in Table 3.9, along with symbols by which the groups are identified in the figures.

It is clear that strong group structure is present in the random sample, as evidenced by the high degree of congruence amongst the phenograms. However, it would be premature at this point in the analysis to attempt to assign taxonomic rank to the groups thus far identified. It will suffice to simply refer to them as groups (e.g. Group 1, Group 2 ..), and reserve taxonomic considerations for Chapter 5.

3.3.2 Patterns of Variation in the Groups

3.3.2.1 Univariate Results

A second round of univariate analysis was performed on each of the four groups found in the Cluster Analysis. A summary table of means and standard deviations for each group and for the entire sample is presented in Table 3.10. Several important observations arise.

In over half of the descriptors, the standard deviation of at least one group was greater than or equal to that for the entire sample. The majority of descriptors that did show a decrease in all four groups were multistate ones. In fact, 5 of the multistate descriptors were invariant in one or more of the groups.

With respect to groups, Group 4 generally showed the least variability. In only one descriptor (MSTAN) was the standard deviation higher in Group 4 than in the entire sample, and in only 2 descriptors (BSHAP, MSTAN) did Group 4 show the highest standard deviation amongst the groups. Conversly, Group 1 exhibited the greatest variability. There were 9 descriptors in this group that showed standard deviations larger than those for the entire sample, and 12 descriptors that showed the highest standard deviations amongst groups.

Departures from normality were again examined with histograms, probability plots and, additionally, statistical tests of significant departures from normality based on the probibility plots. Most of the descriptors showed no significant departures from normality. However, five descriptors showed non-normality in one of the four groups (RLOWP, NTLEN, R1412, FRWID,FRCOB), one was non-normal in two groups (NNUTL) and one in three groups (AANG). Most of these departures, however, appeared to be the result of only one or two outlier individuals (identified as such in the Cluster Analysis) stretching out the distribution.

3.3.2.2 Multivariate Results

At this point in the analysis, a difficulty arose with respect to the number of descriptors being used. As a result of dividing the sample into groups, the sample size was reduced from 112 to 7 (in the case of the smallest group, Group 3). The inherent dimensionality of a dispersion matrix is at most n-1 where n is the sample size. Attempts to invert a dispersion matrix (as is done in ADS and MDA) of dimensionality greater than n-1 will meet

with failure, due to singularity. At this point, the choices available were abandoning ADS and MDA in further analyses, not examining groups with sample size less than 28 (29 descriptors - 1) or reducing the descriptor set to 6 descriptors. It was determined that reducing the descriptor set would be the least damaging to the continuity of the analysis. Even this option presents problems, given the major loss of information that must result from reduction of the descriptor set. On the other hand, reducing the descriptor set to 6 does allow for more careful observation of descriptor behavior and thus a better understanding of patterns of variation in the (reduced) data set.

The choice of the six descriptors to be used in the remainder of the analyses was made on the basis of two criteria. It was determined that continuous and ratio descriptors were preferable on the basis of their generally more normal distributions and because previous analyses of this sort have found these descriptor types to be better behaved (Pimental 1979). Of the 12 continuous and ratio descriptors, five (RLOWP, R1412, AANG, FRWID and NTLEN) were rejected because of some departure from normality. Of the remaining seven descriptors, LFWID was rejected because it was subjectively determined that R3412 incorporated much of the information contained in LFWID, and it was considered desirable to retain a measure of leaf shape independent of size. Thus, for the remainder of the analyses of the random sample, 6 descriptors were used. Four of these are leaf descriptors (LFLEN, PELEN, R3412, BANG), one pertains to thorns (THLEN) and one to flowers (PTLEN). These six descriptors are identified by asterisks in Table 3.10.

Dispersion Structure

Analysis of Dispersion Structure, based on the 6 ranged descriptors, was performed for each of the four groups and for the entire random sample. Results are presented in Tables 3.11 through 3.15. Table 3.11 gives results for the entire sample. This can be compared with Table 3.4 (dispersion structure using 29 ranged descriptors). It is clear that multiple correlation coefficients and pairwise PMCCs are considerably lower for the 6 descriptors when they are considered in isolation. This can be understood as a function of reducing the data set. There are fewer descriptors with which to form strong dispersion structure.

Table 3.11 can also be used as a basis for comparison of dispersion structure in each of the four groups. It is apparent that breaking the sample into groups has had the effect of increasing within-groups multiple correlation coefficients, except in the case of Group 4. There a slight decrease is observed. The same pattern is exhibited in the PMCCs in terms of the increase in the number of high (> .3) values.

The extremely high multiple correlation coefficients (R^2_h) exhibited in Group 3 are of some concern. The possibility exists that the small sample size in this group relative to the number of descriptors (7 OTUs vs. 6 descriptors) has led to a nearly singular dispersion matrix and thus to R^2_h values approaching 1 (Rohlf 1977). On the other hand, Group 2 (with only 2 more OTUs) does not appear to be affected in this way. Presumably, the high pairwise correlations between a few of the descriptors in combination

with the small sample size are forcing the high R^2_h values. Thus, while strong dispersion structure does exist within the group it is probably not as strong as is indicated by the R^2_h values.

With respect to the descriptors, it appears that LFLEN and PELEN are contributing most strongly to dispersion structure in all except Group 2. There, structure is associated with PELEN, R3412 and BANG. THLEN and PTLEN generally contribute the least to dispersion structure.

PCA

Results from a PCA of the entire random sample based on the dispersion matrix of 6 ranged descriptors are presented (in a similar format as in the previous section) in Table 3.16. Results from PCAs of each of the four groups are similarly presented (Group 1-Table 3.17, Group 2-Table 3.18, Group 3 -Table 3.19, Group 4 -Table 3.20).

With respect to the entire sample analysis, Bartlett's test (Table 3.16b) indicates that only the last two components are indistinguishable. 84% of the variation in the sample is accounted for on the first four components (Table 3.16a). Descriptor PELEN appears to contribute most strongly to variation on the first component with lesser contributions from LFLEN, R3412 and BANG. Component 2 is most strongly affected by LFLEN and R3412. These results are supported in tables 3.16c and 3.16d, where the highest correlations and largest percentage variance accounted for are given by PELEN and BANG on the first component, and by LFLEN and R3412 on the second component.

Somewhat different patterns emerge from the PCAs of the groups. In Group 1 only the first two components, accounting for 65% of the variation are significant. On the first of these, LFLEN, PELEN and R3412 are all making equally strong contributions, while PELEN, R3412 and BANG contribute most heavily to variation on the second component.

Results from Group 2 indicate a highly unique pattern of variation. Only the first component (45% of the variation) is significantly different (Table 3.18a). By far the strongest loading on that component is given by PTLEN, although reference to Table 3.18c indicates that PELEN and BANG are also fairly strongly associated with this component. R3412, BANG and PELEN show the highest loading on the second component.

For Group 3, the highest loading on the first component is for PELEN (Table 3.19a). LFLEN is also strongly associated with this component (Table 3.19c), probably as a result of the high correlation between these two descriptors first evidenced in the Analysis of Dispersion Structure. Similarly, R3412 loads heavily on component 2 but both R3412 and THLEN are strongly associated with that component. BANG and PTLEN show a similar relationship on component 3.

In Group 4, all 6 components are significant (Table 3.20b). The first component is dominated by R3412 (Table 3.20a), the second by LFLEN and PELEN and the third by BANG.

The relative contribution of each descriptor to variation on the first two components in each group is graphically summarized in Figure 3.8. In this figure, each descriptor is represented by a line, the terminus of which

is given by the component weights for that descriptor. The scale is arbitrary but is consistent across groups.

In all but Group 2, PELEN and R3412 show the strongest contributions to variation on the first two components (with LFLEN usually assuming a secondary role). Interestingly, the effect of these two descriptors appears to be complementary, with spread in one approximately orthogonal to the other. Given the consistency of this angular relationship, it is tempting to postulate some common effect at work in the three groups. This matter will be dealt with in the Discussion.

MDA

Results from Multi-group Discriminant Analysis of the four groups are presented in Tables 3.21 and 3.22. Table 3.21 gives results from the MANOVA portion of the analysis. Bartlett's test of significance of the discriminant axes, discriminant weights and percentage variance accounted for by each axis are given in Tables 3.22a and 3.22b. Tables 3.22c and 3.22d give factor structure and total communalities for the analysis. In Figure 3.9, a plot of the (standardized) scores for the group centroids on the first two discriminant axes is presented. Superimposed on the plot are lines representing the contribution of each descriptor to discrimination among groups.

In the test of equality of group dispersion determinants, the null hypothesis was rejected. Rigorous statistical methodology, at this juncture, would require abandonment of the analysis on the basis of inequality of

group dispersions. However, since there is some evidence for the robustness of the technique (Cooley and Lohnes 1971), and since prime interest is directed towards revealing patterns of variation rather than making statistical inferences, the remainder of the analysis was carried out. Moreover, observation of the dispersion determinants (Table 3.21) indicates that the Group 3 determinant is several orders of magnitude smaller than the others. A second MDA was performed to examine the effect of removal of this group, the results of which will be presented subsequently.

The test of equality of group centroids led to rejection of the null hypothesis. Thus, at least one centroid was significantly different from the grand centroid. Bartlett's test (Table 3.22b) reveals that all three discriminant axes are making significant contributions to discriminance among groups.

Table 3.22 provides information on the behavior of each of the descriptors in the analysis. The discriminant weights reveal the strongest contribution by LFLEN on the first two axes. PELEN shows the next strongest weight on only the first axis. R3412 is apparently contributing the least to group discrimination. BANG, THLEN and PTLEN show moderate weight on the first two axes but are most heavily loaded on the third. These observations are supported by the factor structure coefficients (Table 3.22c), although PELEN, with low weighting on axis 2, is highly correlated with it. The table of total communalities (Table 3.22d) indicates that only R3412 has overall low importance in discriminating amongst groups.

In Figure 3.9, Group 2 shows the greatest separation from the other groups, apparently on the basis of large positive scores for LFLEN and R3412. Similarly, Group 4 appears to be separating on the basis of high scores for PTLEN, THLEN and BANG.

Multi-group Discriminant Analysis using only 3 groups (Group 3 removed) revealed suprisingly similar results, although reduction to three groups implied a maximum of 2 discriminant axes. These results are presented in Tables 3.23 and 3.24 and Figure 3.10. The removal of Group 3 did have the effect of reducing the level of statistical differences between group dispersion determinants, although the null hypothesis was still rejected ($P<.01$, Table 3.23). It also shifted the discriminant weights somewhat (Table 3.24a), as a result of the reduction to two axes. However, Figure 3.10 is virtually identical to Figure 3.9, except for a minor shift in the grand centroid and a sign change on axis 1.

3.3.3 Patterns of Variation at the Population Level

In the final phase of analysis of the random sample, groups were broken into populations according to site. For example, individuals in Group 1 ($n = 37$) were collected from two sites. Thus, Group 1 was divided into populations 1 ($n = 15$, individuals from site 4) and population 2 ($n = 22$, individuals from site 3). All individuals in Group 2 were collected from the same site and thus made up population 3 ($n = 9$). Group 3 presented a problem in that, of the seven individuals in the group, five were from site 4 and two from site 3. Splitting Group 3 into two populations would require either further reduction in the number of descriptors (and essentially meaningless

results for a population of two OTUs) or complete removal of the group from further consideration. Because neither of these choices was acceptable, the group was treated as a single population (population 4). Group 4 was split into population 5 (n = 40, site 2) and population 6 (n = 19, site 1).

3.3.3.1 Univariate Results

Results of the univariate analysis are presented in Table 3.25. As was found in the univariate analysis at the group level, no general reduction in standard deviations was observed. The lowest standard deviations were exhibited by populations 5 and 6 (from Group 4). The largest standard deviations were found in populations 1 and 2 (from Group 1). Only two minor departures from normality were observed, occurring in population 2 (descriptors LFLEN and BANG).

Figure 3.11 graphically summarizes the univariate results at all three levels of organization for the 6 descriptors. Ranged mean values for each group and population are usually in the range of 0.3 to 0.7 (with the clear exception of Group 2, LFLEN). Standard deviations are also similar, with ranges on the order of 0.4 to 0.5 (except in Group 3, PELEN). The clearest separation amongst groups is seen in descriptors LFLEN (distinguishing Group 2) and BANG (Group 3). Separation amongst populations is clear also in descriptors LFLEN (all populations), PELEN (populations 1 and 2), THLEN (populations 5 and 6) and PTLEN (populations 1 and 2).

3.3.3.2 Multivariate Results

Dispersion Structure

Results from the analysis of dispersion structure for the 6 populations are presented in Tables 3.26 to 3.29 (results for populations 3 and 4 are given in Tables 3.13 and 3.14).

Comparison of Tables 3.26 and 3.27 with Table 3.12 (Group 1) reveals some change in dispersion structure arising from the division of Group 1 into populations. Descriptors R3412 and THLEN display larger R^2_h values in population 1 and thus more dispersion structure is indicated. Otherwise, both populations display the same general structure as Group 1, from which they were derived. Populations 5 and 6 do not differ appreciably from Group 4 with respect to dispersion structure.

Population 4 (Group 3) displays the strongest dispersion structure, although difficulties with the small sample size still apply. Population 3 (Group 2) displays a relatively unique pattern of structure, differing from the other populations in the low R^2_h value for LFLEN and a relatively higher value for BANG.

With respect to descriptors, a pattern similar to that displayed at the group level is evident. LFLEN and PELEN generally make the strongest contribution to structure while THLEN and PTLEN make the least.

PCA

Principal Components Analysis results are presented in Tables 3.30 to 3.33. These tables give results for populations 1, 2, 5 and 6. Results for populations 3 and 4 are given in Tables 3.18 and 3.19. Figure 3.12 gives descriptor vector plots for the first two components for all six populations.

Comparison of significance of components across populations reveals that generally the first (populations 2, 3, 6) or first two components (population 1) are significant. Populations 4 and 5 show significance on many components. Populations 1, 2, 3 and 6 also show similar trends in percentage variance accounted for, with about 60% to 70% of the variance in the population accounted for in the first two components.

Reference to the tables of component weights and to Figure 3.12 reveals quite different patterns of descriptor loading on the first two components, amongst the populations. While descriptor R3412 makes strong contributions in all populations (as it did at the group level) the remainder of the descriptors are not subject to generalization. LFLEN and PELEN, which displayed high loading in previous analyses, make almost no contribution in population 5, and little in population 6. BANG, which showed only moderate loading previously, displays very heavy loading in both populations 5 and 6.

MDA

Results from this part of the population analysis are presented in Tables 3.34 and 3.35. The null hypothesis of equality of group dispersions

was again rejected (Table 3.34, .025<p<.05). As well, the centroids were found to contain significant differences (Table 3.34, $p < .01$) from the grand centroid. Bartlett's test revealed significance in all but the last two discriminant axes.

Not suprisingly, the patterns of factor loadings, structure coefficents and total communalities (Table 3.35) are very similar to those displayed in the MDA at the group level. Figure 3.13 gives the positions of the populations in the space of the first two discriminant axes as well as the relative contributions of each of the descriptors to discrimination. This figure is very close to the previous MDA plot (Figure 3.9) with the exception that the centroids of Groups 1 and 4 are now separated into their respective population centroids. The close proximity of the centroids of populations 1 and 2, along the axis of greatest group separation suggested the possibility that these two populations were statistically indistinguishable (equal dispersions and equal centroids). Conversely, the rather large separation of centroids for populations 5 and 6 indicated the likelihood of their statistical separation. To examine these possibilities, two MANOVAs were performed (results not shown). In both analyses, the dispersion determinants were equal ($p > .05$) but the centroids were significantly different ($p < .05$).

3.4 DISCUSSION

3.4.1 Patterns of Variation in the Random Sample

The pattern of variation being exhibited at the level of the entire random sample is clearly a pattern of grouping. That is, the behavior of the descriptors reflects the presence of well-defined groups in the sample. Analyses using the 29 descriptor data set demonstrate the predominance of the pubescence and gland (multi-state) descriptors in terms of high standard deviations, large contributions to dispersion structure and high loading on the first few principal components. With a PCA based on the correlation matrix (correcting for the large variances of the multi-state descriptors), a few additional descriptors (NSERR, LFLEN, PELEN, LFWID) also showed high loading. These additional descriptors also ranked high in the analysis of dispersion structure.

The importance of this set of descriptors in reflecting group differentiation is clear. The multistate descriptors PDPUB, HYPUB and CXGLN were invariant in at least one of the groups, indicating extremely high consistency in defining these groups. STPUB, LFPUB and NSERR showed reduced standard deviations when groups were formed. Of the remaining descriptors, LFLEN and PELEN were most important in group differentiation in the MDA (see below).

In the six descriptor set, the effect of grouping is not nearly as clear as in the 29 descriptor set. Dispersion structure is weaker, with LFLEN and PELEN contributing most to the structure that was present. However,

MDA results do indicate differentiation between the Groups with the same two descriptors loading most heavily on the first discriminant axis. The heavy loading of R3412 (which contributed least to group discrimination) in the PCA can be explained by its high standard deviation (Table 3.10) and its correlation with PELEN (Table 3.11).

3.4.2 Patterns of Variation in Groups

Once the random sample was divided into separate groups, somewhat different patterns emerged. These patterns can be characterized most easily in terms of the amount of variability exhibited by each group and by the descriptors which contributed most to that variability.

Group 1 exhibited the highest levels of variability. Standard deviations of the 6 descriptors were overall very high, especially in descriptor LFLEN. Strong dispersion structure was exhibited by mostly high R^2_h values, particularly for LFLEN and PELEN. These two descriptors, along wth R3412 played important roles in the PCA. The determinant of the dispersion matrix for this group was the highest of the four groups.

Group 2 showed fairly high levels of variability, but with quite different patterns of descriptor importance. Here, the magnitudes of the standard deviations were variable, with a low value for descriptor LFLEN, and high values for BANG and PTLEN. R^2_h values were particularly high for PELEN, R3412 and BANG. PCA loadings and component descriptor correlations showed a similar pattern with strong contributions from PTLEN and R3412 and, to a lesser extent, from PELEN and BANG. The dispersion

determinant for this group reflected high variability.

A very unusual pattern of variability was exhibited by Group 3. Here, again, standard deviations were variable with particularly high values for PELEN and R3412. Dispersion structure appeared to be extremely high, at least in part due to many large intercorrelations between the descriptors. This high degree of multiple correlation has also had the effect of reducing the dispersion determinant to nearly zero. PELEN and R3412 predominate in PCA loadings but LFLEN also showed high correlation with the first two components and large percentage variance accounted for, as a result of its high correlation with PELEN.

Group 4 displayed the lowest levels of variability in standard deviation magnitude and in dispersion structure. The dispersion determinant was relatively low. Of the variation present in the group, the greatest amount was explained by R3412, PELEN and (to a lesser degree) LFLEN, in terms of PCA loading.

If group 2 is temporarily excluded from consideration, it appears that a common pattern of descriptor variation can be identified. The biological factors giving rise to this pattern are unclear. Suffice it to say that marginally but generally higher variances in descriptors PELEN and R3412 in combination with high correlations between these and LFLEN have led to their major contributions in accounting for variation within Groups. In Group 2, the role of LFLEN is, in effect, replaced by BANG with respect to dispersion structure and by PTLEN in the PCA. Reference to Figure 3.16 provides a possible explanation for this substitution. Group 2 is

characterized by large and quite distinct LFLEN scores. The standard deviation of this descriptor is quite low. Conversely, variance in PTLEN is extremely high. BANG is also reasonably variable in this group. Thus the different pattern presented by the group is a function of its distinctness in descriptor LFLEN.

The major difference between all of the groups then becomes one of levels of variability. This is most clearly demonstrated by the large differences in sizes of dispersion determinants, and also in vastly different levels of dispersion structure. The explanation of these differences most probably lies in the reproductive biology of the series, as will be discussed in Chapter 6.

3.4.3 Patterns of Variation in Populations

In moving from analysis of groups to analysis of populations, only two groups (1 and 4) were subdivided. Thus, the two remaining groups (2 and 3) were treated both as groups and as populations. Because these two groups fit both the definition of group (distinct cluster found in the cluster analysis) and of population (members of a cluster that occur at the same sampling site, except group 3) it was felt that this treatment at both levels of organization was justified.

The patterns of variation that emerge at the population level are similar to those observed in the previous analysis. Populations 1 and 2 generally show the highest standard deviations and dispersion structure and populations 5 and 6 the lowest. A difference in pattern is exhibited,

however, in that the contributions of LFLEN and PELEN in PCA are diminished somewhat, and other descriptors show a consequent increase in importance. In population 2, PTLEN and THLEN display greater importance. Similarly, in populations 5 and 6, BANG and R3412 display extremely high loadings on the first two Principal Components. Otherwise, standard deviations, R^2_h values and dispersion determinants give essentially the same picture as that displayed at the group level. It should be noted, however, that dividing Groups 1 and 4 has had the effect of reducing the dispersion determinant, roughly by half in both cases.

Thus, again, the major difference between populations becomes one of levels of variability. As previously stated, the significance of this result will be dealt with in Chapter 6.

3.4.4 Other Studies

Results from previous morphometric studies of *Crataegus* have demonstrated a tendency towards forming well defined groups or morphotypes. Rickett (1936) studied *C. pruinosa* populations near Columbia, Missouri. He found significant differences between morphotypes (as differentiated by anther colour) in stamen and style number. Other descriptors, such as thorn and nutlet length, also displayed differences between morphotypes. Rickett also found high correlations between anther colour and calyx margination (CXGLN in this study) and between anther colour and leaf pubescence. The different morphotypes were found to occur at the same collection site. Later, Rickett (1937) showed parallel results for morphotypes in *C.* series *crus-galli* in the same area. Sinnott (1978) and

Sinnott and Phipps (1983) examined morphometric variation in *C.* series *Pruinosae* in Ontario. In that study, three strongly differentiated groups (each consisting of two or more morphotypes) were found via Cluster Analysis and PCA. Likewise, Dickinson (1983) and Dickinson and Phipps (1985 in press) discovered morphotype differentiation in *C. crus-galli s.l.* in Ontario. There, as was found by Rickett (1937), differentiation was most marked on the basis of stamen number with other descriptors correlated with stamen number. However, in that study, morphotypes tended to cluster according to local populations.

With respect to the presence of groups in the random sample, the results presented here are intermediate between those of Sinnott and Dickinson. None of Sinnott's groups were restricted to a single site. In this study, however, OTUs of Group 2 were found only at site 3. Dickinson's results showed clustering according to site in all groups. Moreover, the differentiation into morphotypes according to anther color and/or stamen number (and correlated descriptors) is not evident in series *Rotundifoliae* for those or any other descriptors. Nearly all of the *Rotundifoliae* in Ontario have ivory colored anthers, and only the two species rare in Ontario (neither included in this analysis) have 20 stamens (the others have 10). Thus the group structure exhibited in this study is probably not similar to that seen in previous studies. On the other hand, the strong contribution of PDPUB, HYPUB and CXGLN in differentiating groups (Cluster Analysis) and in explaining variation (PCA) in the random sample is mirrored in the results of Dickinson (1983). There, stamen and style number were closely associated with these descriptors in a variety of multivariate analyses using

the same methods used here.

3.4.5 Summary of Groups

Because a majority of the results and discussion in the ramainder of the thesis are presented in a format that reflects the group structure reported in this chapter (e.g. according to the four major groups and their respective populations), a summary of the salient morphological features characterizing each of the groups is presented below, in order to assist interpretation.

Group 1 This group consists of 37 OTUs from sites 3 and 4, and is characterized by generally pubescent leaves (LFPUB) with narrow apices (AANG), numerous leaf serrations (NSERR) and 1 or 2 pairs of glands on the petioles (PEGLN). Young stems are usually pubescent (STPUB). The flowers are small (PTLEN) with sparsely long-villous to glabrous pedicels and hypanthia (PDPUB, HYPUB). The fruits are small (FRWID) and quite red (FRCOB) with 2 or 3 nutlets (NNUTL).

Group 2 This is a noteworthy entity, comprising 9 OTUs from site 3. In this group, the leaves and petioles are larger than in the other groups (LFLEN, LFWID, PELEN) and pubescent (LFPUB), with many sharp serrations (TSERR, NSERR) and several pairs of glands on the petioles (PEGLN). The stems are mostly pubescent (STPUB). The hypanthia and pedicels are pubscent to densely so (HYPUB, PDPUB) and the calyx lobes have many glandular serrations (CXGLN). The fruits are the largest of the four

groups (FRWID) and dark red (FRCOB) with 3 or 4 long nutlets (NNUTL, NTLEN).

Group 3 In this group of 7 OTUs from sites 3 and 4, the leaves are the smallest of the four groups (LFLEN, LFWID) and obovate (RLOWP) with narrow bases (BANG). The leaves, stems, hypanthia and pedicels are all glabrous (LFPUB, STPUB, HYPUB, PDPUB), the thorns are somewhat more recurved (THCUR) than in the other groups and the fruits are small (FRWID) and greenish (FRCOA).

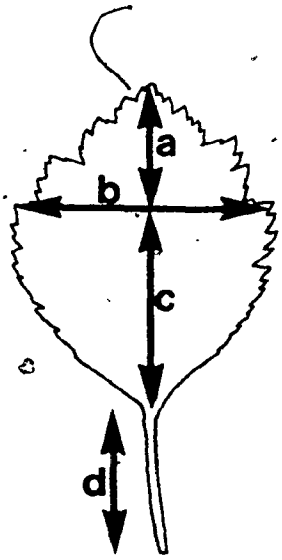
Group 4 This is the largest of the four groups in terms of numbers of OTUs collected (59), all of which were found at sites 1 and 2. No other groups were represented at these two sites. As in Group 3, OTUs of Group 4 are virtually glabrous (LFPUB, STPUB, HYPUB, PDPUB) and possess obovate leaves (RLOWP). This group differs from Group 3 in the possession of larger leaves (LFLEN, LFWID) with wider apices (AANG) and bases (BANG) and fewer, more crenate leaf serrations (NSERR, TSERR).

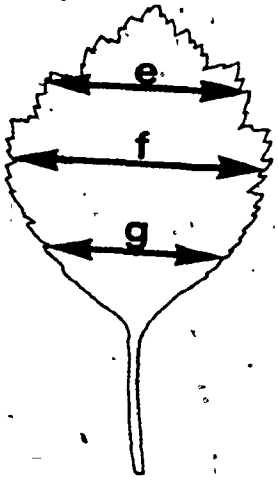
Table 3.1 Descriptors used in morphometric analysis of *C.* series *Rotundifoliae* individuals.

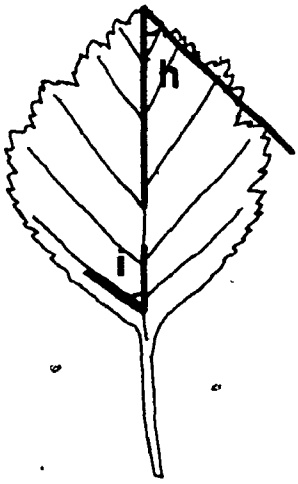
DESCRIPTOR	ABBREVIATION	NOTES
Continuous Descriptors (units)		
Leaf Length (cm.)	LFLEN	
Petiole Length (cm.)	PELEN	
Leaf Width (cm.)	LFWID	Measured at widest point
Angle of Leaf Base (deg.)	BANG	Midvein and 1st side vein
Angle of Leaf Apex (deg.)	AANG	Apex to widest point
Thorn Length (cm.)	THLEN	
Fruit Width (mm.)	FRWID	
Nutlet Length (mm.)	NTLEN	
Petal Length (mm.)	PTLEN	
Ratio Descriptors		
Length from Petiole to Widest Point	RLOWP	Relative to Total Length
Width at 3/4 Length	R3412	Relative to Width at 1/2 Length
Width at 1/4 Length	R1412	Relative to Width at 1/2 Length
Meristic Descriptors		
No. of Leaf Serrations	NSERR	Along 2 cm. distance from apex
No. of Nerves (secondary veins)	NNERV	Along one side of the leaf
No. of Petiole Glands	PEGLN	Pairs of glands counted
No. of Nutlets	NNUTL	Per fruit
Modal Stamen No.	MSTAN	Per flower
No. of Flowers	NFLWS	Per inflorescence
Multi-state Descriptors		
Leaf Pubescence	LFPUB	Glabrous to pubescent, 3 states
Shape of Leaf Base	BSHAP	Cuneate to Rounded, 3 states
Type of Leaf Serration	TSERR	Serrate to Crenate, 5 states
Thorn Curvature	THCUR	Straight to Recurved, 3 states
Thorn Color	THCOL	No Silver-Grey to Silver-Grey, 4 states
Stem Pubescence	STPUB	Glabrous to Pubescent, 3 states
Fruit Color (A)	FRCOA	No Green to Mostly Green, 3 states
Fruit Color (B)	FRCOB	Yellowish Red to Dark Red, 3 states
Pedicel Pubescence	PDPUB	Glabrous to Pubescent, 3 states
Hypanthium Pubescence	HYPUB	Glabrous to Pubescent, 3 states
Calyx Glands (margination)	CXGLN	None to Many, 3 states

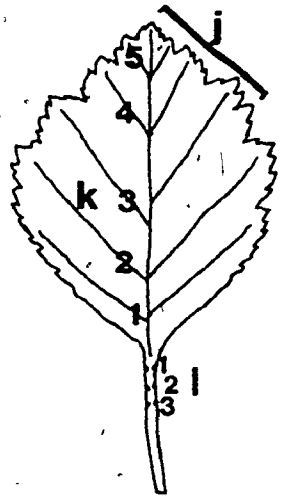
Figure 3.1a Illustration of selected leaf descriptors used in morphometric analyses of *Crataegus* series *Rotundifoliae* individuals. Descriptor abbreviations are explained in Table 3.1.

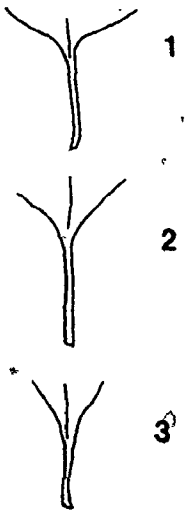
LFLEN – Line segments a+c
LFWID – Line segment b
PELEN – d
RLOWP – c/c+a
R3412 – e/f
R1412 – g/f
AANG – h
BANG – i
NSERR – j
NNERV – k
PEGLN – l
BSHAP – m (3 states shown)
TSERR – n (5 states shown)

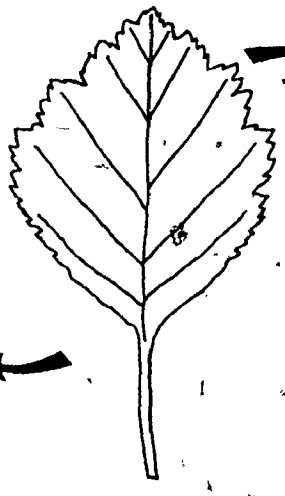

a
b
c
d

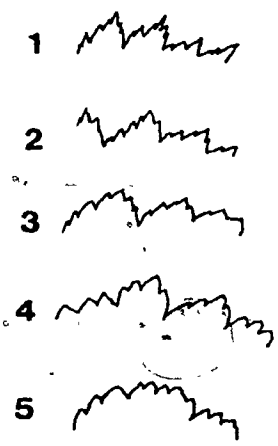

e
f
g


h
i


j
5
4
k
3
2
1
1
2
3
l


1
2
3


n
m


1
2
3
4
5

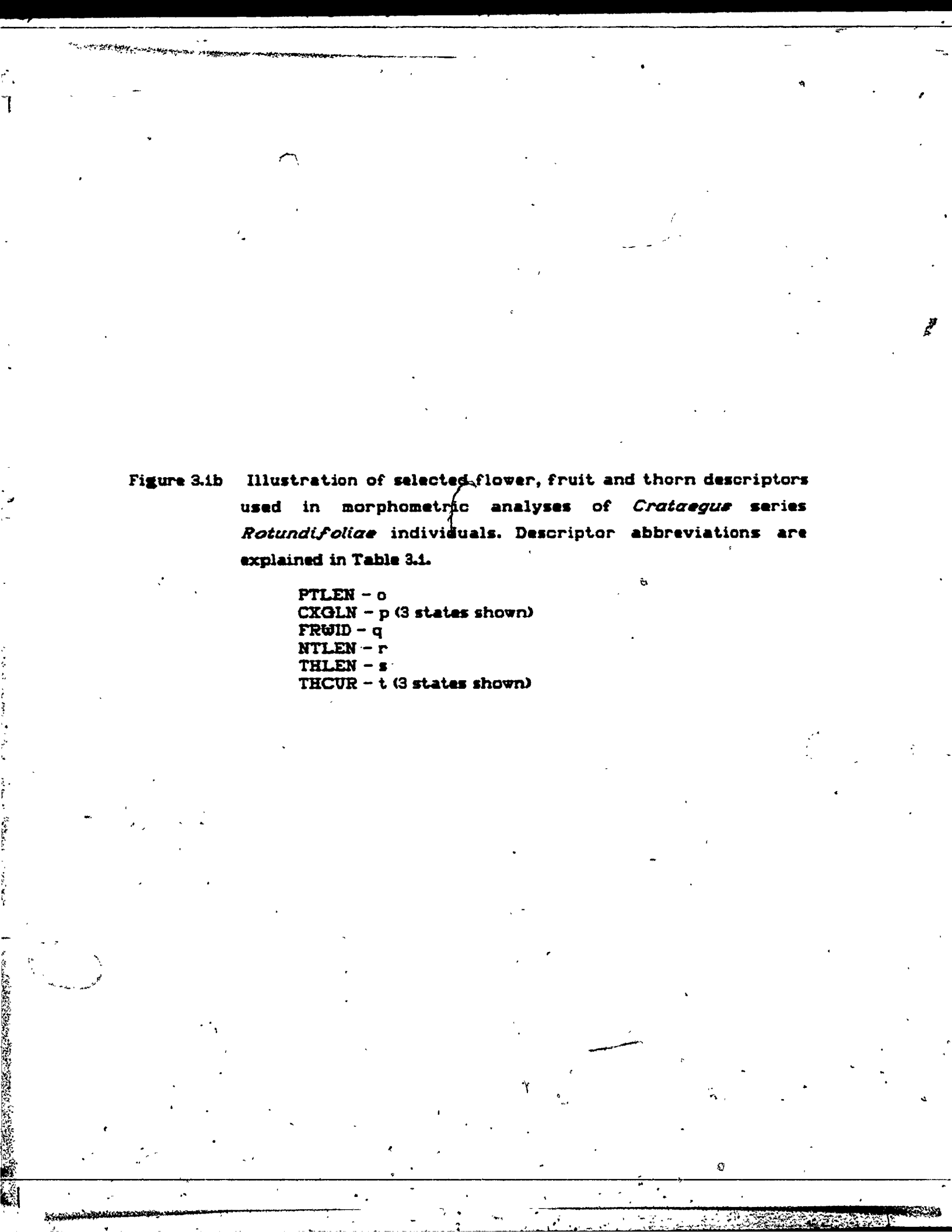
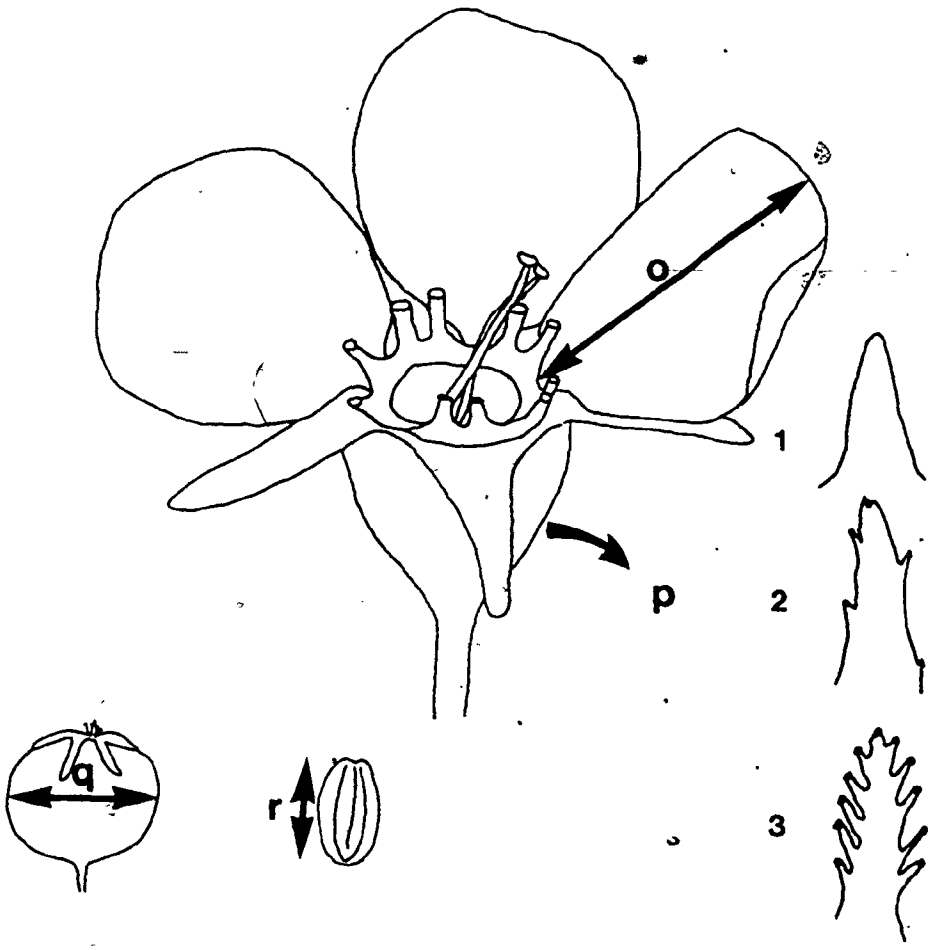


Figure 3.ib Illustration of selected flower, fruit and thorn descriptors used in morphometric analyses of *Crataegus* series *Rotundifoliae* individuals. Descriptor abbreviations are explained in Table 3.1.

PTLEN – o
CXGLN – p (3 states shown)
FRWID – q
NTLEN – r
THLEN – s
THCUR – t (3 states shown)


o
p
1
2
3
q
r
s
t
1
2
3

DESCRIPTOR	MEAN	STANDARD DEVIATION	C.V.
Continuous Descriptors			
LFLEN	4.15	0.45	0.11
PELEN	1.66	0.28	0.17
LFWID	3.17	0.37	0.12
BANG	36.19	3.38	0.09
AANG	40.24	4.05	0.10
THLEN	4.35	0.69	0.16
FRWID	9.22	0.88	0.10
NTLEN	5.59	0.94	0.17
PTLEN	6.36	0.72	0.11
Ratio Descriptors			
RLOWP	0.52	0.05	0.10
R3412	0.71	0.06	0.08
R1412	0.69	0.06	0.08
Meristic Descriptors			
NSERR	14.14	3.02	0.21
NNERV	4.06	0.54	0.13
PEGLN	1.64	0.73	0.45
NNUTL	2.86	0.31	0.11
MSTAN	7.07	2.11	0.30
NFLWS	8.85	2.02	0.23
Multi-state Descriptors			
LFPUB	1.76	0.89	0.51
BSHAP	1.88	0.68	0.36
TSERR	3.51	1.03	0.29
THCUR	1.37	0.55	0.40
THCOL	2.55	0.73	0.29
STPUB	1.72	0.90	0.52
FRCOA	1.67	0.58	0.34
FRCOB	3.11	0.89	0.29
PDPUB	1.71	0.91	0.53
HYPUB	1.62	0.85	0.53
CXGLN	2.08	0.60	0.29

Table 3.2 Basic descriptive statistics of raw scores for 29 descriptors, characterizing the entire sample (112 OTUs). Descriptor abbreviations are explained in Table 3.1.

DESCRIPTOR	MEAN	STANDARD DEVIATION	C.V.
Continuous Descriptors			
LFLEN	0.48	0.19	0.39
PELEN	0.51	0.22	0.43
LFWID	0.46	0.16	0.35
BANG	0.54	0.20	0.37
AANG	0.44	0.19	0.44
THLEN	0.57	0.17	0.29
FRWID	0.49	0.17	0.34
NTLEN	0.26	0.19	0.73
PTLEN	0.46	0.19	0.41
Ratio Descriptors			
RLOWP	0.37	0.18	0.48
R3412	0.39	0.20	0.52
R1412	0.55	0.22	0.44
Meristic Descriptors			
NSERR	0.42	0.25	0.60
NNERV	0.35	0.18	0.51
PEGLN	0.32	0.37	1.14
NNUTL	0.38	0.14	0.36
MSTAN	0.67	0.23	0.35
NFLWS	0.49	0.24	0.48
Multi-state Descriptors			
LFPUB	0.38	0.45	1.17
BSHAP	0.44	0.34	0.77
TSERR	0.63	0.26	0.41
THCUR	0.18	0.28	1.51
THCOL	0.52	0.24	0.47
STPUB	0.36	0.45	1.25
FRCOA	0.33	0.29	0.86
FRCOB	0.70	0.30	0.42
PDPUB	0.36	0.46	1.28
HYPUB	0.31	0.43	1.38
CXGLN	0.54	0.30	0.56

Table 3.3 Basic descriptive statistics of ranged scores for 29 descriptors, characterizing the entire sample (112 OTUs). Descriptor abbreviations are explained in Table 3.1.

Table 3.4 Analysis of Dispersion Structure for 29 ranged descriptors characterizing the entire sample. S_{hh} = total variance, S_{hs} = descriptor-specific variance, S_{hc} = common variance, R^2_h = multiple correlation coefficient, PMCC = product moment correlation coefficient (see text for details).

Table 3.5 **Principal Components Analysis of the dispersion matrix of 29 ranged descriptors characterizing the entire sample. Results for the first 8 Principal Components are presented**

Table 3.5 **Principal Components Analysis of the dispersion matrix of 29 ranged descriptors characterizing the entire sample. Results for the first 8 Principal Components are presented**

COMPONENT WEIGHTS

DESCRIPTORS	COMPONENTS 1	2	3	4	5	6	7	8
LFLEN	.046	.285	-.114	-.250	.120	.202	-.068	-.104
PELEN	.104	.071	-.294	-.286	.058	.098	.004	-.128
LFWID	.010	.220	-.046	-.254	.162	.147	.027	-.131
RLOWP	-.063	.093	-.288	.200	-.248	-.151	.146	.232
R3412	.059	-.047	-.014	.242	-.166	-.210	.015	-.048
R1412	-.097	.007	.205	.081	.192	-.050	.002	-.295
NSERR	.211	-.102	-.049	.200	-.132	-.018	.006	-.128
BANG	-.066	.082	.149	-.222	.016	.067	-.032	.146
AANG	-.102	.112	.042	.087	-.177	-.051	.187	.249
LFPUB	.401	-.149	.102	-.013	.076	.046	.469	.263
NNERV	.088	-.089	-.114	-.039	.185	-.013	.028	-.020
BSHAP	-.212	.150	.369	.225	.246	.364	.508	-.059
TSERR	-.186	-.065	.213	-.080	.027	-.023	.150	.012
PEGLN	.219	.370	-.189	.528	.421	-.071	-.007	-.213
THLEN	-.087	.043	-.028	.091	.014	.124	-.028	-.026
THCUR	.054	-.186	-.334	-.177	.300	-.510	.439	-.050
THCOL	.003	-.095	.052	-.224	-.338	.029	.348	-.622
STPUB	.426	-.107	.114	-.162	.022	.188	-.019	.150
FRWID	.006	.093	-.225	.108	-.223	.149	.036	-.031
FRCOA	-.100	-.325	-.426	.025	.188	.391	.004	.044
FRCOB	.199	.112	.305	-.172	.126	-.252	-.156	-.167
NNUTL	-.021	.156	-.022	.054	-.081	.006	-.086	-.178
NTLEN	.129	.080	-.137	.074	-.099	.199	.050	-.008
PTLEN	-.061	.161	-.202	-.018	-.032	.066	.087	.171
MSTAN	.073	.088	-.203	.128	-.386	.019	.166	-.178
NFLWS	.008	.021	-.041	-.024	.074	-.139	.038	.142
PDPUB	.435	-.010	.115	.014	-.077	.271	.042	.014
HYPUB	.398	.083	.107	.067	-.073	-.106	-.120	-.017
CXGLN	-.002	.618	-.092	-.250	-.114	-.110	.178	.194
EIGENVALUES	.993	.170	.150	.096	.090	.070	.067	.061
% VARIANCE	46.3%	7.9%	7.0%	4.5%	4.2%	3.3%	3.1%	2.1%
CUM. VAR.	46.3%	54.3%	61.3%	65.8%	70.0%	73.2%	76.4%	79.2%

Table 3.5a Component Weights (eigenvector elements), Component Roots (eigenvalues) and absolute and cumulative variance accounted for by each Principal Component.

COMPONENT-DESCRIPTOR CORRELATIONS

DESCRIPTORS	COMPONENTS 1	2	3	4	5	6	7	8
LFLEN	.242	.621	-.232	-.410	.189	.282	-.093	-.136
PELEN	.475	.135	-.520	-.406	.079	.112	.004	-.145
LFWID	.058	.562	-.112	-.489	.300	.240	.043	-.201
RLOWP	-.350	.215	-.062	.348	-.417	-.223	.212	.320
R3412	.290	-.096	-.027	.370	-.245	-.273	.019	-.058
R1412	-.444	.014	.366	.116	.265	-.061	.002	-.335
NSERR	.835	-.167	-.076	.246	-.158	-.019	.007	-.126
BANG	-.332	.171	.290	-.348	.024	.089	-.042	.181
AANG	-.526	.239	.084	.140	-.274	-.069	.251	.319
LFPUB	.894	-.138	.088	-.009	.051	.027	.272	.145
NNERV	.483	-.089	-.243	-.067	.308	-.018	.040	-.027
BSHAP	-.622	.182	.419	.205	.217	.283	.386	-.042
TSERR	-.717	-.104	.320	-.096	.031	-.023	.151	.011
PEGLN	.594	.416	-.200	.447	.344	-.051	-.005	-.143
THLEN	-.518	.106	-.065	.170	.026	.196	-.043	-.038
THCUR	.193	-.278	-.467	-.199	.326	-.487	.411	-.045
THCOL	.013	-.160	.082	-.284	-.415	.031	.369	-.628
STPUB	.941	-.098	.098	-.112	.015	.110	-.011	.082
FRWID	.034	.231	-.524	.202	-.402	.236	.056	-.045
FRCOA	-.344	-.467	-.573	.027	.196	.359	.004	.038
FRCOB	.667	.155	.396	-.179	.127	-.224	-.136	-.138
NNUTL	-.151	.474	-.063	.124	-.178	.012	-.163	-.322
NTLEN	.671	.172	-.277	.120	-.155	.274	.068	-.010
PTLEN	-.320	.350	-.412	-.030	-.051	.092	.189	.223
MSTAN	.309	.154	-.335	.169	-.493	.021	.183	-.188
NFLWS	.046	.051	-.094	-.044	.131	-.216	.057	.207
PDPUB	.949	-.009	.097	.009	-.051	.156	.024	.007
HYPUB	.930	.081	.098	.049	-.052	-.066	-.073	-.010
CXGLN	-.005	.847	-.182	-.258	-.114	-.097	.153	.159

Table 3.5b Correlation between each descriptor and each Principal Component.

PERCENTAGE VARIANCE ACCOUNTED FOR

DESCRIPTORS	COMPONENTS 1	2	3	4	5	6	7	8
LFLEN	5.86	38.59	5.39	16.85	3.58	7.94	0.86	1.85
PELEN	22.53	1.81	27.02	16.45	0.62	1.41	0.00	2.10
LFWID	0.34	31.62	1.24	23.91	9.01	5.78	0.19	4.03
RLOWP	12.29	4.62	0.39	12.09	17.38	4.98	4.51	10.24
R3412	8.44	0.93	0.07	13.72	5.98	7.46	0.04	0.34
R1412	19.70	0.02	13.38	1.34	7.05	0.38	0.00	11.24
NSERR	69.70	2.78	0.57	6.06	2.48	0.35	0.00	1.59
BANG	11.00	2.93	8.43	12.07	0.06	0.80	0.18	3.27
AANG	27.72	5.71	0.71	1.95	7.53	0.48	6.31	10.17
LFPUB	79.94	1.90	0.78	0.01	0.26	0.07	7.41	2.11
NNERV	23.35	0.79	5.92	0.45	9.49	0.04	0.16	0.07
BSHAP	38.65	3.31	17.57	4.22	4.72	8.00	14.92	0.18
TSERR	51.48	1.08	10.22	0.91	0.10	0.05	2.28	0.01
PEGLN	35.30	17.32	3.98	19.96	11.85	0.26	0.00	2.05
THLEN	26.86	1.13	0.41	2.87	0.06	3.85	0.19	0.14
THCUR	3.74	7.76	21.79	3.96	10.62	23.74	16.90	0.20
THCOL	0.02	2.56	0.67	8.07	17.22	0.10	13.61	39.44
STPUB	88.54	0.96	0.96	1.25	0.02	1.22	0.01	0.68
FRWID	0.11	5.35	27.47	4.06	16.20	5.57	0.31	0.21
FRCOA	11.85	21.79	32.79	0.07	3.83	12.89	0.00	0.14
FRCOB	44.46	2.41	15.66	3.19	1.61	5.01	1.85	1.92
NNUTL	2.28	22.44	0.40	1.55	3.17	0.01	2.66	10.40
NTLEN	44.96	2.97	7.70	1.44	2.39	7.53	0.46	0.01
PTLEN	10.21	12.26	16.96	0.09	0.26	0.85	1.41	4.96
MSTAN	9.53	2.37	11.22	2.84	24.32	0.04	3.34	3.52
NFLWS	0.21	0.26	0.88	0.20	1.71	4.66	0.33	4.28
PDPUB	89.97	0.01	0.94	0.01	0.26	2.45	0.06	0.00
HYPUB	86.50	0.65	0.95	0.24	0.26	0.43	0.53	0.00
CXGLN	0.00	71.76	1.39	6.66	1.29	0.93	2.33	2.53

Table 3.5c Percentage descriptor variance accounted for by each Principal Component

BARTLETT'S TEST

LAST *n* COMPONENTS	χ^2	D.F.	SIGNIFICANCE
5	35.1	14	***
4	24.0	9	***
3	10.7	5	NS
2	7.0	2	NS

***	P < .01
**	P < .025
*	P < .05
NS	No Significant Difference

Table 3.6 Bartlett's test of the equality of the last *n* (= 5 to-2) Principal Components, in PCA of the dispersion matrix of 29 ranged descriptors characterizing the entire sample. χ^2 = chi-square test statistic, D.F. = degrees of freedom.

2

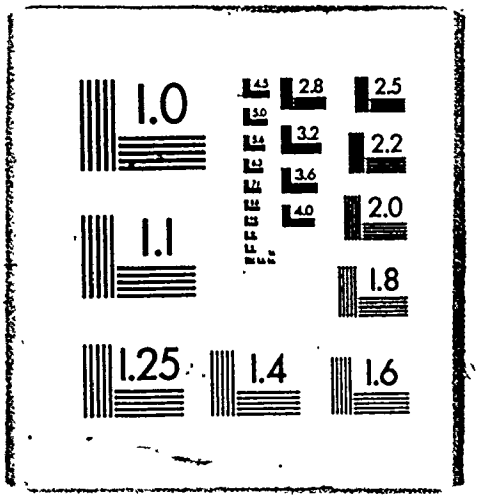


Table 3.7 Principal Components Analysis of the correlation matrix of 29 ranged descriptors characterizing the entire sample. Results for the first 8 Principal Components are presented

COMPONENT WEIGHTS

DESCRIPTORS	COMPONENTS 1	2	3	4	5	6	7	8
LFLEN	.103	.443	-.243	.090	.034	.027	-.017	.054
PELEN	.192	.227	-.038	.315	.062	-.052	.240	.024
LFWID	.037	.390	-.349	.131	.047	.047	.152	-.029
RLOWP	-.132	.088	.344	-.267	-.324	-.062	.136	.055
R3412	.104	-.112	.220	-.156	-.058	.328	.425	-.125
R1412	-.162	-.088	-.202	-.061	.142	.449	.138	.061
NSERR	.287	-.108	.162	-.088	.085	.056	-.063	.141
BANG	-.139	.066	-.244	-.102	-.098	-.407	-.202	.176
AANG	-.201	.106	.152	-.219	-.233	-.226	-.003	-.207
LFPUB	.292	-.118	-.027	-.054	-.069	-.109	-.049	-.058
NNERV	.186	-.003	-.059	.261	-.188	.138	.048	.060
BSHAP	-.228	.019	-.131	-.003	.185	-.033	-.010	.114
TSERR	-.262	-.093	-.115	-.035	.040	-.058	.144	.014
PEGLN	.206	.140	.025	-.105	-.091	.282	-.163	.028
THLEN	-.184	.111	.083	.026	.070	.153	-.294	.421
THCUR	.086	-.046	.078	.385	-.222	-.047	.287	-.062
THCOL	.006	-.031	.019	-.046	.412	-.453	.418	.194
STPUB	.310	-.074	-.075	-.053	-.009	-.107	-.119	-.031
FRWID	.029	.289	-.408	.003	.206	-.057	-.161	.104
FRCOA	-.096	-.047	.209	.498	.087	.016	-.121	.067
FRCOB	.215	-.049	-.270	-.228	-.099	-.085	.025	.185
NNUTL	-.042	.277	.039	-.244	.270	.218	.119	.226
NTLEN	.239	.175	.175	-.039	.113	-.001	-.231	-.028
PTLEN	-.101	.310	.173	.122	-.171	-.013	-.118	-.299
MSTAN	.123	.135	.301	-.077	.128	-.064	.245	.162
NFLWS	.019	.045	.017	.058	-.548	.010	.138	.613
PDPUB	.314	-.041	-.034	-.106	-.004	-.055	-.107	-.002
HYPUB	.309	-.027	-.035	-.156	-.040	.058	-.007	-.084
CXGLN	.001	.411	-.093	-.184	-.159	-.040	.174	-.229
EIGENVALUES	8.48	3.15	2.40	2.06	1.43	1.28	1.14	0.94
% VARIANCE	29.2	10.9%	8.3%	7.1%	4.9%	4.4%	3.9%	3.2%
CUM. VAR.	29.2%	40.1%	48.4%	55.5%	60.4%	64.9%	68.8%	72.0%

Table 3.7a Component Weights (eigenvector elements), Component Roots (eigenvalues) and absolute and cumulative variance accounted for by each Principal Component.

COMPONENT-DESCRIPTOR CORRELATIONS

DESCRIPTORS	COMPONENTS 1	2	3	4	5	6	7	8
LFLEN	.300	.787	-.377	.129	.041	.030	-.018	.052
PELEN	.562	.403	-.059	.452	.074	-.059	.257	.023
LFWID	.109	.693	-.541	.189	.057	.053	.163	-.028
RLOWP	-.383	.156	.533	-.383	-.388	-.070	.146	.053
R3412	.302	-.197	.342	-.224	-.070	.372	.454	-.121
R1412	-.472	-.157	-.312	-.088	.170	.508	.148	.059
NSERR	.837	-.193	.252	-.126	.102	.064	-.068	.137
BANG	-.389	.118	-.378	-.147	-.117	-.461	-.216	.171
AANG	-.586	.189	.236	-.314	-.279	.255	-.004	-.201
LFPUB	.852	-.197	-.042	-.077	-.083	-.123	-.052	-.056
NNERV	.540	-.005	-.092	.374	-.225	.156	.051	.058
BSHAP	-.666	.034	-.204	-.193	-.004	.209	-.035	-.010
TSERR	-.764	-.165	-.178	-.050	.047	-.065	.154	.013
PEGLN	.599	.249	.039	-.152	-.108	.320	-.174	.027
THLEN	-.537	.196	.128	.038	.084	.173	-.314	.408
THCUR	.251	-.081	.122	.553	-.266	-.053	.306	-.060
THCOL	.017	-.056	.029	-.066	.493	-.513	.447	.188
STPUB	.902	-.131	-.116	-.076	-.011	-.121	-.127	-.030
FRWID	.083	.513	.631	.004	.246	-.065	-.172	.101
FRCOA	-.281	-.084	.324	.715	.104	.018	-.129	.065
FRCOB	.628	-.086	-.418	-.327	-.118	-.096	.027	.180
NNUTL	-.122	.492	.060	-.350	.323	.246	.127	.220
NTLEN	.695	.310	.271	-.056	.135	-.001	-.247	-.027
PTLEN	-.294	.550	.268	.176	-.205	-.014	-.126	-.290
MSTAN	.358	.241	.466	-.111	.153	-.073	.262	.157
NFLWS	.055	.079	.027	.082	-.656	.011	.148	.595
PDPUB	.914	-.073	-.053	-.152	-.005	-.062	-.114	-.002
HYPUB	.901	-.049	-.054	-.224	-.047	.066	-.008	-.082
CXGLN	.003	.730	-.144	-.264	-.191	-.045	.186	-.223

Table 3.7b Correlation between each descriptor and each Principal Component.

PERCENTAGE VARIANCE ACCOUNTED FOR

	COMPONENTS							
DESCRIPTORS	1	2	3	4	5	6	7	8
LFLEN	9.01	61.90	14.22	1.65	0.17	0.09	0.03	0.03
PELEN	31.55	16.23	0.35	20.41	0.54	0.35	6.59	0.05
LFWID	1.18	48.02	29.27	3.56	0.32	0.28	2.65	0.08
RLOWP	14.69	2.44	28.38	14.69	15.06	0.49	2.12	0.28
R3412	9.14	3.90	11.66	5.00	0.48	13.82	20.60	1.46
R1412	22.31	2.47	9.76	0.77	2.87	25.80	2.18	0.35
NSERR	70.01	3.71	6.34	1.60	1.04	0.40	0L46	1.88
BANG	15.10	1.40	14.30	2.16	1.38	21.26	4.67	2.92
AANG	34.28	3.58	5.55	9.86	7.78	6.52	0.00	4.03
LFPUB	72.60	3.87	0.18	0.59	0.69	1.52	0.28	0.32
NNERV	29.18	0.00	0.84	14.02	5.07	2.43	0.26	0.34
BSHAP	44.30	0.12	4.14	3.73	0.00	4.38	0.12	0.01
TSERR	58.48	2.72	3.15	0.25	0.22	0.42	2.38	0.02
PEGLN	35.92	6.20	0.15	2.30	1.18	10.22	3.02	0.07
THLEN	28.82	3.86	1.64	0.14	0.71	3.00	9.86	16.66
THCUR	6.29	0.66	1.48	30.54	7.08	0.28	9.38	0.36
THCOL	0.03	0.31	0.08	0.44	24.35	26.31	19.94	3.54
STPUB	81.31	1.72	1.35	0.58	0.01	1.46	1.62	0.09
FRWID	0.69	26.31	39.87	0.00	6.05	0.42	2.96	1.01
FRCOA	7.89	0.71	10.49	51.18	1.08	0.03	1.67	0.42
FRCOB	39.40	0.76	17.49	10.68	1.39	0.92	0.07	3.23
NNUTL	1.48	24.19	0.36	12.21	10.43	6.07	1.61	4.83
NTLEN	48.36	9.62	7.34	0.31	1.81	0.00	6.09	0.07
PTLEN	8.66	30.25	7.16	3.08	4.19	0.02	1.58	8.43
MSTAN	12.83	5.79	21.74	1.22	2.34	0.53	6.84	2.47
NFLWS	0.30	0.63	0.07	0.68	43.05	0.01	2.18	35.35
PDPUB	83.48	0.53	0.28	2.33	0.00	0.38	1.30	0.00
HYPUB	81.14	0.24	0.29	5.01	0.22	0.43	0.00	0.67
CXGLN	0.00	53.34	2.08	7.00	3.63	0.20	3.45	4.96

Table 3.7c Percentage descriptor variance accounted for by each Principal Component

Figure 3.2 Phenogram resulting from Cluster Analysis of the random sample (112 OTUs) based on 29 ranged descriptors. The RM-CM combination used was Sums of Squares Agglomeration of Euclidean Distances. Scale at left represents the resemblance level (Euclidean Distance) at which groups of OTUs fused. OTUs are represented by symbols identifying the four groups into which the sample was ultimately divided.

Symbols:
Group 1 - closed squares
Group 2 - closed triangles
Group 3 - open squares
Group 4 - closed circles

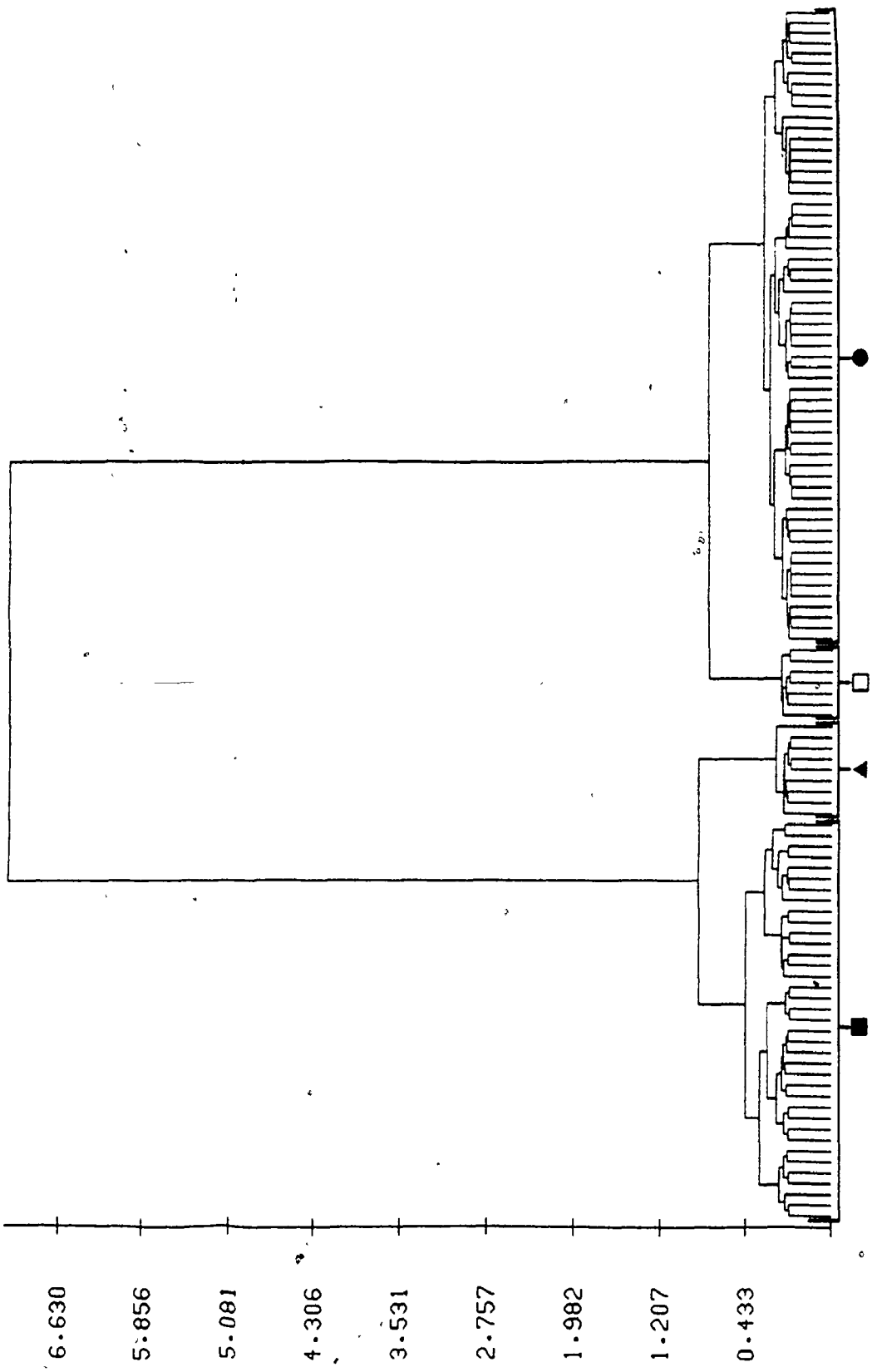

6.630
5.856
5.081
4.306
3.531
2.757
1.982
1.207
0.433

Figure 3.3 Phenogram resulting from Cluster Analysis of the random sample (112 OTUs) based on 29 ranged descriptors. The RM-CM combination used was Unweighted Pairs Grouped by Mathematical Averages of Euclidean Distances. Scale at left represents the resemblance level (Euclidean Distance) at which groups of OTUs fused. OTUs are represented by symbols identifying the four groups into which the sample was ultimately divided.

Symbols:
Group 1 - closed squares
Group 2 - closed triangles
Group 3 - open squares
Group 4 - closed circles

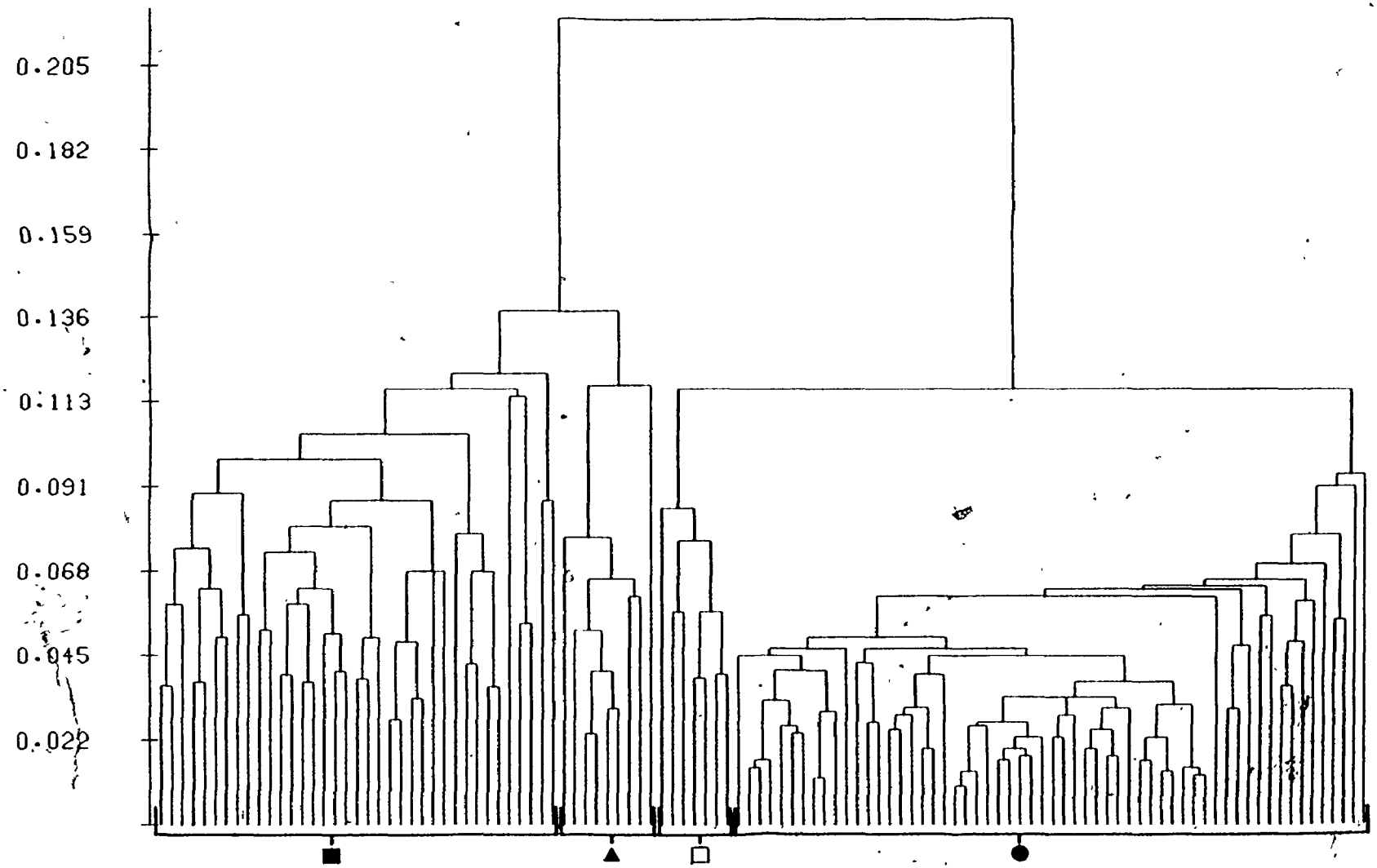

0.205
0.182
0.159
0.136
0.113
0.091
0.068
0.045
0.022

Figure 3.4 Phenogram resulting from Cluster Analysis of the random sample (112 OTUs) based on 29 ranged descriptors. The RM-CM combination used was Single Linkage of Euclidean Distances. Scale at left represents the resemblance level (Euclidean Distance) at which groups of OTUs fused. OTUs are represented by symbols identifying the four groups into which the sample was ultimately divided.

Symbols:
Group 1 - closed squares
Group 2 - closed triangles
Group 3 - open squares
Group 4 - closed circles

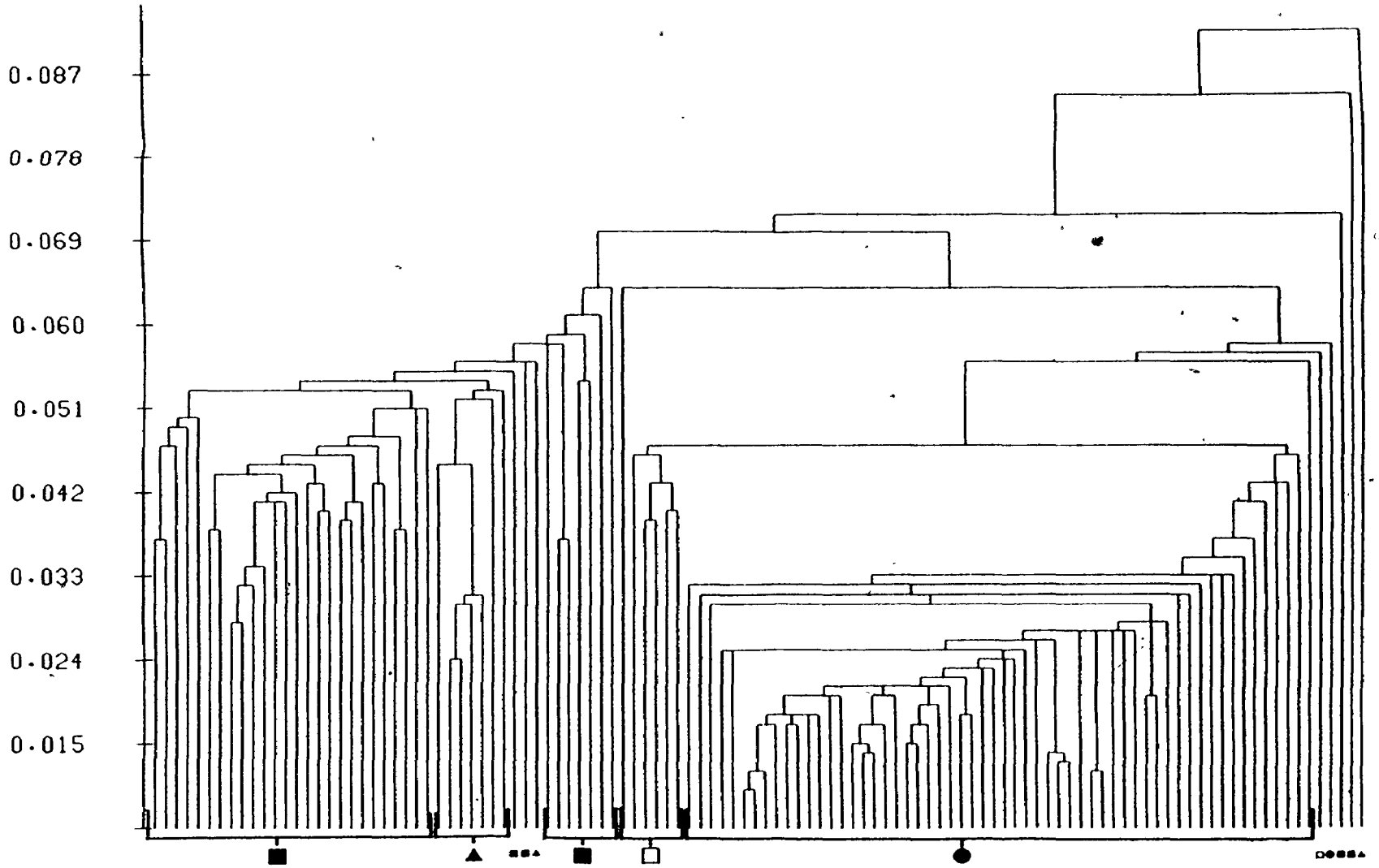

0.087
0.078
0.069
0.060
0.051
0.042
0.033
0.024
0.015

Figure 3.5 Phenogram resulting from Cluster Analysis of the random sample (112 OTUs) based on 29 ranged descriptors. The RM-CM combination used was Unweighted Pairs Grouped by Mathematical Averages of Centered Cosines. Scale at left represents the resemblance level (Centered Cosines) at which groups of OTUs fused. OTUs are represented by symbols identifying the four groups into which the sample was ultimately divided.

Symbols:
Group 1 - closed squares
Group 2 - closed triangles
Group 3 - open squares
Group 4 - closed circles

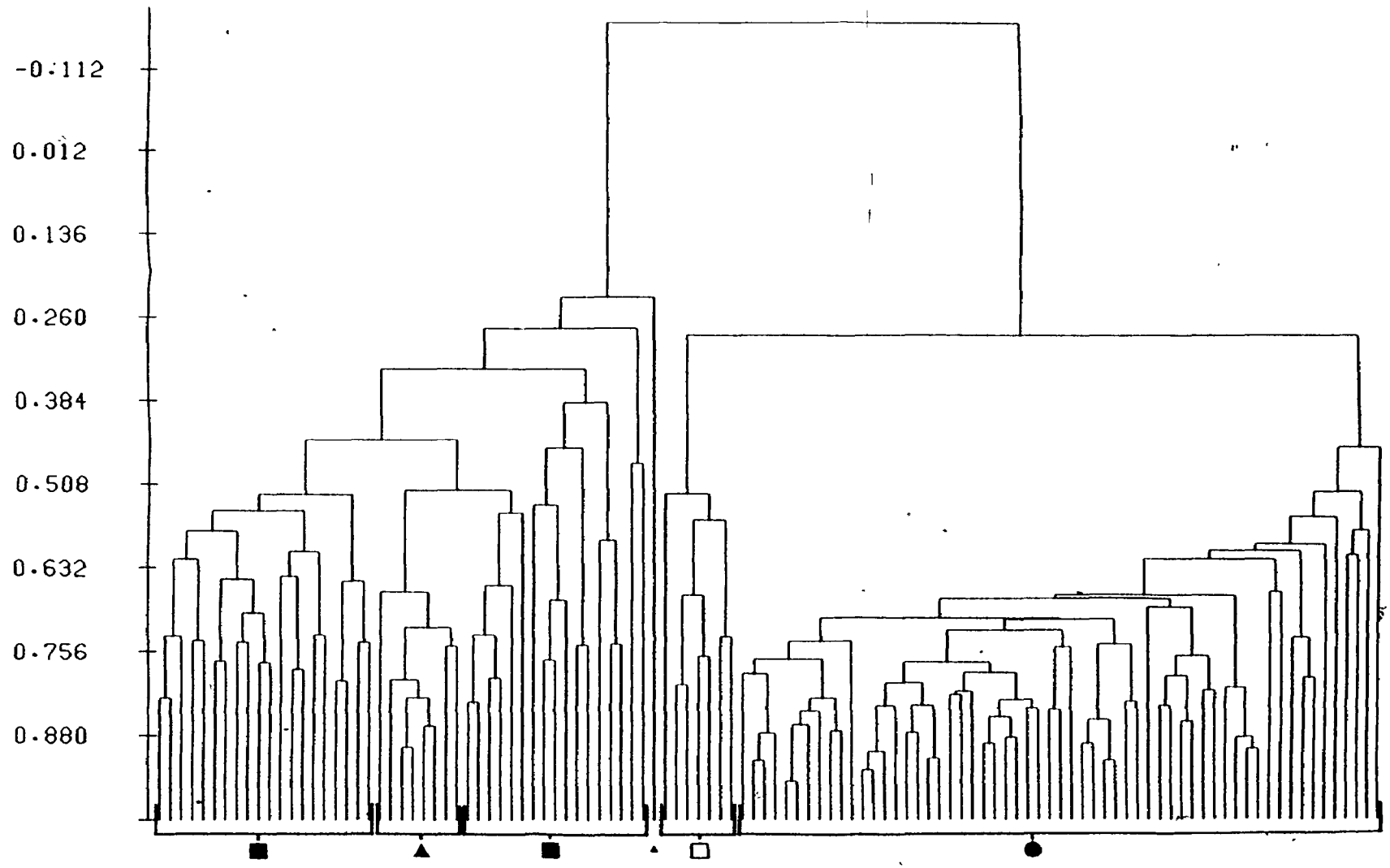
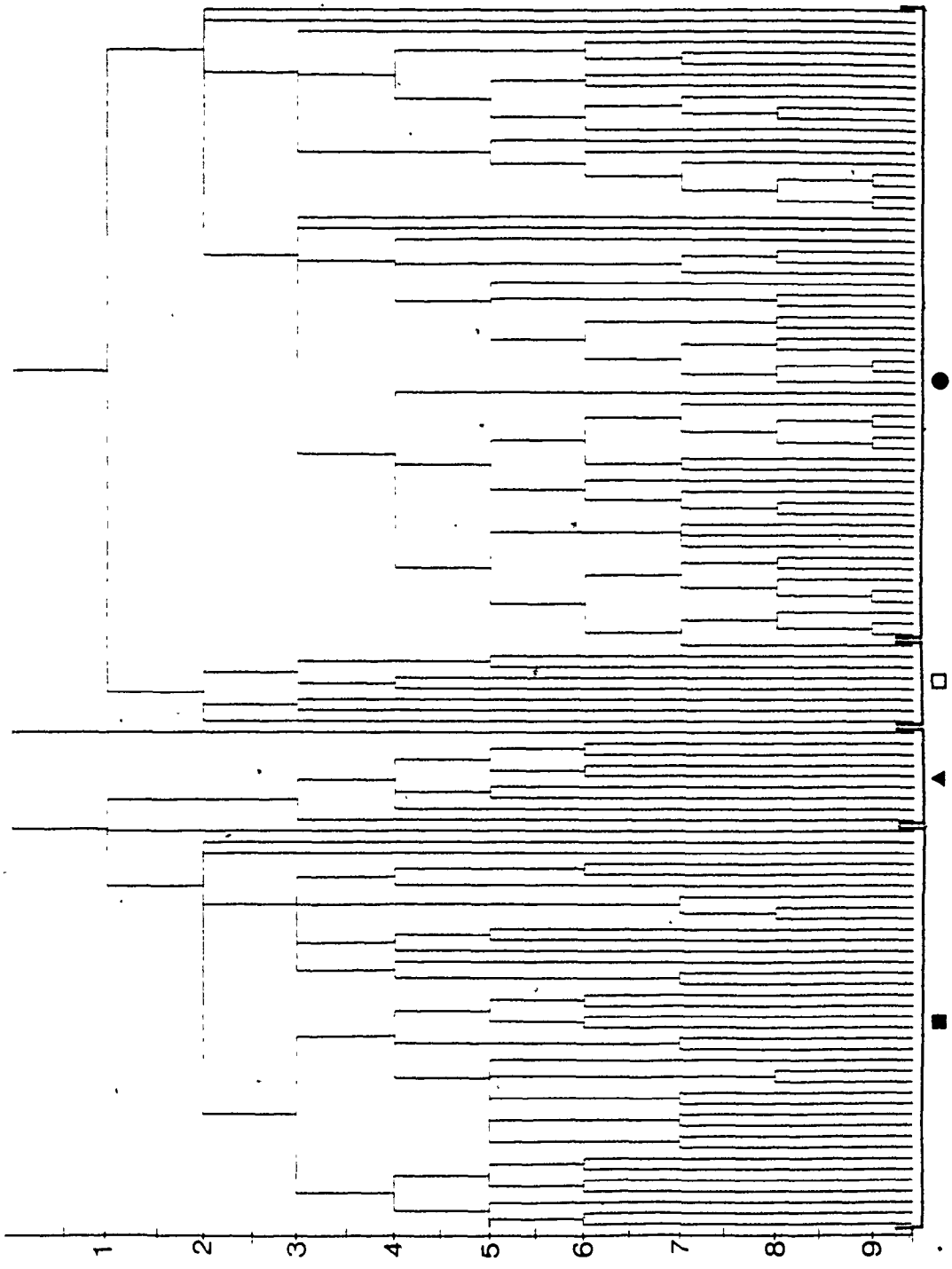

-0.112
0.012
0.136
0.260
0.384
0.508
0.632
0.756
0.880

Figure 3.6 Durchschnitt consensus tree, illustrating the information about group structure shared between the four previous phenograms. Scale at left represents the fusion level at which consensus clusters were formed (see text and Smith and Phipps 1984 for details). OTUs are represented by symbols identifying the four groups into which the sample was ultimately divided. The actual sequence of OTU numbers in this figure are given in Appendix 3.

Symbols:
Group 1 - closed squares
Group 2 - closed triangles
Group 3 - open squares
Group 4 - closed circles


1
2
3
4
5
6
7
8
9

Resemblance Matrix - Phenogram Comparisons

RM-CM Combination	Coph. Corr. Coeff.
EUCD - SL	.796
EUCD-UPGMA	.874
EUCD - SSA	.835
CC - UPGMA	.918

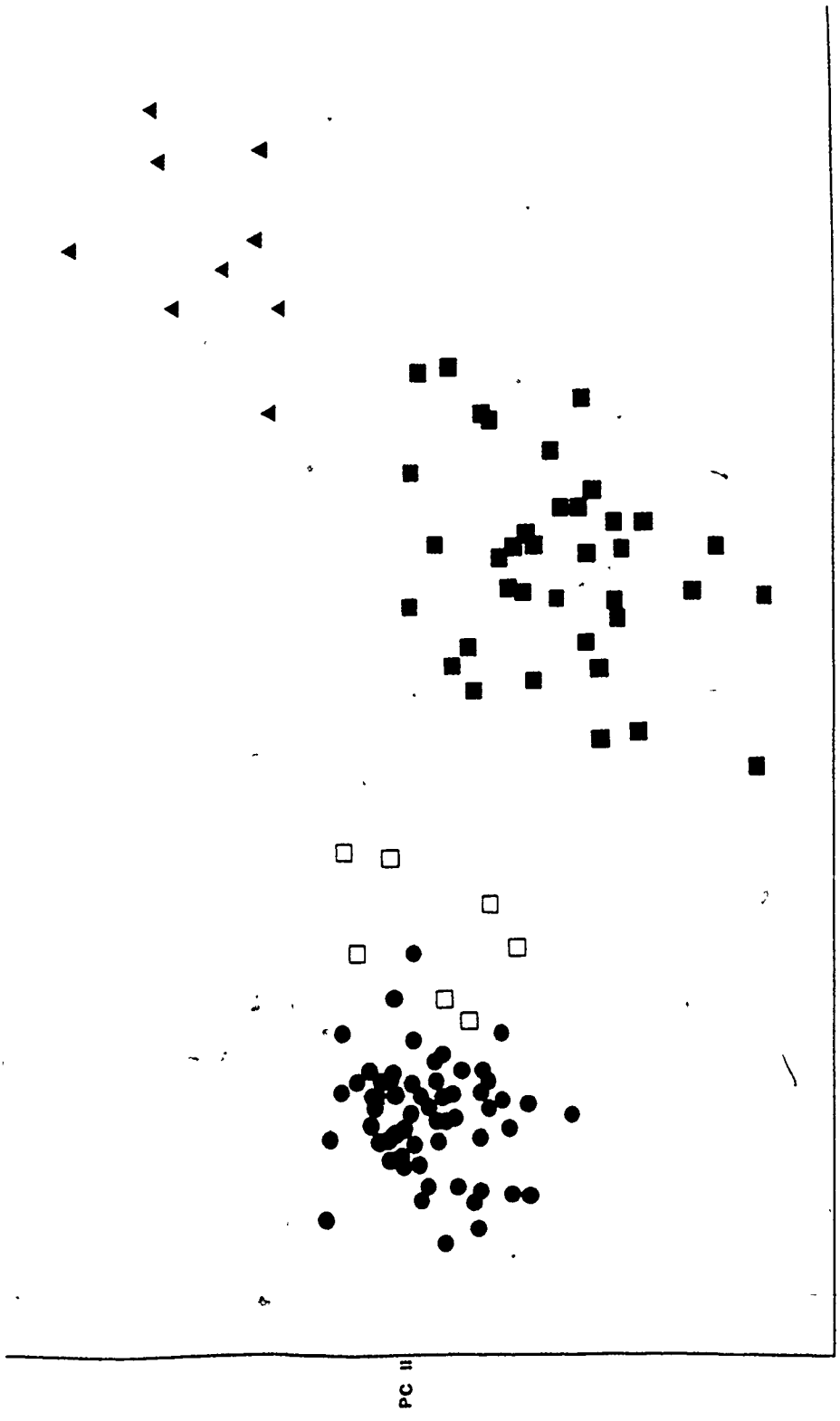
Phenogram - Phenogram Comparisons

	EUCD - UPGMA	EUCD - SSA	CC - UPGMA
EUCD - SL	.902	.791	-.888
EUCD - UPGMA		.955	-.985
EUCD - SSA			-.957

Table 3.8 Cophenetic Correlation Coefficients between resemblance matrices and phenograms produced from them and between different phenograms for each of the four RM-CM combinations used in Cluster Analysis of the entire sample. EUCD = Euclidean Distances, CC = Centered Cosines, SL = Single linkage Clustering, UPGMA = Unweighted Pairs Grouped by Mathematical Averages, SSA = Sums of Squares Agglomeration.

Figure 3.7 Scatter diagram of the position of each individual in the random sample (112 OTUs) in the space of the first two Principal Components (PC I and II) of a Principal Components Analysis of the correlation matrix (29 ranged descriptors).

Symbols:
Group 1 - closed squares
Group 2 - closed triangles
Group 3 - open squares
Group 4 - closed circles


PC I
PC II

GROUP 1 Symbol- ■

PS101, PS103, PS115, PS119, PS161, PS162, PS163, PS164, PS165, PS167, PS169, PS170, PS172, PS174, PS177, PS178, PS179, PS183, PS187, PS188, PS189, PS191, PS195, PS197, PS198, PS200, PS201, PS202, PS203, PS204, PS205, PS206, PS208, PS209, PS210, PS211, PS215

GROUP 2 Symbol- ▲

PS158, PS159, PS160, PS175, PS176, PS181, PS182, PS196, PS214

GROUP 3 Symbol- □

PS104, PS141, PS168, PS190, PS192, PS207, PS216

GROUP 4 Symbol- ●

PS218, PS219, PS220, PS221, PS222, PS223, PS224, PS225, PS226, PS227, PS228, PS229, PS230, PS231, PS232, PS233, PS235, PS236, PS237, PS238, PS239, PS240, PS241, PS242, PS243, PS244, PS245, PS246, PS247, PS248, PS249, PS250, PS251, PS252, PS253, PS254, PS255, PS256, PS257, PS258, PS259, PS261, PS262, PS263, PS264, PS265, PS266, PS267, PS268, PS269, PS270, PS271, PS272, PS273, PS274, PS275, PS276, PS277, PS278

Table 3.9 Assignment of the 112 OTUs making up the entire sample into the four groups discovered via Cluster Analysis. Symbols by which the groups are identified in all figures are also indicated.

		Entire Sample	Group 1	Group 2	Group 3	Group 4
LFLEN*	Mean	.480	.436	.880	.298	.469
	S.D.	.189	.189	.117	.157	.108
PELEN*	Mean	.510	.605	.709	.615	.407
	S.D.	.219	.218	.191	.344	.144
LFWID	Mean	.464	.425	.662	.304	.477
	S.D.	.162	.156	.165	.186	.131
RLOWP	Mean	.370	.253	.389	.434	.433
	S.D.	.179	.205	.118	.179	.128
R3412*	Mean	.394	.452	.484	.490	.334
	S.D.	.203	.224	.171	.241	.171
R1412	Mean	.550	.457	.389	.483	.641
	S.D.	.217	.197	.293	.183	.180
NSERR	Mean	.417	.644	.784	.498	.210
	S.D.	.252	.132	.137	.148	.084
BANG*	Mean	.540	.483	.418	.227	.632
	S.D.	.199	.167	.189	.129	.164
AANG	Mean	.441	.293	.376	.360	.553
	S.D.	.193	.193	.077	.180	.126
LFPUB	Mean	.379	.851	.778	.143	.051
	S.D.	.446	.285	.363	.244	.152
NNERV	Mean	.354	.450	.481	.428	.265
	S.D.	.181	.161	.242	.163	.135
BSHAP	Mean	.442	.202	.222	.285	.644
	S.D.	.341	.249	.264	.267	.279
TSERR	Mean	.627	.486	.167	.500	.801
	S.D.	.258	.186	.177	.289	.121

Table 3.10 Summary of univariate means and standard deviations of 29 ranged descriptors for the entire sample and for each of the four groups discovered via Cluster Analysis. Asterisks indicate those descriptors used in the reduced data set (see text).

		Entire Sample	Group 1	Group 2	Group 3	Group 4
PEGLN	Mean	.321	.432	.944	.214	.169
	S.D.	.367	.376	.167	.267	.256
THLEN*	Mean	.574	.455	.528	.643	.646
	S.D.	.167	.163	.123	.142	.133
THCUR	Mean	.183	.270	.111	.433	.110
	S.D.	.277	.303	.220	.178	.247
THCOL	Mean	.518	.558	.481	.479	.503
	S.D.	.245	.295	.294	.182	.209
STPUB	Mean	.362	.865	.833	.004	.017
	S.D.	.451	.225	.354	.011	.091
FRWID	Mean	.494	.428	.723	.733	.473
	S.D.	.166	.144	.124	.119	.135
FRCOA	Mean	.335	.283	.000	.928	.347
	S.D.	.288	.251	.000	.189	.232
FRCOB	Mean	.702	.928	1.00	.048	.593
	S.D.	.298	.195	.000	.126	.186
NNUTL	Mean	.377	.286	.590	.370	.402
	S.D.	.136	.115	.101	.143	.104
NTLEN	Mean	.263	.330	.707	.341	.144
	S.D.	.192	.127	.183	.107	.076
PTLEN*	Mean	.464	.333	.585	.613	.511
	S.D	.189	.199	.246	.205	.114
MSTAN	Mean	.675	.721	.901	.762	.601
	S.D.	.235	.184	.067	.175	.254
NFLWS	Mean	.492	.490	.500	.339	.510
	S.D.	.236	.254	.313	.187	.216
PDPUB	Mean	.357	.838	1.00	.000	.000
	S.D.	.457	.265	.000	.000	.000
HYPUB	Mean	.308	.703	.944	.000	.000
	S.D.	.426	.343	.167	.000	.000
CXGLN	Mean	.540	.392	1.00	.214	.602
	S.D.	.301	.240	.000	.267	.259

Table 3.10 continued.

Table 3.11 Entire Sample

DESCRIPTOR	S_{hh}	S_{hs}	S_{hc}	R^2_h
LFLEN	.036	.022	.013	.374
PELEN	.048	.030	.018	.370
R3412	.041	.034	.007	.180
BANG	.040	.033	.007	.168
THLEN	.028	.024	.004	.141
PTLEN	.036	.032	.004	.120

DESCRIPTOR PAIR	|PMCC| (>.3)	VALUE IS
LFLEN & PELEN	.52	OK
R3412 & BANG	.31	OK

Tables 3.11 to 3.15

Analysis of Dispersion Structure in the reduced data set (6 ranged descriptors) for the entire sample, and for each of the four groups. S_{hh} = total variance, S_{hs} = descriptor-specific variance, S_{hc} = common variance, R^2_h = multiple correlation coefficient, PMCC = product moment correlation coefficient. See text for details.

Table 3.12 - GROUP 1

DESCRIPTOR	S_{hh}	S_{hs}	S_{hc}	R^2_h
LFLEN	.036	.013	.022	.625
PELEN	.047	.019	.029	.604
R3412	.050	.030	.020	.406
BANG	.028	.023	.005	.189
THLEN	.027	.022	.004	.168
PTLEN	.040	.031	.009	.229

DESCRIPTOR PAIR	|PMCCI (>.3)	VALUE IS
LFLEN & PELEN	.74	OK
LFLEN & R3412	.44	OK
LFLEN & PTLEN	.30	HIGH
PELEN & PTLEN	.36	OK
R3412 & BANG	.41	OK
R3412 & THLEN	.31	OK
R3412 & PTLEN	.32	OK

Table 3.13 - GROUP 2

DESCRIPTOR	S_{hh}	S_{hs}	S_{hc}	R^2_h
LFLEN	.014	.011	.003	.196
PELEN	.037	.015	.021	.580
R3412	.029	.012	.018	.606
BANG	.036	.014	.022	.605
THLEN	.015	.012	.003	.205
PTLEN	.060	.043	.018	.294

DESCRIPTOR PAIR	|PMCCI (>.3)	VALUE IS
LFLEN & BANG	.39	OK
PELEN & R3412	.57	OK
PELEN & BANG	.34	HIGH
PELEN & THLEN	.36	HIGH
PELEN & PTLEN	.32	HIGH
BANG & THLEN	.31	HIGH
BANG & PTLEN	.41	OK

Table 3.14 - GROUP 3

DESCRIPTOR	S_{hh}	S_{hs}	S_{hc}	R^2_h
LFLEN	.024	.000	.024	.999
PELEN	.118	.000	.118	.999
R3412	.058	.000	.058	.997
BANG	.017	.000	.017	.995
THLEN	.020	.001	.019	.933
PTLEN	.042	.010	.032	.756

DESCRIPTOR PAIR	|PMCC| (>.3)	VALUE IS
LFLEN & PELEN	.91	OK
LFLEN & THLEN	.44	HIGH
LFLEN & PTLEN	.47	HIGH
PELEN & BANG	.39	LOW
R3412 & THLEN	.83	OK
BANG & PTLEN	.51	HIGH

Table 3.15 - GROUP 4

DESCRIPTOR	S_{hh}	S_{hs}	S_{hc}	R^2_h
LFLEN	.012	.008	.004	.363
PELEN	.021	.014	.006	.302
R3412	.029	.027	.002	.071
BANG	.027	.026	.000	.016
THLEN	.018	.016	.002	.119
PTLEN	.013	.012	.001	.054

DESCRIPTOR PAIR	|PMCC| (>.3)	VALUE IS
LFLEN & PELEN	.51	OK

Table 3.16a COMPONENT WEIGHTS

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	.347	-.556	-.067	.066	-.072	.746
PELEN	.714	-.032	-.070	.083	.229	-.557
R3412	.340	.540	.108	.734	-.109	.176
BANG	-.403	-.352	-.550	.530	-.300	-.202
THLEN	-.296	-.194	.208	.358	.836	.059
PTLEN	-.067	-.382	.780	.200	-.376	-.240
EIGENVALUES	.069	.061	.036	.027	.021	.015
% VARIANCE	30.3%	26.7%	15.6%	11.7%	9.2%	6.6%
CUM. VAR.	30.3%	57.0%	72.6%	84.3%	93.4%	100.0%

Table 3.16b BARTLETT'S TEST

LAST *n* COMPONENTS	χ^2	D.F.	SIGNIFICANCE
5	63.8	14	***
4	21.5	9	***
3	9.0	5	**
2	3.1	2	NS

***	P < .01
**	P < .025
*	P < .05
NS	No Significant Difference

Table 3.16 Principal Components Analysis of the dispersion matrix of 6 ranged descriptors for the entire sample.

3.16a Component Weights (eigenvector elements), Component Roots (eigenvalues) and absolute and cumulative variance accounted for by each Principal Component.

3.16b Bartlett's test of the equality of the last *n* (= 5 to 2) Principal Components. χ^2 = chi-square test statistic, D.F. = degrees of freedom.

Table 3.16c COMPONENT-DESCRIPTOR CORRELATIONS

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	.482	-.724	-.067	.059	-.055	.482
PELEN	.859	-.340	-.146	.062	.151	-.312
R3412	.440	.655	.101	.591	-.077	.106
BANG	-.533	-.437	-.522	.436	-.218	-.124
THLEN	-.465	-.286	.235	.350	.723	.043
PTLEN	-.093	-.500	.778	.173	-.287	-.155

TABLE 3.16d PERCENTAGE VARIANCE ACCOUNTED FOR

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	23.21	52.44	0.44	0.35	0.30	23.26
PELEN	73.86	11.59	2.14	0.39	2.29	9.73
R3412	19.38	42.95	1.01	34.93	0.60	1.12
BANG	28.39	19.07	27.25	19.00	4.76	1.54
THLEN	21.60	8.18	5.52	12.25	52.26	0.19
PTLEN	0.86	25.05	60.45	2.98	8.25	2.40

Table 3.16c Correlation between each descriptor and each Principal Component.

3.16d Percentage descriptor variance accounted for by each Principal Component

Table 3.17a COMPONENT WEIGHTS

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	.501	.217	-.229	-.327	.023	.736
PELEN	.538	.533	-.214	.004	.169	-.594
R3412	-.511	.619	-.072	.277	.459	.251
BANG	.168	-.486	-.053	-.015	.856	-.022
THLEN	.196	-.192	-.480	.806	-.160	.137
PTLEN	.363	.111	.815	.407	.053	.153
EIGENVALUES	.100	.047	.034	.019	.019	.008
% VARIANCE	43.9%	20.7%	14.9%	8.5%	8.3%	3.7%
CUM. VAR.	43.9%	64.6%	79.5%	88.0%	96.3%	100.0%

Table 3.17b BARTLETT'S TEST

LAST *n* COMPONENTS	χ^2	D.F.	SIGNIFICANCE
5	26.2	14	**
4	15.0	9	NS
3	6.5	5	NS
2	3.1	2	*

***	P < .01
**	P < .025
*	P < .05
NS	No Significant Difference

Table 3.17 Principal Components Analysis of the dispersion matrix of 6 ranged descriptors for Group 1.

3.17a Component Weights (eigenvector elements), Component Roots (eigenvalues) and absolute and cumulative variance accounted for by each Principal Component.

3.17b Bartlett's test of the equality of the last *n* (*n* = 5 to 2) Principal Components. χ^2 = chi-square test statistic, D.F. = degrees of freedom.

Table 3.17c COMPONENT-DESCRIPTOR CORRELATIONS

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	.838	.249	-.223	-.241	.017	.358
PELEN	.781	.532	-.181	.002	.107	-.251
R3412	-.720	.600	-.060	.172	.281	.103
BANG	.318	-.632	-.058	-.013	.704	-.012
THLEN	.380	-.256	-.541	.687	-.134	.078
PTLEN	.575	.121	.753	.284	.037	.071

Table 3.17d PERCENTAGE VARIANCE ACCOUNTED FOR

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	70.15	6.20	4.98	5.80	0.03	12.84
PELEN	61.01	28.29	3.26	0.00	1.14	6.30
R3412	51.77	35.96	0.35	2.94	7.90	1.06
BANG	10.09	40.00	0.34	0.02	49.54	0.01
THLEN	33.10	1.47	56.73	8.06	0.13	0.50

Table 3.17c Correlation between each descriptor and each Principal Component.

3.17d Percentage descriptor variance accounted for by each Principal Component.

Table 3.18a COMPONENT WEIGHTS

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	-.104	-.110	.246	.681	-.571	-.356
PELEN	-.453	.446	.467	-.419	-.021	-.450
R3412	-.235	.691	-.072	.268	-.178	.598
BANG	-.400	-.514	.529	-.095	-.047	.534
THLEN	-.200	.046	.169	.529	.798	-.108
PTLEN	.727	.215	.638	.016	.042	-.126
EIGENVALUES	.086	.044	.031	.013	.011	.005
% VARIANCE	45.1%	23.3%	16.2%	6.8%	5.8%	2.8%
CUM. VAR.	45.1%	68.4%	84.6%	91.4%	97.2%	100.0%

Table 3.18b BARTLETT'S TEST

LAST *n* COMPONENTS	χ^2	D.F.	SIGNIFICANCE
5	7.0	14	NS
4	3.7	9	NS
3	0.8	5	NS
2	0.8	2	NS

***	P < .01
**	P < .025
*	P < .05
NS	No Significant Difference

Table 3.18 Principal Components Analysis of the dispersion matrix of 6 ranged descriptors for Group 2.

3.18a Component Weights (eigenvector elements), Component Roots (eigenvalues) and absolute and cumulative variance accounted for by each Principal Component.

3.18b Bartlett's test of the equality of the last *n* (= 5 to 2) Principal Components. χ^2 = chi-square test statistic, D.F. = degrees of freedom.

Table 3.18c COMPONENT-DESCRIPTOR CORRELATIONS

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	-.260	-.199	.371	.665	-.513	-.224
PELEN	-.694	.491	.430	-.249	-.012	-.172
R3412	-.402	.850	-.081	.178	-.109	.256
BANG	-.620	-.572	.492	-.057	-.026	.207
THLEN	-.477	.079	.242	.490	.681	-.064
PTLEN	.869	.185	.457	.008	.018	.038

Table 3.18d PERCENTAGE VARIANCE ACCOUNTED FOR

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	6.77	3.95	13.75	44.16	26.36	5.00
PELEN	48.22	24.11	18.47	6.21	0.01	2.97
R3412	16.18	72.22	0.66	3.17	1.20	6.57
BANG	38.41	32.70	24.20	0.33	0.07	4.29
THLEN	22.77	0.62	5.83	24.00	46.40	0.41
PTLEN	75.54	3.41	20.87	0.00	0.03	0.14

Table 3.18c Correlation between each descriptor and each Principal Component.

3.18d Percentage descriptor variance accounted for by each Principal Component.

Table 3.19a COMPONENT WEIGHTS

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	.397	-.085	-.104	-.156	-.361	.818
PELEN	.880	.246	.070	-.007	.014	-.400
R3412	-.094	.810	-.301	-.199	-.392	.227
BANG	.118	.029	.500	.628	-.475	.339
THLEN	-.131	.396	-.171	.643	.616	-.065
PTLEN	-.166	.346	.784	-.357	.334	.002
EIGENVALUES	.146	.080	.045	.008	.002	.000
% VARIANCE	52.3%	28.5%	16.0%	2.7%	0.6%	0.0%
CUM. VAR.	52.3%	80.7%	96.7%	99.4%	100.0%	100.0%

Table 3.19b BARTLETT'S TEST

LAST *n* COMPONENTS	χ^2	D.F.	SIGNIFICANCE
5	32.7	14	***
4	21.6	9	**
3	9.5	5	*
2	3.9	2	NS

***	P < .01
**	P < .025
*	P < .05
NS	No Significant Difference

Table 3.19 Principal Components Analysis of the dispersion matrix of 6 ranged descriptors for Group 3.

3.19a Component Weights (eigenvector elements), Component Roots (eigenvalues) and absolute and cumulative variance accounted for by each Principal Component.

3.19b Bartlett's test of the equality of the last *n* (= 5 to 2) Principal Components. χ^2 = chi-square test statistic, D.F. = degrees of freedom.

Table 3.19c COMPONENT-DESCRIPTOR CORRELATIONS

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	.969	-.154	-.141	-.086	.096	.016
PELEN	.978	.202	.043	-.002	.002	-.004
R3412	-.150	.948	-.264	-.071	-.068	.003
BANG	.351	.064	.819	.421	-.153	.008
THLEN	-.354	.789	-.256	.393	-.181	-.001
PTLEN	-.389	.475	.807	-.150	.068	.000

Table 3.19d PERCENTAGE VARIANCE ACCOUNTED FOR

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	93.97	2.37	1.98	0.74	0.91	0.02
PELEN	95.75	4.06	0.18	0.00	0.00	0.00
R3412	2.24	89.81	7.00	0.51	0.46	0.00
BANG	12.34	0.41	67.13	17.77	2.34	0.01
THLEN	12.54	62.20	6.55	15.44	3.26	0.00
PTLEN	9.58	22.58	65.12	2.26	0.46	0.00

Table 3.19c Correlation between each descriptor and each Principal Component.

3.19d Percentage descriptor variance accounted for by each Principal Component.

Table 3.20a COMPONENT WEIGHTS

DESCRIPTORS	COMPONENTS 1	2	3	4	5	6
LFLEN	.015	.490	-.011	.197	.088	.844
PELEN	.285	.679	-.049	.277	-.454	-.416
R3412	.924	-.330	.039	.119	.022	.145
BANG	.027	.135	.979	-.110	.091	-.050
THLEN	-.252	-.407	.193	.695	-.484	.132
PTLEN	-.023	.085	-.024	.613	.737	-.269
EIGENVALUES	.030	.029	.027	.014	.013	.006
% VARIANCE	25.4%	24.2%	22.7%	11.9%	10.9%	4.9%
CUM. VAR.	25.4%	49.6%	72.3%	84.2%	95.1%	100.0%

Table 3.20b BARTLETT'S TEST

LAST *n* COMPONENTS	χ^2	D.F.	SIGNIFICANCE
5	42.8	14	***
4	31.5	9	***
3	11.7	5	**
2	9.0	2	**

***	P < .01
**	P < .025
*	P < .05
NS	No Significant Difference

Table 3.20 Principal Components Analysis of the dispersion matrix of 6 ranged descriptors for Group 4.

3.20a Component Weights (eigenvector elements), Component Roots (eigenvalues) and absolute and cumulative variance accounted for by each Principal Component.

3.20b Bartlett's test of the equality of the last *n* (= 5 to 2) Principal Components. χ^2 = chi-square test statistic, D.F. = degrees of freedom.

Table 3.20c COMPONENT-DESCRIPTOR CORRELATIONS

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	.024	.767	-.017	.216	.093	.596
PELEN	.345	.803	-.056	.230	-.361	-.222
R3412	.938	-.327	.037	.083	.015	.065
BANG	.029	.140	.984	-.080	.064	-.231
THLEN	-.329	-.520	.238	.622	-.415	.076
PTLEN	-.035	.127	-.035	.638	.736	-.180

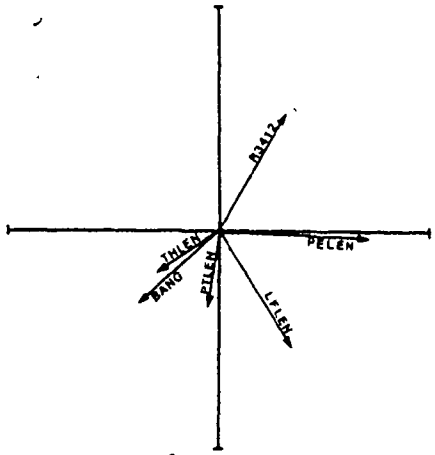
Table 3.20d PERCENTAGE VARIANCE ACCOUNTED FOR

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	0.06	58.89	0.03	4.67	0.86	35.48
PELEN	11.90	64.54	0.31	5.28	13.04	4.92
R3412	88.03	10.70	0.14	0.68	0.02	0.42
BANG	0.08	1.95	96.86	0.64	0.40	0.05
THLEN	10.84	27.03	5.68	38.63	17.25	0.57

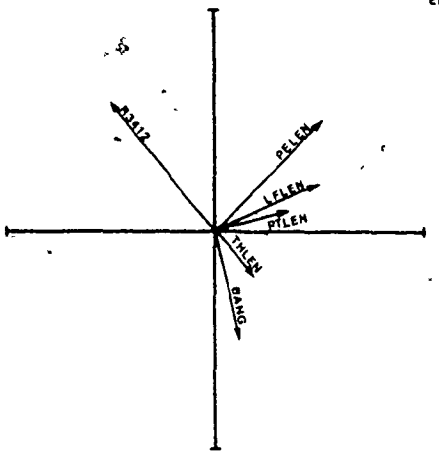
Table 3.20c Correlation between each descriptor and each Principal Component.

3.20d Percentage descriptor variance accounted for by each Principal Component.

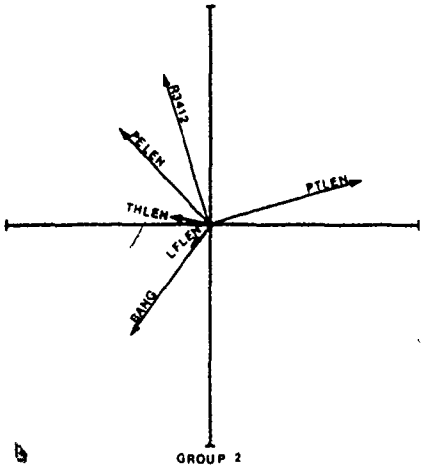
Figure 3.8 Descriptor vector plots, showing the loading (Component Weights) of each descriptor on the first two components of a Principal Components Analysis based on the covariance matrix of six ranged descriptors. Plots are presented for the entire sample (112 OTUs) and for each of the four Groups discovered via Cluster Analysis. Horizontal axes represent the first Principal Component, vertical axes the second. Descriptor abbreviations are explained in Table 3.1.



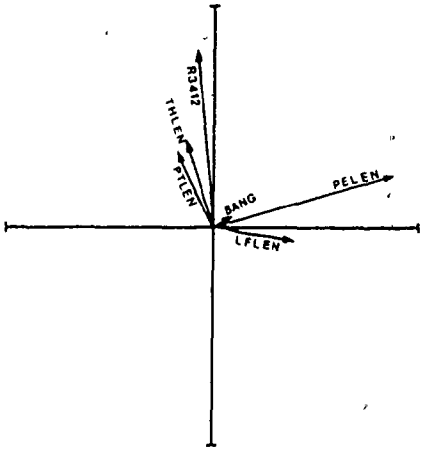
ENTIRE SAMPLE



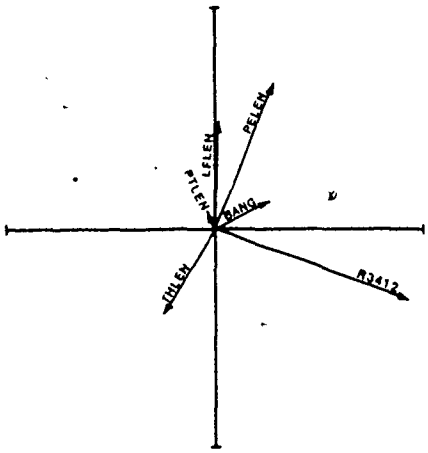
GROUP 1



GROUP 2



GROUP 3



GROUP 4

DISPERSION DETERMINANTS (X 10^{-10})

Group 1	Group 2	Group 3	Group 4
4.947	.9179	.0006	.2569

TEST OF H_1 – (equality of dispersion determinants) – $\Delta_k = \Delta$ (k = 1-4)

Box's M = 141.37

$F_{app.}$ = 1.63 (63 & 1439 D.F.)

H_1 rejected, P < .01

TEST OF H_2 – (equality of centriods) – $\mu_k = \mu$ (k = 1-4)

Wilk's Λ = .1359

Rao's F = 16.62 (18 & 291 D.F.)

H_2 rejected, P < .01

Table 3.21 Multivariate Analysis of Variance of four *C.* series *Rotundifoliae* groups, testing the equality of group dispersion determinants and the equality of group centriods.

Table 3.22a DISCRIMINANT WEIGHTS

DESCRIPTORS	DISCRIMINANT AXES 1	2	3
LFLEN	4.97	3.84	-1.11
PELEN	-3.25	0.46	0.98
R3412	0.87	1.37	0.41
BANG	1.02	-1.75	-3.15
THLEN	1.34	-1.76	2.82
PTLEN	1.13	-0.89	3.27
ROOTS	1.513	.962	.492
% VARIANCE	51.0%	32.4%	16.6%
CUM. VAR.	51.0%	83.4%	100.0%

Table 3.22b BARTLETT'S TEST

ROOTS REMOVED	χ^2	D.F.	SIGNIFICANCE
0	211.6	18	***
1	113.9	10	***
2	42.3	4	***

***	P < .01
**	P < .025
*	P < .05
NS	No Significant Difference

Table 3.22 Multi-Group Discriminant Analysis of four *C.* series *Rotundifoliae* groups dscovered via Cluster Analysis.

3.22a Discriminant Weights (eigenvector elements), Discriminant Roots (eigenvalues) and absolute and cumulative variance accounted for by each Discriminant Axis.

3.22b Bartlett's test of the equality of the Discriminant Axes, with successive roots removed. χ^2 = chi-square test statistic, D.F. = degrees of freedom.

Table 3.22c FACTOR STRUCTURE

DESCRIPTORS	DISCRIMINANT AXES 1	2	3
LFLEN	.597	.677	-.039
PELEN	-.294	.648	.134
R3412	-.236	.364	.131
BANG	.416	-.477	-.585
THLEN	.409	-.517	.397
PTLEN	.444	-.131	.646

Table 3.22d TOTAL COMMUNALITIES

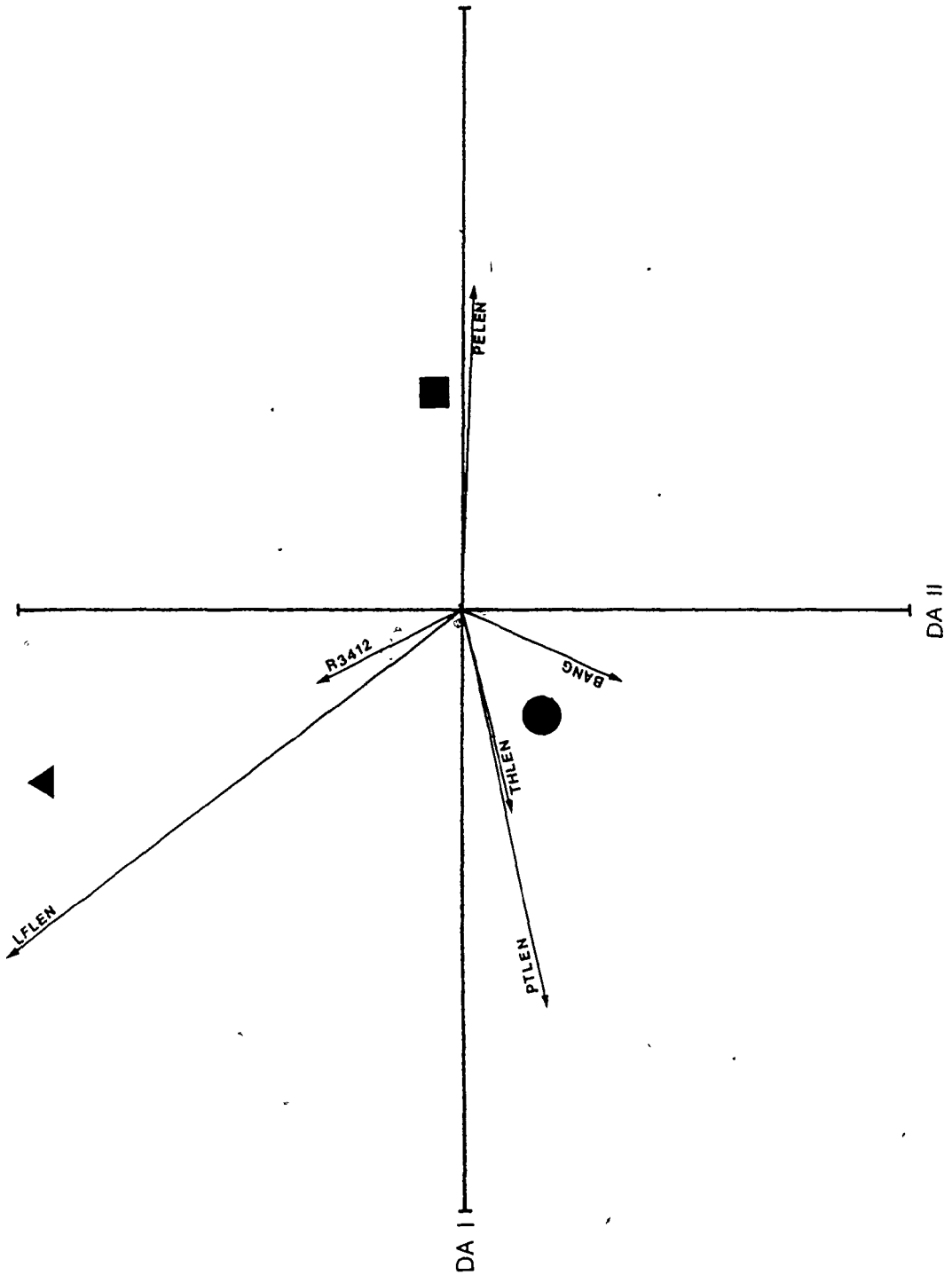
LFLEN	.817
PELEN	.524
R3412	.205
BANG	.742
THLEN	.592
PTLEN	.632

Table 3.22c Factor Structure coefficients (correlation between each descriptor and each Discriminant Axis).

3.22d Total Communalities (contribution of each descriptor to discriminance amongst the groups).

Figure 3.9 Scatter diagram of the position of the centroids of each of four *Crataegus* series *Rotundifoliae* groups in the space of the first two Discriminant Axes (DA I and II) of a Multi-group Discriminant Analysis based on six ranged descriptors. Superimposed on the diagram are vectors illustrating the loading (Discriminant Weights) of each descriptor on the Discriminant Axes. Descriptor abbreviations are explained in Table 3.1.

Symbols:
Group 1 centriod - closed square
Group 2 centroid - closed triangle
Group 3 centroid - open square
Group 4 centroid - closed circle


PELEN
DA II
R3412
BANG
THLEN
LFLEN
PTLEN
DA I

DISPERSION DETERMINANTS (X 10^{-10})

Group 1	Group 2	Group 4
4.947	.9179	.2569

TEST OF H_1 - (equality of dispersion determinants) - $\Delta_k = \Delta$ (k = 1-3)

Box's M = 72.7

$F_{app.}$ = 1.39 (42 & 1659 D.F.)

H_1 rejected, P < .01

TEST OF H_2 - (equality of centriods) - $\mu_k = \mu$ (k = 1-3)

Wilk's Λ = .2006

Rao's F = 19.93 (12 & 194 D.F.)

H_2 rejected, P < .01

Table 3.23 Multivariate Analysis of Variance of three *C.* series *Rotundifoliae* groups (Group 3 excluded), testing the equality of group dispersion determinants and the equality of group centriods.

Table 3.24a DISCRIMINANT WEIGHTS

DESCRIPTORS	DISCRIMINANT AXES 1	2
LFLEN	-3.52	4.69
PELEN	3.19	-0.04
R3412	-0.76	1.50
BANG	-0.72	-1.68
THLEN	-2.21	-0.95
PTLEN	-2.02	-0.39
ROOTS	1.451	1.034
% VARIANCE	58.4%	41.6%
CUM. VAR.	58.4%	100.0%

Table 3.24b BARTLETT'S TEST

ROOTS REMOVED	χ^2	D.F.	SIGNIFICANCE
0	159.8	12	***
1	70.6	5	***

***	P < .01
**	P < .025
*	P < .05
NS	No Significant Difference

Table 3.24 Multi-Group Discriminant Analysis of three *C.* series *Rotundifoliae* groups (Group 3 excluded).

3.24a Discriminant Weights (eigenvector elements), Discriminant Roots (eigenvalues) and absolute and cumulative variance accounted for by each Discriminant Axis.

3.24b Bartlett's test of the equality of the Discriminant Axes, with successive roots removed. χ^2 = chi-square test statistic, D.F. = degrees of freedom.

Table 3.24c FACTOR STRUCTURE

DESCRIPTORS	DISCRIMINANT AXES 1	2
LFLEN	-.384	.808
PELEN	.410	.614
R3412	.273	.326
BANG	-.356	-.439
THLEN	-.616	-.349
PTLEN	-.646	.072

Table 3.24d TOTAL COMMUNALITIES

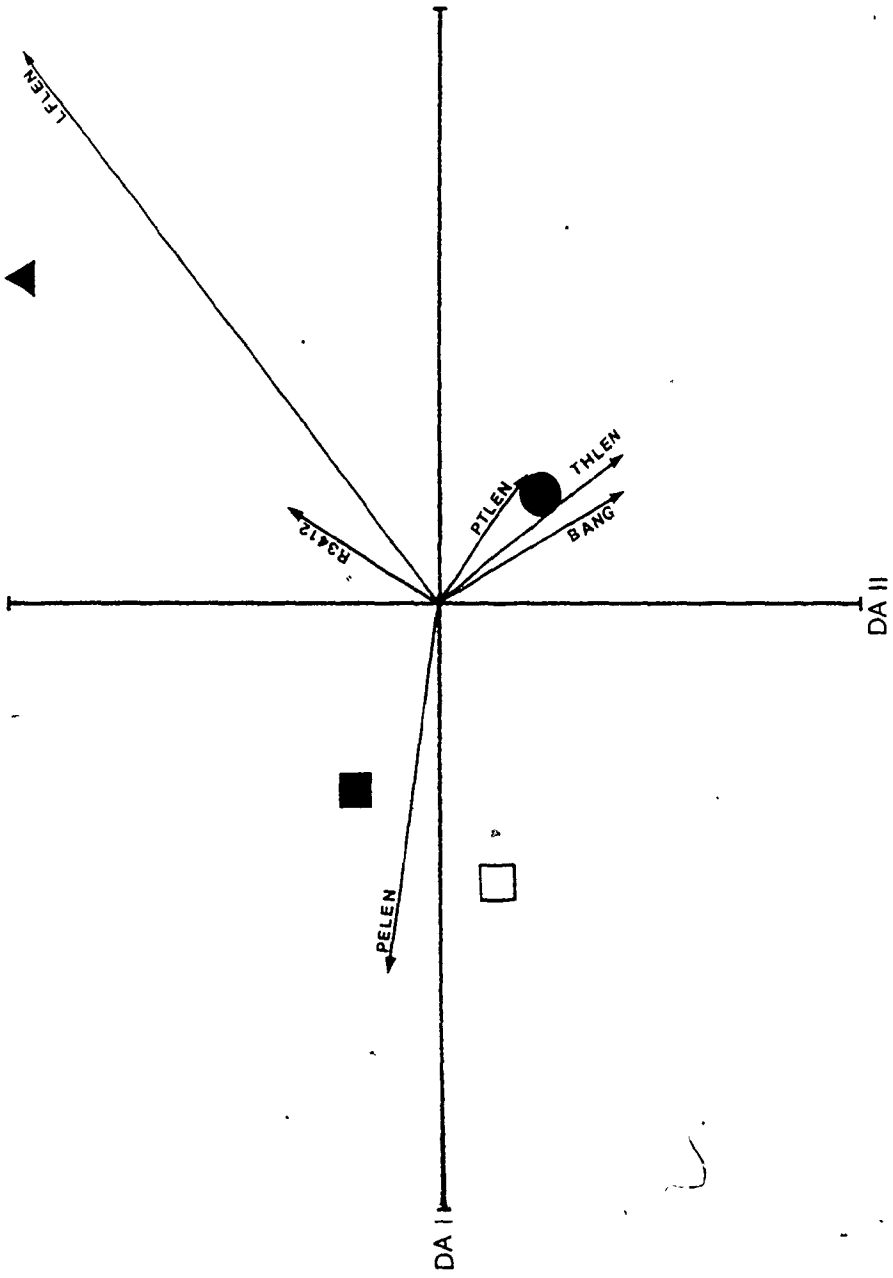
LFLEN	.800
PELEN	.546
R3412	.181
BANG	.370
THLEN	.501
PTLEN	.432

Table 3.24c Factor Structure coefficients (correlation between each descriptor and each Discriminant Axis).

3.24d Total Communalities (contribution of each descriptor to discriminance amongst the groups).

Figure 3.10 Scatter diagram of the position of the centroids of each of three *Crataegus* series *Rotundifoliae* groups (Group 3 removed) in the space of the first two Discriminant Axes (DA I and II) of a Multi-group Discriminant Analysis based on six ranged descriptors. Superimposed on the diagram are vectors illustrating the loading (Discriminant Weights) of each descriptor on the Discriminant Axes. Descriptor abbreviations are explained in Table 3.1.

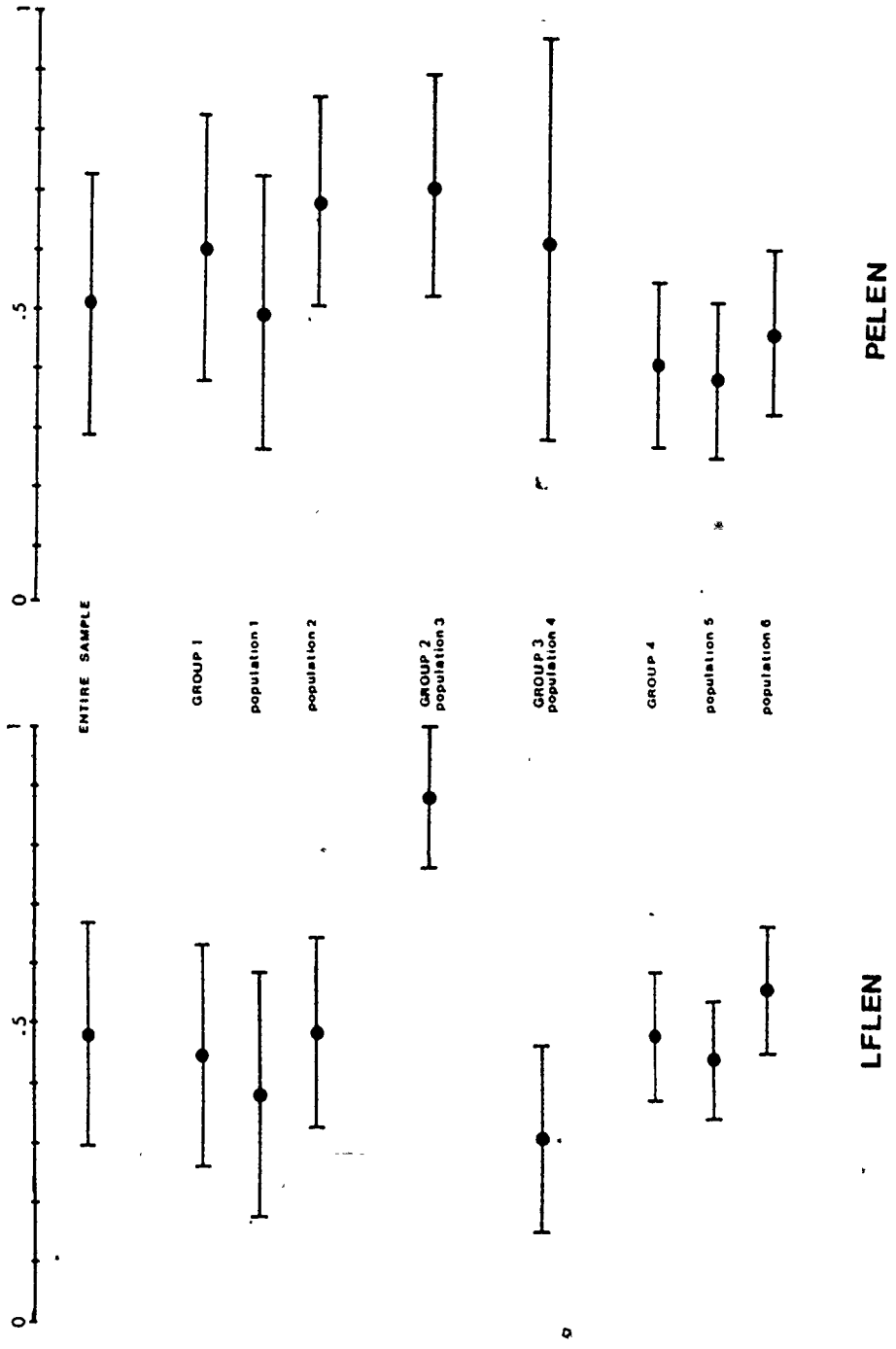
Symbols:
Group 1 centriod - closed square
Group 2 centroid - closed triangle
Group 4 centroid - closed circle

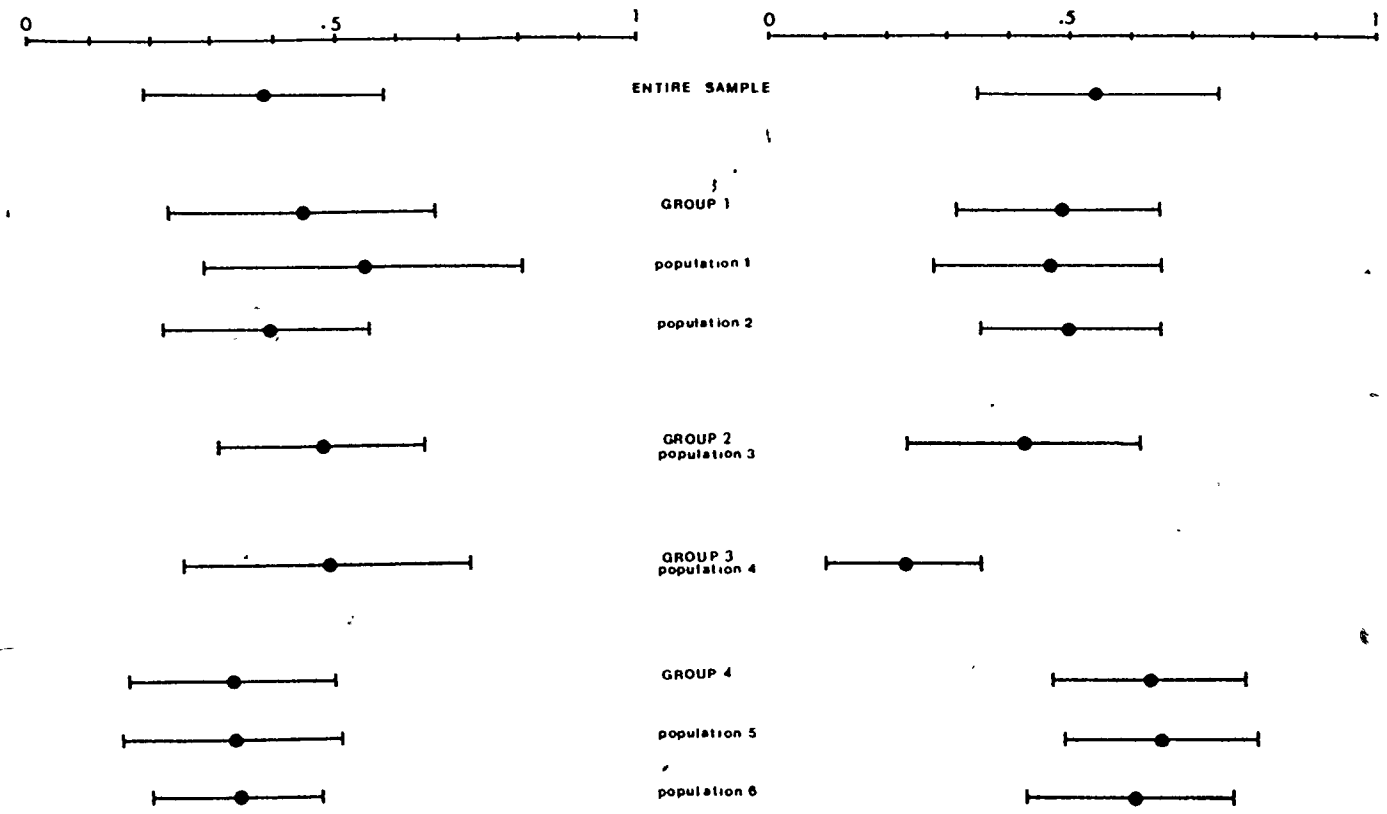

LFLEN
R3412
PTLEN
THLEN
BANG
DA II
PELEN
DA I

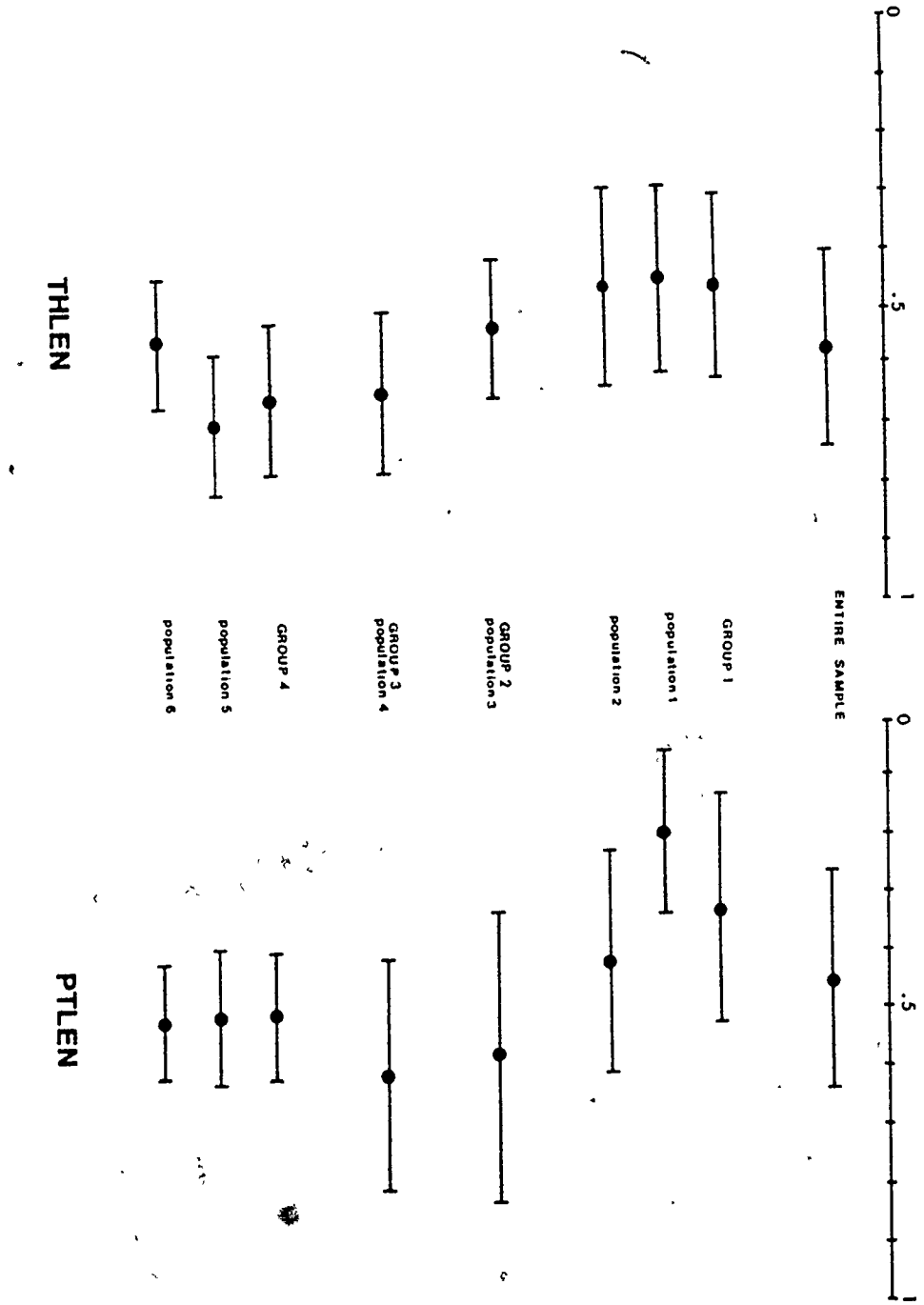
DESCRIPTORS	POPULATIONS					
	1	2	3	4	5	6
LFLEN						
Mean	.361	.481	.880	.298	.432	.546
S.D.	.208	.165	.117	.157	.090	.106
PELEN						
Mean	.492	.682	.709	.615	.379	.466
S.D.	.229	.177	.191	.344	.136	.144
R3412						
Mean	.548	.386	.484	.490	.331	.338
S.D.	.260	.173	.171	.241	.185	.144
BANG						
Mean	.459	.500	.418	.227	.647	.601
S.D.	.195	.147	.189	.129	.158	.175
THLEN						
Mean	.447	.461	.528	.643	.693	.549
S.D.	.157	.170	.123	.116	.116	.114
PTLEN						
Mean	.200	.423	.585	.508	.508	.518
S.D.	.137	.186	.246	.121	.121	.103

Table 3.25 Summary of univariate means and standard deviations of 6 ranged descriptors for each of six *C.* series *Rotundifoliae* populations.

Figure 3.11 - Means (dots) and Standard Deviations (horizontal bar = 2 S.D.) for each of six ranged descriptors for the entire sample, the four groups discovered via Cluster Analysis and the six populations into which the Groups were divided. Descriptor abbreviations are explained in Table 3.1.


PELEN
LFLEN
0
.5
1
ENTIRE SAMPLE
GROUP 1
population 1
population 2
GROUP 2
population 3
GROUP 3
population 4
GROUP 4
population 5
population 6


0
.5
1
0
.5
1
ENTIRE SAMPLE
GROUP 1
population 1
population 2
GROUP 2
population 3
GROUP 3
population 4
GROUP 4
population 5
population 6
R3412
BANG


THLEN
PTLEN
0
.5
1
ENTIRE SAMPLE
GROUP 1
population 1
population 2
GROUP 2
population 3
GROUP 3
population 4
GROUP 4
population 5
population 6

Tables 3.26 to 3.29

Analysis of Dispersion Structure in the reduced data set (6 ranged descriptors) for each of six *C.* series *Rotundifoliae* populations. S_{hh} = total variance, S_{hs} = descriptor–specific variance, S_{hc} = common variance, R^2_h = multiple correlation coefficient, PMCC = product moment correlation coefficient (see text for details).

Table 3.26 POPULATION 1

DESCRIPTOR	S_{hh}	S_{hs}	S_{hc}	R^2_h
LFLEN	.043	.015	.029	.662
PELEN	.052	.016	.036	.687
R3412	.068	.016	.052	.760
BANG	.038	.021	.017	.453
THLEN	.025	.009	.017	.649
PTLEN	.019	.014	.005	.250

DESCRIPTOR PAIR	|PMCC| (>.3)	VALUE IS
LFLEN & PELEN	.66	OK
LFLEN & R3412	.33	OK
R3412 & BANG	.53	OK
R3412 & THLEN	.61	OK

Table 3.27 POPULATION 2

DESCRIPTOR	S_{hh}	S_{hs}	S_{hc}	R^2_h
LFLEN	.027	.009	.018	.654
PELEN	.031	.012	.019	.604
R3412	.030	.022	.009	.286
BANG	.022	.017	.005	.221
THLEN	.029	.022	.007	.229
PTLEN	.034	.030	.004	.123

DESCRIPTOR PAIR	|PMCC| (>.3)	VALUE IS
LFLEN & PELEN	.76	OK
LFLEN & R3412	.44	OK
PELEN & R3412	.33	HIGH
R3412 & PTLEN	.30	OK
BANG & THLEN	.37	OK

Table 3.28 POPULATION 5

DESCRIPTOR	S_{hh}	S_{hs}	S_{hc}	R^2_h
LFLEN	.008	.005	.003	.347
PELEN	.019	.013	.005	.293
R3412	.034	.029	.005	.157
BANG	.025	.024	.001	.034
THLEN	.014	.013	.001	.046
PTLEN	.015	.014	.001	.054

DESCRIPTOR PAIR	|PMCCI (>.3)	VALUE IS
LFLEN & PELEN	.46	OK

Table 3.29 POPULATION 6

DESCRIPTOR	S_{hh}	S_{hs}	S_{hc}	R^2_h
LFLEN	.011	.008	.003	.302
PELEN	.021	.015	.005	.264
R3412	.021	.020	.001	.055
BANG	.031	.025	.006	.195
THLEN	.013	.012	.001	.054
PTLEN	.010	.008	.002	.225

DESCRIPTOR PAIR	|PMCCI (>.3)	VALUE IS
LFLEN & PELEN	.42	OK

Table 3.30a COMPONENT WEIGHTS

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	-.402	.499	.239	-.285	-.570	.354
PELEN	-.129	.772	.009	.100	.490	-.372
R3412	.769	.263	.203	-.250	.178	.452
BANG	-.365	-.279	.688	-.182	.502	.176
THLEN	-.311	.016	-.607	-.034	.383	.622
PTLEN	.015	.089	.247	.901	-.070	.338
EIGENVALUES	.099	.079	.032	.019	.012	.004
% VARIANCE	40.4%	32.3%	13.1%	7.8%	4.7%	1.8%
CUM. VAR.	40.4%	72.6%	85.7%	93.5%	98.2%	100.0%

Table 3.30b BARTLETT'S TEST

LAST *n* COMPONENTS	χ^2	D.F.	SIGNIFICANCE
5	24.6	14	**
4	10.0	9	NS
3	5.2	5	NS
2	2.2	2	NS

***	P < .01
**	P < .025
*	P < .05
NS	No Significant Difference

Table 3.30 Principal Components Analysis of the dispersion matrix of 6 ranged descriptors for Population 1.

3.30a Component Weights (eigenvector elements), Component Roots (eigenvalues) and absolute and cumulative variance accounted for by each Principal Component.

3.30b Bartlett's test of the equality of the last *n* (= 5 to 2) Principal Components. χ^2 = chi-square test statistic, D.F. = degrees of freedom.

Table 3.30c COMPONENT-DESCRIPTOR CORRELATIONS

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	-.607	.674	.206	-.189	-.294	.112
PELEN	-.178	.949	.007	.061	.230	-.107
R3412	.929	.284	.140	-.133	.074	.114
BANG	-.088	-.402	.630	-.129	.276	.059
THLEN	-.622	.030	-.690	-.030	.261	.260
PTLEN	.034	.184	.323	.912	-.055	.162

Table 3.30d PERCENTAGE VARIANCE ACCOUNTED FOR

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	36.93	45.39	4.22	3.59	8.63	1.24
PELEN	3.16	90.05	0.00	0.37	5.28	1.14
R3412	86.36	8.09	1.94	1.77	0.54	1.30
BANG	34.54	16.14	39.71	1.66	7.59	0.35
THLEN	38.68	0.09	47.61	0.09	6.80	6.73
PTLEN	0.12	3.38	10.44	83.11	0.30	2.64

Table 3.30c Correlation between each descriptor and each Principal Component.

3.30d Percentage descriptor variance accounted for by each Principal Component

Table 3.31a COMPONENT WEIGHTS

DESCRIPTORS	COMPONENTS 1	2	3	4	5	6
LFLEN	.537	-.025	-.350	.106	-.024	.759
PELEN	.553	-.200	-.406	.121	.355	-.591
R3412	-.434	-.268	-.112	.733	.406	.167
BANG	.160	-.302	.549	-.337	.655	.198
THLEN	.276	-.606	.426	.312	-.521	-.078
PTLEN	.341	.655	.465	.474	.094	-.069
EIGENVALUES	.069	.035	.028	.022	.013	.006
% VARIANCE	39.9%	20.4%	16.2%	12.8%	7.3%	3.4%
CUM. VAR.	39.9%	60.2%	76.5%	89.3%	96.6%	100.0%

Table 3.31b BARTLETT'S TEST

LAST *n* COMPONENTS	χ^2	D.F.	SIGNIFICANCE
5	15.9	14	NS
4	11.7	9	NS
3	7.5	5	NS
2	2.4	2	NS

***	P < .01
**	P < .025
*	P < .05
NS	No Significant Difference

Table 3.31 Principal Components Analysis of the dispersion matrix of 6 ranged descriptors for Population 2.

3.31a Component Weights (eigenvector elements), Component Roots (eigenvalues) and absolute and cumulative variance accounted for by each Principal Component.

3.31b Bartlett's test of the equality of the last *n* (= 5 to 2) Principal Components. χ^2 = chi-square test statistic, D.F. = degrees of freedom.

Table 3.31c COMPONENT-DESCRIPTOR CORRELATIONS

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	.858	-.028	-.357	.096	-.017	.354
PELEN	.823	-.213	-.386	.102	.226	-.257
R3412	-.658	-.290	-.108	.630	.264	.069
BANG	.286	-.386	.625	-.341	.501	.104
THLEN	.426	-.670	.419	.273	-.344	-.035
PTLEN	.483	.664	.420	.381	.057	-.028

Table 3.31d PERCENTAGE VARIANCE ACCOUNTED FOR

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	73.65	0.08	12.74	0.93	0.03	12.57
PELEN	67.79	4.54	14.89	1.05	5.12	6.61
R3412	43.24	8.42	1.17	39.71	6.97	0.48
BANG	8.20	14.91	39.08	11.63	25.11	1.07
THLEN	18.13	44.82	17.60	7.46	11.86	0.12
PTLEN	23.37	44.03	17.65	14.53	0.32	0.08

Table 3.31c Correlation between each descriptor and each Principal Component.

3.31d Percentage descriptor variance accounted for by each Principal Component

Table 3.32a COMPONENT WEIGHTS

DESCRIPTORS	COMPONENTS 1	2	3	4	5	6
LFLEN	-.116	.093	.417	.097	.037	.890
PELEN	.153	-.039	.876	-.043	.231	-.391
R3412	.969	.006	-.117	.093	.101	.166
BANG	.007	.979	-.027	-.154	.106	-.076
THLEN	-.152	-.012	-.176	.526	.818	-.027
PTLEN	.011	.176	.118	.825	-.505	-.140
EIGENVALUES	.036	.026	.021	.015	.012	.004
% VARIANCE	31.2%	22.4%	18.8%	13.4%	10.3%	3.9%
CUM. VAR.	31.2%	53.6%	72.4%	85.8%	96.1%	100.0%

Table 3.32b BARTLETT'S TEST

LAST *n* COMPONENTS	χ^2	D.F.	SIGNIFICANCE
5	28.8	14	**
4	21.7	9	***
3	13.6	5	**
2	8.3	2	***

***	P < .01
**	P < .025
*	P < .05
NS	No Significant Difference

Table 3.32 Principal Components Analysis of the dispersion matrix of 6 ranged descriptors for Population 5.

3.32a Component Weights (eigenvector elements), Component Roots (eigenvalues) and absolute and cumulative variance accounted for by each Principal Component.

3.32b Bartlett's test of the equality of the last *n* (= 5 to 2) Principal Components. χ^2 = chi-square test statistic, D.F. = degrees of freedom.

Table 3.32c COMPONENT-DESCRIPTOR CORRELATIONS

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	-.244	.165	.679	.133	.045	.658
PELEN	.212	-.046	.939	-.039	.184	-.190
R3412	.990	.005	-.093	.062	.059	.060
BANG	.008	.989	-.025	-.120	.073	-.032
THLEN	-.246	-.016	-.222	.558	.760	-.016
PTLEN	.017	.234	.142	.844	-.454	-.077

Table 3.32d PERCENTAGE VARIANCE ACCOUNTED FOR

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	5.93	2.72	46.13	1.77	0.20	43.25
PELEN	4.50	0.21	88.15	0.15	3.38	3.61
R3412	98.04	0.00	0.86	0.38	0.35	0.36
BANG	0.01	97.86	0.06	1.44	0.53	0.10
THLEN	6.07	0.03	4.91	31.15	57.83	0.02
PTLEN	0.03	5.45	2.03	71.31	20.58	0.60

Table 3.32c Correlation between each descriptor and each Principal Component.

3.32d Percentage descriptor variance accounted for by each Principal Component

Table 3.33a COMPONENT WEIGHTS

DESCRIPTORS	COMPONENTS 1	2	3	4	5	6
LFLEN	.245	.269	.209	.479	.312	.705
PELEN	.487	.404	.545	.004	-.475	-.277
R3412	-.052	-.702	.677	.117	.166	-.068
BANG	.826	-.322	-.343	-.046	.267	-.150
THLEN	-.018	-.283	-.288	.725	-.556	-.047
PTLEN	-.130	.296	.024	.479	.518	-.630
EIGENVALUES	.037	.025	.017	.012	.011	.005
% VARIANCE	34.5%	23.1%	16.1%	11.1%	10.2%	5.0%
CUM. VAR.	34.5%	57.6%	73.7%	84.8%	95.0%	100.0%

Table 3.33b BARTLETT'S TEST

LAST *n* COMPONENTS	χ^2	D.F.	SIGNIFICANCE
5	9.4	14	NS
4	5.0	9	NS
3	2.7	5	NS
2	2.3	2	NS

***	P < .01
**	P < .025
*	P < .05
NS	No Significant Difference

Table 3.33 Principal Components Analysis of the dispersion matrix of 6 ranged descriptors for Population 6.

3.33a Component Weights (eigenvector elements), Component Roots (eigenvalues) and absolute and cumulative variance accounted for by each Principal Component.

3.33b Bartlett's test of the equality of the last *n* (= 5 to 2) Principal Components. χ^2 = chi-square test statistic, D.F. = degrees of freedom.

Table 3.33c COMPONENT-DESCRIPTOR CORRELATIONS

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	.445	.399	.259	.493	.308	.488
PELEN	.649	.441	.496	.003	-.343	-.141
R3412	-0.07	-.768	.617	.089	.120	-.035
BANG	.906	-.289	-.257	-.029	.159	-.063
THLEN	-.030	-.390	-.331	.692	-.508	-.030
PTLEN	-.244	.453	.030	.507	.525	-.449

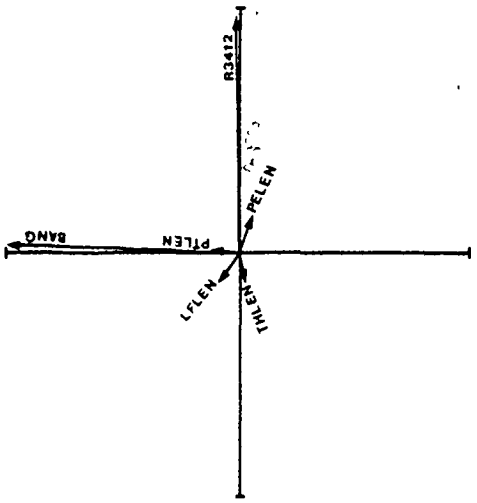
Table 3.33d PERCENTAGE VARIANCE ACCOUNTED FOR

	COMPONENTS					
DESCRIPTORS	1	2	3	4	5	6
LFLEN	19.77	15.94	6.69	24.33	9.46	23.81
PELEN	42.16	19.47	24.59	0.00	11.80	1.99
R3412	0.49	59.03	38.12	0.79	1.45	0.12
BANG	82.06	8.35	6.58	0.08	2.53	0.39
THLEN	0.93	15.21	10.98	47.85	25.78	0.09
PTLEN	5.94	20.52	0.09	25.72	27.59	20.14

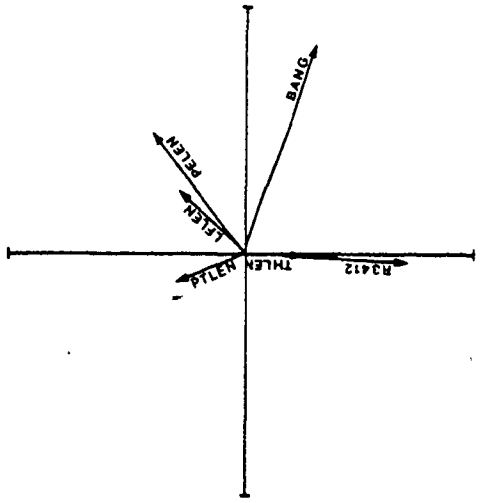
Table 3.33c Correlation between each descriptor and each Principal Component.

3.33d Percentage descriptor variance accounted for by each Principal Component

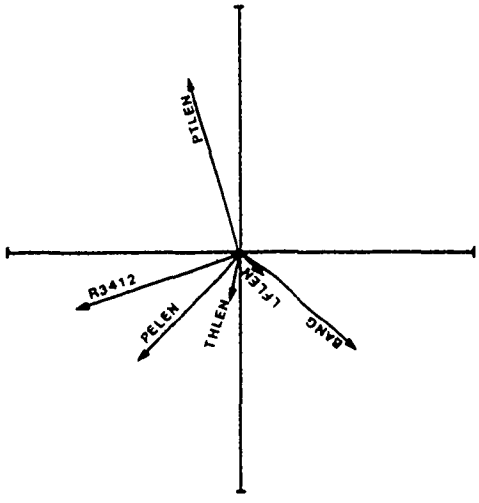
Figure 3.12 Descriptor vector plots, showing the loading (Component Weights) of each descriptor on the first two components of a Principal Components Analysis based on the covariance matrix of six ranged descriptors. Plots are presented for each of six *Crataegus* series *Rotundifoliae* populations. Horizontal axes represent the first Principal Component, vertical axes the second. Descriptor abbreviations are explained in Table 3.1.



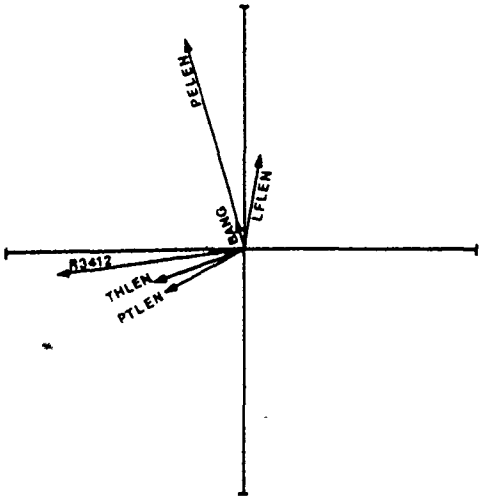
POPULATION 5



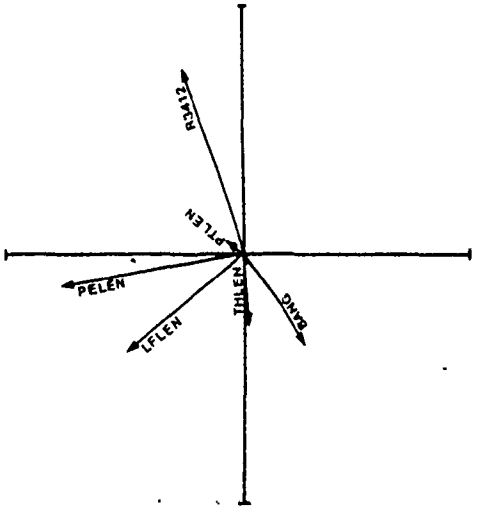
POPULATION 6



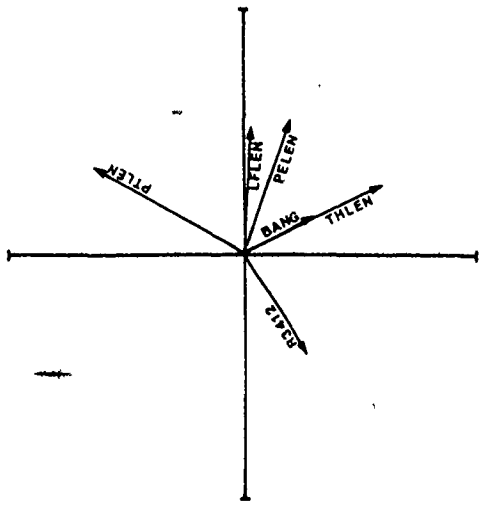
POPULATION 3



POPULATION 4



POPULATION 1



POPULATION 2

DISPERSION DETERMINANTS (X 10^{-10})

POPULATIONS					
1	2	3	4	5	6
2.38	1.14	.918	.0006	.154	.106

TEST OF H_1 -(equality of dispersion determinants)- $\Delta_k = \Delta$ (k = 1-6)

Box's M = 174.6

$F_{app.}$ = 1.27 (105 & 3478 D.F.)

H_1 rejected, P < .05

TEST OF H_2 - (equality of centriods) - $\mu_k = \mu$ (k = 1-6)

Wilk's Λ = .0813

Rao's F = 11.81 (30 &.406 D.F.)

H_2 rejected, P < .01

Table 3.34 Multivariate Analysis of Variance of six *C.* series *Rotundifoliae* populations, testing the equality of group dispersion determinants and the equality of group centriods.

Table 3.35a DISCRIMINANT WEIGHTS

DESCRIPTORS	DISCRIMINANT AXES 1	2	3	4	5
LFLEN	5.56	2.71	-1.47	1.55	0.22
PELEN	-2.94	0.92	2.37	-2.59	3.28
R3412	0.97	0.89	-0.48	2.90	0.63
BANG	0.68	-1.57	-2.38	-2.97	1.47
THLEN	0.74	-2.83	1.85	3.05	4.35
PTLEN	1.16	-1.12	3.79	-1.59	-2.38
EIGENVALUES	1.537	1.350	.653	.217	.025
% VARIANCE	40.6%	35.7%	17.3%	5.7%	0.7%
CUM. VAR.	40.6%	76.4%	93.6%	99.3%	100.0%

Table 3.35b BARTLETT'S TEST

ROOTS REMOVED	χ^2	D.F.	SIGNIFICANCE
0	263.5	30	***
1	165.7	20	***
2	76.0	12	***
3	23.2	6	***
4	2.6	2	NS

***	P < .01
**	P < .025
*	P < .05
NS	No Significant Difference

Table 3.35 Multi-Group Discriminant Analysis of six *C.* series *Rotundifoliae* populations.

3.35a Discriminant Weights (eigenvector elements), Discriminant Roots (eigenvalues) and absolute and cumulative variance accounted for by each Discriminant Axis.

3.35b Bartlett's test of the equality of the Discriminant Axes, with successive roots removed. χ^2 = chi-square test statistic, D.F. = degrees of freedom.

Table 3.35c FACTOR STRUCTURE

DESCRIPTORS	DISCRIMINANT AXES 1	2	3	4	5
LFLEN	.740	.531	.117	-.228	.288
PELEN	-.191	.657	.382	-.329	.540
R3412	-.186	.353	-.029	.591	.051
BANG	.313	-.489	-.479	-.524	.278
THLEN	.270	-.692	.267	.294	.515
PTLEN	.446	-.243	.729	-.243	-.309

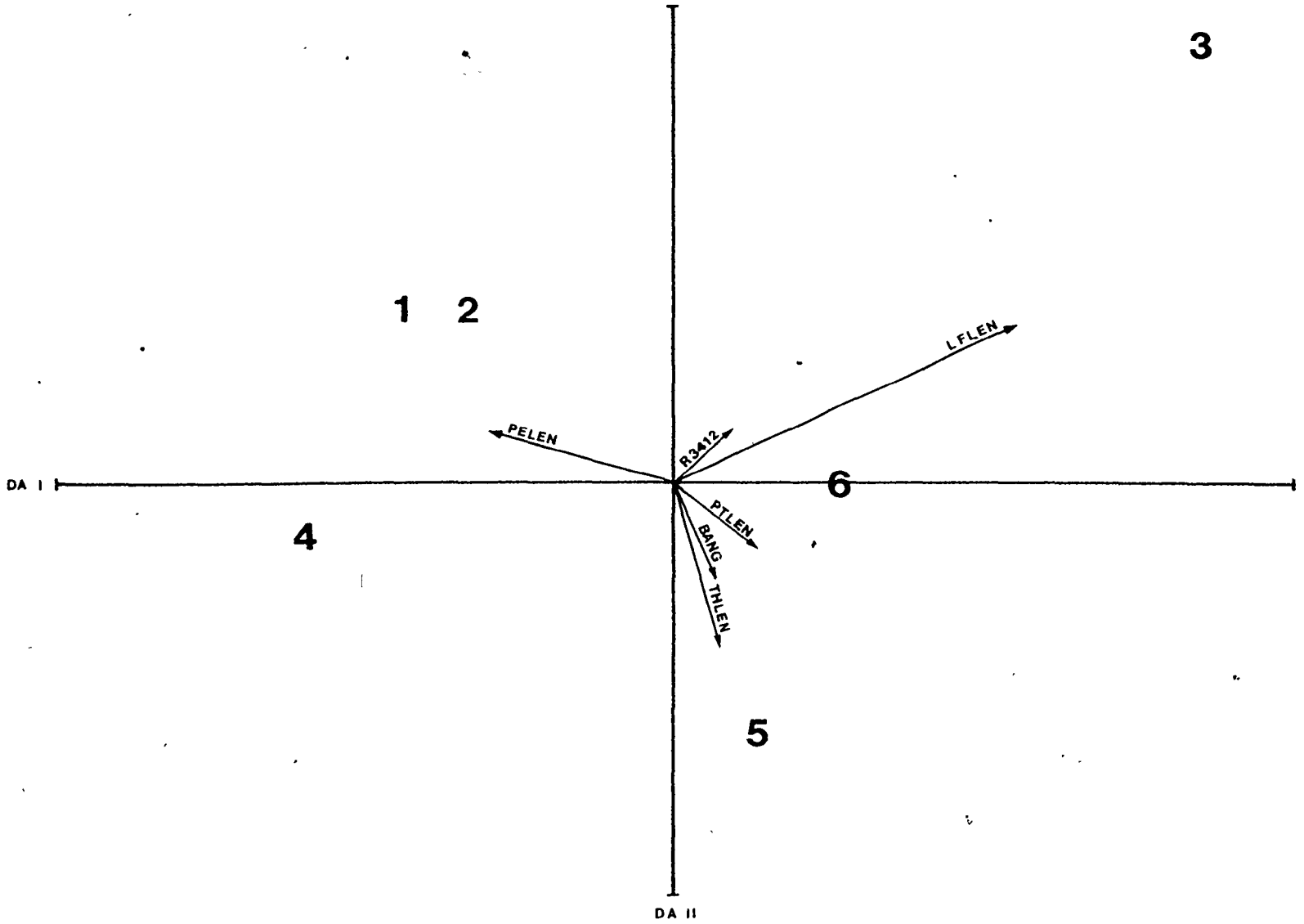
Table 3.35d TOTAL COMMUNALITIES

LFLEN	.977
PELEN	.994
R3412	.512
BANG	.918
THLEN	.975
PTLEN	.943

Table 3.35c Factor Structure coefficients (correlation between each descriptor and each Discriminant Axis).

3.35d Total Communalities (contribution of each descriptor to discriminance amongst the groups).

Figure 3.13 Scatter diagram of the position of the centroids of each of six *Crataegus* series *Rotundifoliae* populations in the space of the first two Discriminant Axes (DA I and II) of a Multi-group Discriminant Analysis based on six ranged descriptors. Superimposed on the diagram are vectors illustrating the loading (Discriminant Weights) of each descriptor on the Discriminant Axes. Population centroids are indicated by the number given to that population in section 3.3.3 of the text. Descriptor abbreviations are explained in Table 3.1.


3
1 2
LFLEN
PELEN
R3412
DA I
6
PTLEN
4
BANG
THLEN
5
DA II

CHAPTER 4

REPRODUCTIVE BEHAVIOR

4.1 INTRODUCTION

The second area of research outlined in Chapter 1 was investigation of certain aspects of reproductive behavior of series *Rotundifoliae* individuals in the random sample. The results of that examination are presented in this chapter. The primary purpose of examining reproductive behavior is to establish how much apomixis (specifically apospory, see below) exists in members of the series in Ontario, as delimited by the cluster analysis in Chapter 3. To this end, histological examinations of ovule development in selected individuals from the random sample were performed. However, polyploidy and hybridization are frequently associated with agamospermy; it is the combined effect of these three processes that gives rise to agamic complexes (Stebbins 1950, Grant 1981). An integral part of establishing the verity of the thesis, then, is examination of the degree to which polyploidy and hybridization have contributed to the formation of the agamic complex in Ontario series *Rotundifoliae*. Levels of polyploidy can be obtained directly through chromosome counts. Investigation of hybridization is not, however, quite so simply undertaken. Here, pollen

viabilities, pollen/ovule ratios, degrees of self- and out-crossing and flowering phenology were all examined in order to obtain information bearing on the possibility of, and insights into the effect of hybridization in this series.

In total, five separate studies of reproductive behavior were performed. These are

1. Histological studies of ovule development, estimating the frequency and establishing the type of agamospermy present in the series.
2. Cytological studies of chromosome numbers, giving ploidy levels.
3. Pollen counts and viability studies to examine pollen/ovule ratios and to determine amounts of viable pollen produced.
4. Pollination experiments to examine the degree of self fertility in the series and to investigate the role of pseudogamy.
5. Phenological studies of flowering to establish the amount of overlap in flowering time and thus obtain estimates of the likelihood of sympatric cross-pollination.

Some clarification of terminology, specifically with respect to apomixis, is in order. In this thesis, following Asker (1979), apomixis is equated with agamospermy (asexual formation of seed). Several pathways to the formation of agamospermous seed exist. Reviews of the process are given in Gustaffson (1946-47), Battaglia (1963), Nygren (1967) and others. Two types of apomixis

are defined. In gametophytic apomixis, an unreduced gametophyte is formed. In adventitious embryony, the gametophyte stage is bypassed and an embryo is formed directly from a somatic cell of the nucellus. In gametophytic apomixis, if the gametophyte (embryo sac) arises from an unreduced embryo mother cell (EMC), this is called diplospory. Apospory occurs when the gametophyte arises from a somatic cell of the nucellus. In diplospory or apospory, the functional embryo can arise either from the unreduced egg cell of the embryo sac (diploid parthenogenesis), or from some other cell of the gametophyte (apogamety). Diploid parthenogenesis appears to be by far the more common pathway in the angiosperms (Grant 1981). Apospory, followed by diploid parthenogenesis is the most common pathway in apomictic *Crataegi* (Muniyamma and Phipps 1979a, 1984b) but diplospory - parthenogenisis also occurs (Muniyamma and Phipps 1984a). Agamospermy may be facultative or obligate. In facultative apomicts, the ratio between sexually formed and apomictically formed embryo sacs may vary from plant to plant or seasonally. In the facultative apomict *Dichanthium aristatum* (Poir.) C. E. Hubbard, the ratio between sexual and apomictic embryo sacs is influenced by environmental (light regime) factors (Knox and Heslop-Harrison 1963).

In many agamospermous plants, embryogenesis requires pollination despite the fact that syngamy does not occur. The pollination process is then referred to as pseudogamy. In the case of pseudogamous agamosperms, then, direct evidence for the presence of agamospermy can only be obtained through direct observation of ovule (Embryo Sac or ES) development. Otherwise, production of seed by emasculated (and isolated) flowers would

be sufficient proof of agamospermy. On the other hand, lack of seed production under such conditions, in a proven agamosperm is sufficient evidence for the presence of pseudogamy.

Even with examination of ovule development, establishing the presence of agamospermy is often very difficult. For example, diplospory can be distinguished from sexual ES development only through observation of meiotic events in the EMC. If this stage of development is passed through quickly, as has been suggested for *Crataegus* (Dickinson 1983), accumulation of evidence for diplospory is extremely arduous. Diplospory may, however, be inferred in triploids showing apparently normal (sexual) development of the gametophyte. On the other hand, apospory is readily established by the presence of multiple ESs or aposporous initial cells that give rise to aposporous ESs.

Many of the investigations presented in this chapter are based on techniques that have already been applied in other investigations of *Crataegus*. As a result, a considerable store of information on *Crataegus* reproductive biology already exists. Some of this information will be discussed in the final section of this chapter. Otherwise, the reader is referred to Muniyamma and Phipps (1979a,b, 1984a,b) and Dickinson (1983) for accounts of developmental stages in both sexual and agamospermous ovules as well as androecial development. Meiotic and mitotic events are described in Muniyamma and Phipps (1979b). Sinnott (1978) and Dickinson (1983) discuss pollen stainability analyses.

4.2 MATERIALS AND METHODS

Because of the labor intensive nature of the techniques presented in this chapter, and the short flowering season, 44 OTUs were examined for floral and reproductive biology. In order to make this sample representative of that used in the morphometric studies, the OTUs initially chosen for this investigation were selected on the basis of their position in the durchschnitt consensus tree (Figure 3.10) resulting from cluster analysis. Individuals from both central and outlying positions were picked from each of the four groups (see Chapter 3). In the pollination experiments (section 4.2.4), sufficient flowers for treatment were a limiting factor. There, a few individuals not in the random sample (but included in the cluster analyses in Chapter 5) were used. The majority of individuals used for examination of reproductive behavior in series *Rotundifoliae* were, however, from the random sample.

4.2.1 Histological Studies

Thirty individuals were examined in this part of the analysis. Samples of 20 or more flower buds, at or near anthesis, were collected from each individual. The buds were fixed in Farmer's fluid (3:1 absolute alcohol : Glacial Acetic Acid) for 24 to 48 hours, and then transferred to 70% alcohol for cold storage.

From each sample, 10 (usually open) flowers were dissected to partially reveal the ovules. These were then dehydrated using a graded alcohol : TBA series according to standard methods (Berlyn and Miksche 1976). The series

was then continued using an increasing percentage of paraffin. Each bud was then cast in a paraffin block for sectioning. The blocks were sectioned at 7 µm thickness with a rotary microtome. Ribbons of sectioned material were fixed to slides using chrome alum (Berlyn and Miksche 1976).

A variety of standard staining techniques were used but Johansen's haemotoxylin (Clark 1973) was found produce the best contrast between nucleoli, nuclei and the cytoplasm. Slides were mounted in permount and stored for later examination. Each ovule was examined to determine whether a single ES (indicating sexual reproduction) or multiple ESs (indicating apospory) were present.

As stated in the introduction, evidence for the presence of diplospory is much harder to obtain than that for apospory. Given that the main thrust of the thesis is to establish the presence of agamospermy (of whatever form) in the series, most of the effort in this analysis was directed towards finding evidence for apospory. Thus, some cases of diplospory may have been overlooked. Both apospory and diplospory have been found in the genus (Muniyamma and Phipps 1979a, 1984a,b, Dickinson 1983).

4.2.2 Cytological Studies

Flower buds in early stages of development were collected from each individual in early spring, split with a razor blade and floated in 0.33% colchicine for 2 hours. This material was then fixed with Farmer's Fluid for 24 to 48 hours and subsequently cold stored in 70% alcohol. Chromosome counts were made, for the most part, using developing petal material in

mitotic division. Of the actively dividing tissue on the plant, developing petals were the easiest to obtain in quantity. Some counts were made using Pollen Mother Cells (PMCs) in meiotic division.

A variety of hydrolysis and staining schedules were tested. The schedule giving the best results (presented below) was used for all chromosome counts.

Material for staining was removed from storage and washed with distilled water. It was then placed in 45% acetic acid for 30 minutes, and then hydrolysed in 5 N HCl for 1 hour. A ceramic spot plate was found to be effective for processing the small amounts of material being used. The material was again washed in distilled water and transferred to leuco-basic fucsin (Berlyn and Miksche 1976) for 3 hours of staining. After staining, cell walls were lysed with a 5% (w/v, aqu.) solution of pectinase (fungal, technical powder, ICN Pharmaceuticals) and cellulase (fungal powder, Calbiochem-Behring Ltd.) for at least 2 hours.

For meiotic counts, anthers were removed from the flower buds and squashed on a glass slide in a few drops of 45% acetic acid to free the PMCs. A coverslip was then applied and the slide was alternately tapped, heated and micropressed to spread the material. For mitotic counts, petal material was macerated in 45% acetic acid with fine forceps, transferred to a glass slide with a micropipette and then heated and micropressed to spread the material.

4.2.3 Pollen Analysis

For this part of the examination of reproductive behavior, mature unopened flower buds were used. Buds were collected, fixed and stored in alcohol as in the previous analyses. A total of 50 anthers from each individual were dissected out, macerated and stained overnight in a solution of Malachite Green - Acid Fuchsin - Orange G (Alexander 1969). Using this solution, pollen walls stain green while cytoplasm stains red (and masks green staining). Thus, aborted grains can be identified by a characteristic green color while non-aborted (presumably viable) grains are red.

Following staining, 0.25 ml glycerin jelly was added and a known volume of the solution examined using an Hematocytometer. Both aborted and viable pollen grains were counted and the counts converted to total and viable pollen grains per anther. Finally, percentage pollen viability for each individual was calculated.

At the time of dissection of anthers, the number of stamens and styles per flower was also recorded. These counts were combined with pollen viabilities to obtain numbers of viable pollen grains per flower and numbers of ovules per flower (based on two ovules per style). Finally, pollen/ovule ratios (P/O ratios) were calculated.

P/O ratios (Cruden 1977, Cruden and Miller-Ward 1981) provide a good estimate of the degree of outcrossing (or inbreeding) practiced by a given plant group. The theory underlying P/O ratios is that any species tends to minimize pollen production towards that level at which optimal seed set is most likely. Autogamous plants tend to produce far fewer pollen grains

per ovule than do xenogamous ones (other factors being equal) because selfing implies more efficient pollen transfer. The increase in P/O ratios in moving from obligate autogamy to obligate xenogamy is logarithmic (Cruden 1977).

4.2.4 Pollination Experiments

The two objectives of this set of experiments were to determine if the populations under study are self-fertile and to determine (given that agamospermy can be shown to exist in the populations) whether the populations are pseudogamous. It was assumed that the populations are largely xenogamous. This assumption seems reasonable, *a priori*, given the results of Dickinson (1983) showing outcrossing in populations of *C. crus-galli*, and those of Power (unpubl.) where floral visitation by pollen carrying bees (primarily *Apis*) was directly observed in populations of *C. crus-galli*.

The pollination experiments were performed in the spring of 1982 at sites 1, 2 and 3 and included individuals from Groups 1, 2 and 4. There was not sufficient time to include site 4 in the experiments.

Three treatments were used in these experiments (Table 4.1). Prior to treatment, the selected individuals were further screened for the presence of sufficient inflorescences (with five or more closed flowers) to perform at least one set of the three treatments. If the individual was chosen, the treatments were performed sequentially until insufficient inflorescences remained for another full set of treatments. After treatment, each inflorescence was tagged and the number of treated buds counted. In

Treatment I (open pollination), the inflorescences were tagged and all flowers counted. In Treatment II (emasculation and isolation) all open flowers in the inflorescence were discarded. Next, each remaining flower was emasculated by removing the stamens with fine forceps, and the number of styles was recorded. Inflorescences were isolated with small, fine mesh polyester bags. Treatment III (manual selfing) was identical to Treatment II except that stigmatic surfaces were smeared with pollen from newly dehisced anthers from the same individual (usually from the culled, open flowers from the same inflorescence). In this way a total of 240 inflorescences from 18 individuals were treated.

After fruit formation had started, an initial assessment of results was made. Isolation bags were removed from those inflorescences that had not formed fruit. Those inflorescences from Treatment I that had begun to develop fruit were covered with isolation bags in order to avoid loss of the fruit.

In the fall, all fruits from treated inflorescences were harvested. The nutlets in each fruit were counted and subsequently split to assess seed set. Occasionally, seeds were aborted (shriveled) even though apparently good fruit had been formed.

Percentage conversion of styles to seed was computed for each Treatment, for each of the three groups represented in the experiment. Aborted seed were given the same score as developed seed, reflecting the fact that seed formation had been initiated. It was assumed that factors other than pollen incompatibility had caused abortion. For Treatment I, the

number of styles was taken as the mean number of styles computed from the other two treatments, for that OTU. For all treatments, 0.1 was added to percentage conversion scores to avoid difficulties in statistical tests using 0% scores. Adjusted percentages were then arcsine-square root transformed to produce more normal sampling distributions (Sokal and Rohlf 1981). Two hypotheses were tested. First, an analysis of variance (ANOVA) was performed to test the null hypothesis that the means for each treatment were the same. Second the null hypothesis that mean percentage transformed scores were equal to .3218 (arcsine-square root transform of 0% + 0.1 adjusted percentage conversion scores) was tested for each treatment. All tests (proceedures TTEST and ONEWAY) were performed using the MINITAB statistical package (Ryan, Joiner and Ryan 1981).

4.2.5 Flowering Phenology

During spring sampling in the period 1979-1983, the time of anthesis of each of the populations was monitored. Primarily for purposes of recollection of flowering material from the population in a subsequent year, the time of anthesis was expressed in terms of Accumulated Heat values for London, Ontario. Accumulated Heat (units = degree days or D.D.) was defined as the number of degrees of mean daily temperature over 5° C, accumulated daily, starting on April, 1 of each year. Each spring, Accumulated Heat values were calculated daily until all populations had gone through anthesis. This information provided a means of predicting anthesis in the populations as well as an estimate of the degree of overlap in flowering amongst the populations. Of particular interest was whether or

not the sequence of flowering (expressed in terms of Accumulated Heat) was consistent from year to year.

In order to address the specific question of whether or not two populations occurring at the same site flower at the same time, a different technique was used. Estabrook et al (1982) have presented an application of the Kolmogorov-Smirnov two-sample test (of the difference between two distributions - see Sokal and Rohlf 1981) that is useful for testing differences in flowering phenology. The test is based upon maximum absolute differences between the relative cumulative frequency distributions of two samples. In the specific application presented here (referred to as Estabrook's test), the number of flowers (on individuals from each of two populations) that had opened since the onset of flowering was monitored at intervals throughout the flowering season. For each interval, the number of flowers that had opened since the onset of flowering was expressed as a fraction of the total number of flowers that eventually opened, for that population. The maximum absolute difference between fractions (one from each population) among all intervals was then used as a non-parametric test statistic. A significant difference in flowering times is indicated when the maximum absolute difference between fractions is greater than a critical value computed (for a given alpha probability level) on the basis of sample sizes. The formula for the $\alpha = 0.05$ critical value is

$$P\,(D > 1.36[(m+n)/(mxn)]^{1/2}) = 0.05$$

where D = maximum absolute difference

m = sample size in population 1

n = sample size in population 2

Tests of differences in flowering time were made at sites 3 and 4. At site 3, one individual each from populations 1, 3 and 4 were monitored. At site 4, two individuals each from populations 2 and 4 were monitored. Tests were not performed at sites 1 and 2 because only one population occurred at these two sites.

4.3 RESULTS

Results from the first three investigations of reproductive behavior in the series are presented in Table 4.2. This table, arranged by population, gives chromosome number, pollen viability, P/O ratio and numbers of ovules with sexual or aposporous embryo sacs for each individual examined. Clearly, information from all the aspects of reproductive behavior studied here could not be obtained for all individuals. At the end of the table, the same information is given for those individuals that were not part of the random sample. Of these individuals, one (PS144) is from site 1 and three (PS108, PS111, PS113) are from site 2. All four of these OTUs were found (Chapter 5) to have taxonomic affinity with Group 4. The other OTU (PS140) is from site 4 and has affinity with Group 1. Chromosome counts are also presented for two OTUs (PS284, PS285) from additional sampling site 10. These OTUs were chosen (Chapter 5) as exemplars of *C. chrysocarpa*.

4.3.1 Histological Studies

In scoring ovules, apospory was determined to be present when one of three conditions was observed. Either multiple embryo sacs were observed

directly (Plate 1a,b), or aposporous initial cells were observed along with a mass of degenerated cells (degenerated megagametophyte) in the middle of the nucellus (Plate 1c), or aposporous initial cells were observed, usually along with a single, developed embryo sac (Plate 1d). In the latter case, it could not be determined whether the single ES was formed via apospory or sexually. Sexual formation of an ES was determined to be present when only a single ES (6 or 8 nucleate) was observed in the ovule (Plate 1e,f). This condition is not differentiable from diplospory.

It is apparent that all of the populations are producing aposporous ESs. The proportion of ovules showing evidence of apospory (Class 2, Table 4.2) is about the same in all populations. It also appears that many are producing sexual ESs, albeit at a low level in most. There is, however, some difficulty with establishing the degree of apospory on the basis of numbers of ovules showing aposporous ESs. The sectioned material represents only the stage of ovule development that was present in the flowers at the time of collection. Subsequent events, specifically the later formation of aposporous ESs in ovule showing only sexual development at the time of collection, may give a different assessment of the degree of apospory in a given individual. Thus, the slightly greater numbers of ovules with single ESs apparent in Group 4 (populations 5 and 6) may be a result of the immaturity of the collections rather than a real difference in the commitment to apospory in the group. This may also explain the presence of ostensibly sexual ESs in OTUs with triploid chromosome numbers. Otherwise, one must assume that a single aposporous ES has arisen (this does not frequently occur) or that the ES is diplosporous.

4.3.2 Cytological Studies

Results from the cytological investigations are presented in Figure 4.1. Here, the first two components of a PCA-correlation are plotted, with ploidy level indicated for those individuals that were studied. Groups 1, 2 and 3 are mostly tetraploid, while Group 4 is mixed triploid and tetraploid. No diploid individuals were found in the analysis. Plate 2 illustrates both meiotic (tetraploid) and mitotic (triploid and tetraploid) divisions.

Some difficulty was encountered in obtaining countable preparations. Thus, relatively few counts were obtained for the amount work involved. Similar difficulty has been encountered in other studies of *Crataegus* cytology (Longley 1924, Muniyamma and Phipps 1979b, Dickinson 1983), and has been attributed to large numbers of very small (< 5 µm.) chromosomes and to the small amount of tissue in which counts can even be sought; when root tips are generally unusable (see Muniyamma and Phipps 1979b).

4.3.3 Pollen Analysis

Percentage viable pollen is plotted against P/O ratio in Figure 4.2. With one exception, all individuals exhibited extremely high pollen viability (>80%). These results are notable given the presence of at least three triploids in the sample (see section 4.4 for further discussion). Pollen/ovule ratios were in the same range as those found in *C. crus-galli* (Dickinson 1983) and as other facultative outcrossers (Cruden 1977). However, Group 4 displays consistently lower P/O ratios than the other groups, possibly indicating a slightly greater reliance on autogamy.

4.3.4 Pollination Experiments

Results from this set of experiments are presented in Tables 4.3 and 4.4. Table 4.3 gives results from the ANOVA portion of the analysis. Table 4.4 presents the results of the t-tests.

Results from Table 4.3 (ANOVA) indicate that in all three groups, the treatment means are significantly different from each other. Examination of Table 4.4 indicates that in all three groups, mean percentage conversion of styles to seed is significantly different from 0 in both Treatment I and Treatment III. Thus, manual self pollination is capable of stimulating seed set. Whether any of the seeds produced were the product of sexual processes could not be determined. It is assumed that at least some of the seeds were the result of the sexual process and, therefore, that the groups are self-fertile.

Results from the t-tests for Treatment II are perhaps the most interesting. Here, the means for Treatment II in Groups 1 and 2 are not significantly different from the transformed 0% score. It is concluded that, in these two groups, pollination is required for seed set and thus, pseudogamy is present. However, the Treatment II mean for Group 4 does show a significant difference. Two interpretations of these results arise. Either the seed set for this Treatment indicates a low (5 seeds of 1159 styles treated) level of non-pseudogamous production of apomictic seed in this group, or experimental technique for this group was poorer than for the other groups (e.g. accidental pollination during emasculation by pollen from contaminated forceps). In any case, it is clear that these groups are at

least predominantly pseudogamous.

4.3.5 Flowering Phenology

No formal tests of differences between populations, with respect to critical Accumulated Heat values for flowering, were made. However, during the course of sampling it became clear that populations 1 through 4 (sites 3 and 4) flowered at lower Accumulated Heat values (computed for London, Ontario) than did populations 5 and 6 (sites 1 and 2). The sequence of flowering at the sites, and associated London Accumulated Heat values were as follows; population 3 220-250 D.D., populations 1, 2 and 4 240-260 D.D., populations 5 and 6 280-300 D.D. This sequence was observed to be consistent over the five year collecting period, although the spacing between population flowering varied from year to year.

Results from Estabrook's test are presented in Table 4.5. It is evident that Groups 1 and 3 do not differ in time of flowering but that Group 2 flowers significantly earlier than both Groups 1 and 3, by about 4 days. It is also clear that, despite differences in peak flowering times, some overlap in flowering (some flowers from all populations open on the same day) occurs at both sites. Thus, at least for 1983 (when the experiment was performed), the possibility for cross-pollination between populations was present.

4.4 DISCUSSION

The results presented in this chapter indicate a suprisingly high degree

individuals of each group appear to be mostly pseudogamous, facultative agamosperms, reproducing both sexually and via apospory. A few individuals showed evidence for exclusive production of aposporous ESs. Low levels of sexuality (10 - 15%, on average) are indicated in those individuals showing evidence for facultative apomixis. All four groups are predominantly (or completely) tetraploid and apparently self-fertile with high pollen viability and P/O ratios in the range of facultative xenogamy.

Group 2 differs, to some degree, from this trend. It is the only group in which no triploids were found. As well, results from Estabrook's test indicated a significant difference in peak flowering (Group 2 flowered earlier) from Groups 1 and 3. Group 4, while sharing most of the characteristics of the other groups, differs to some degree, as well. A greater frequency of sexual ESs was observed in this group, although this difference may be more apparent than real (see section 4.3.1). This was the only group for which firm counts of triploids were obtained. P/O ratios in Group 4 were (with a few exceptions) the lowest of all groups. Finally, results from the pollination experiments indicated that this group may be less commited to pseudogamy for seed set than Groups 1 and 2.

The last two observations may be related in that the lack of an absolute requirement for pseudogamous pollination (indicated for this group) may allow a reduction in pollen production (and thus P/O ratio) without a consequent reduction in seed set.

The requirement for pseudogamous pollination, generally indicated for all groups, may also play a part in maintaining levels of sexuality. If there

were no such general requirement, then sexual embryo sacs may have a reduced chance of being fertilized. Similarly, the presence of a generally pseudogamous habit bears importantly on the possibility of polyploidization via fertilization of unreduced embryo sacs. As discussed in Chapter 6, polyploidy appears to have played an important role in the evolutionary history of series *Rotundifoliae*.

There is also an anomaly in the results that should be noted. This concerns the ability of triploid individuals to produce viable pollen. Of the five triploids for which pollen viability data were available, none displayed viability less than 73%. Given that meiosis in triploids almost invariably produces chromsome disturbances and imbalanced divisions, the high viabilities displayed here present a problem. However, the problem is not unique to this study. Dickinson (1983) found several triploids in 10- and 20-stamen morphotypes of *C. crus-galli*. While most of these displayed pollen viabilities in the range of 10 to 40%, a few showed much higher viability (> 60%). Similarly, Muniyamma and Phipps (1979b) list a possible triploid individual of *C. pruinosa* as having a pollen viability of 66%. Thus, there appear to be sufficient examples of this phenomenon to rule out the possibility of mistaken chromosome counts. An alternate explanation is that the pollen viability estimates, based on the presence of stainable cytoplasm, are based on mistaken assumptions. This could be so if the supposedly viable pollen contained no functional nucleus (Hauser and Morrison 1964). On the other hand, Dickinson (1983) reports crossing experiments in which seed set depended upon pollination by pollen showing high viability. Pollen viability was assessed in that experiment in the same way as in the present study.

Ultimately, intensive studies of pollen meiosis in these triploids are required to fully explain the phenomenon. It may be that only unreduced pollen is being produced. To date, insufficient information exists to clarify the matter further.

The results presented here are generally congruent with the few and limited earlier studies in which members of series *Rotundifoliae* were examined. As stated in Chapter 1, apospory has been found in a member of the series from Ontario (Muniyamma and Phipps 1984). Longley (1924) reported diploid, triploid and tetraploid individuals in the series. Moffett (1931) reports tetraploidy for *C. chrysocarpa*. Muniyamma and Phipps (1979b) identified three ploidy levels for the Ontario *Rotundifoliae*. They list *C. flavida* as possessing diploid, triploid and tetraploid members, *C. dodgei* as diploid, and *C. aboriginum* as tetraploid.

The frequency of triploidy (ca. 20%) displayed in the individuals studied here is very close to the 16% triploidy given by Muniyamma and Phipps (1979b) for Ontario *Crataegi*. This is an important confirmation, given earlier claims (specifically by Longley 1924) of up to 80% triploidy in the genus.

Muniyamma and Phipps (1979b) also report pollen viabilities in the series in the range of 50 to 75%. Sinnott (1978) reports very high pollen viability (ca. 90%) for a series *Rotundifoliae* individual used as an outgroup in his study of series *Pruinosae*.

The major difference between the results presented here and previous reports on series *Rotundifoliae* is the absence of diploids in this study. It may be that the relatively small sample of individuals for which

chromosome counts were obtained is at fault for the lack of diploid counts. However, if diploids are present in the populations studied here, it seems likely that they occur at a relatively low frequency.

TREATMENTS		
I	II	III
OPEN POLLINATION	TEST OF PSEUDOGAMY	TEST OF SELF-FERTILITY
Flowers Counted	Flowers Counted	Flowers Counted
	Flowers Emasculated and Isolated	Flowers Emasculated and Isolated
		Flowers Manually Self-Fertilized

Table 4.1 Tabulation of the three treatments used in pollination experiments testing levels of pseudogamy and self-fertility.

OTU	2N	Pollen Viability	P/O Ratio	Histological Studies Class 1	Class 2	Class 3	Total
GROUP 1							
Population 1							
PS101	68	87%	3066	1	28	3	32
PS103	68	82%	2239				
PS162	68			0	20	4	24
PS169	68						
PS187	68			0	14	7	21
PS191				0	34	0	34
Population 2							
PS115	68?	90%	2669	0	33	1	34
PS174	51?	93%	2636	0	19	1	20
PS200	68	85%	1594	5	26	0	31
PS208	68	90%	1487	3	30	0	33
PS210		92%	4124	0	30	3	33
PS215	68?	88%	2391	1	38	1	40
GROUP 2							
Population 3							
PS160	68?			5	31	0	36
PS175	68	81%	2268	6	21	7	34
PS182	68	96%	2458	0	15	1	16
PS196				6	25	0	31
PS214	68?			1	29	1	31
GROUP 3							
Population 4							
PS104	68?						
PS141				0	20	1	21
PS207	68?	90%	1969	0	37	1	38
PS216	51?	73%	3489	0	24	0	24

Table 4.2 Summary of results from cytological, pollen and histological studies of *C.* series *Rotundifoliae* individuals. P/O ratio = number of viable pollen grains per flower / number of ovules per flower, Histological Studies - Class 1 = # of ovules in which the origin (sexual or aposporous) of the ESs could not be determined, Class 2 = # of ovules showing evidence for aposporous ES development, Class 3 = # of ovules showing evidence of sexual ES development (see text for details).

GROUP 4

Population 5				Histological Studies Class			
OTU	2N	Pollen Viability	P/O Ratio	1	2	3	Total
PS219	68	84%	1081	0	36	2	38
PS221	51	84%	1186	2	28	1	31
PS222	68	89%	1290	0	37	1	38
PS226	51	90%	1314	0	24	0	24
PS228	51?	91%	1056	5	32	4	41
PS232		93%	1807	1	27	1	29
PS237	68	88%	1592	0	29	6	35
PS246	68	93%	980				
Population 6							
PS262	68						
PS263		89%	681	15	13	7	35
PS268	68						
PS271		95%	1604	12	13	7	32
PS272	68?						
PS273	68	94%	1539	2	8	22	32
PS274		94%	1524	5	18	3	26
OTHERS							
PS108		90%	1226				
PS111	51	96%	1491				
PS113		93%	1498	15	10	13	38
PS140	68	91%	2583				
PS144	68	93%	1738				
PS284	68						
PS285	68						

Table 4.2 continued.

Plate 1 Sexual and aposporous embryo sac development in selected *Crataegus* series *Rotundifoliae* individuals.

(a) Multiple aposporous embryo sacs from open flowers of PS175 (*C. aboriginum*).

(b) early development of multiple aposporous ESs in PS141 (*C. flavida*).

(c) aposporous initial cells and mass of degenerated megaspore or megaspore mother cell (arrow) in PS182 (*C. chrysocarpa*).

(d) sexual embryo sac with evidence of developing aposporous initial cells (arrow) in PS191 (*C. ?lumaria*).

(e and f, adjacent sections) sexual embryo sac development in PS221 (*C. flavida*).

all photographs to same scale, given in (c).

Figure 4.1 Diagram showing the ploidy level (triploid or tetraploid) of *Crataegus* series *Rotundifoliae* individuals from the random sample for which chromosome counts were obtained. Crosses represent the position of members of each of the four groups in the space of the first two Principal Components (PC I and II) of the PCA-correlation originally presented in Figure 3.7.

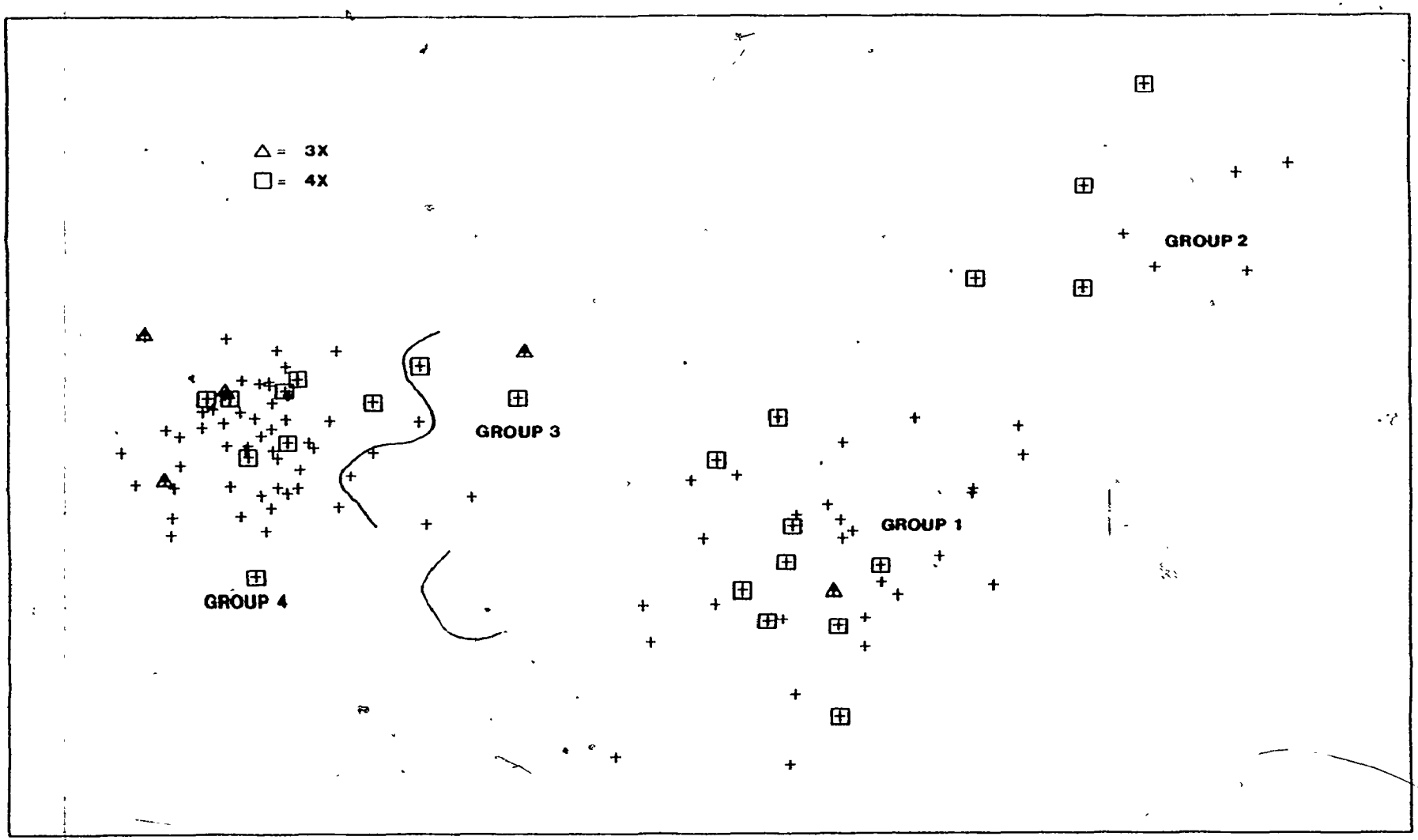

△ = 3X
□ = 4X
GROUP 2
GROUP 3
GROUP 1
GROUP 4

Plate 2 Mitotic metaphase and meiosis in selected *Crataegus* series *Rotundifoliae* individuals.

(a) diakinesis in triploid PS111 (*C. flavida*).

(b) metaphase I in tetraploid PS222 (*C. flavida*), to same scale as (d).

(c) mitotic metaphase in tetraploid PS144 (*C. ?lumaria*), to same scale as (e).

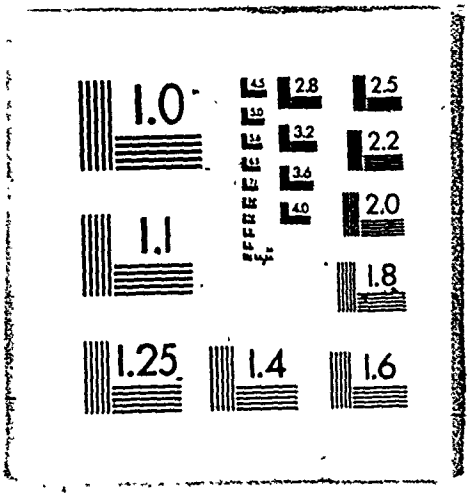
(d) mitotic metaphase in triploid PS226 (*C. flavida*).

(e) mitotic metaphase in tetraploid PS285 (*C. chrysocarpa*).

(f) mitotic metaphase in tetraploid PS208 (*C. ?lumaria*), to same scale as (d).

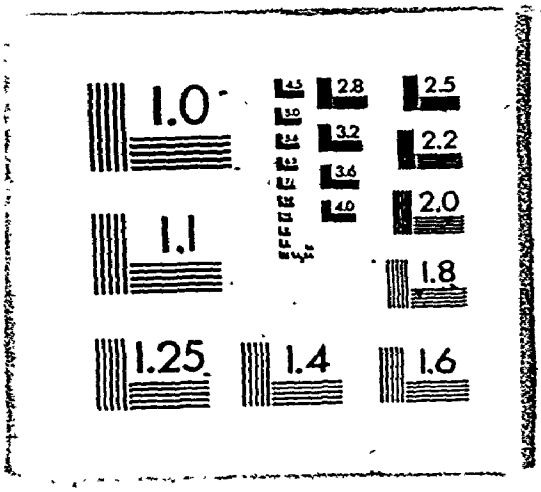
3 3

OF / DE



3 3

OF / DE



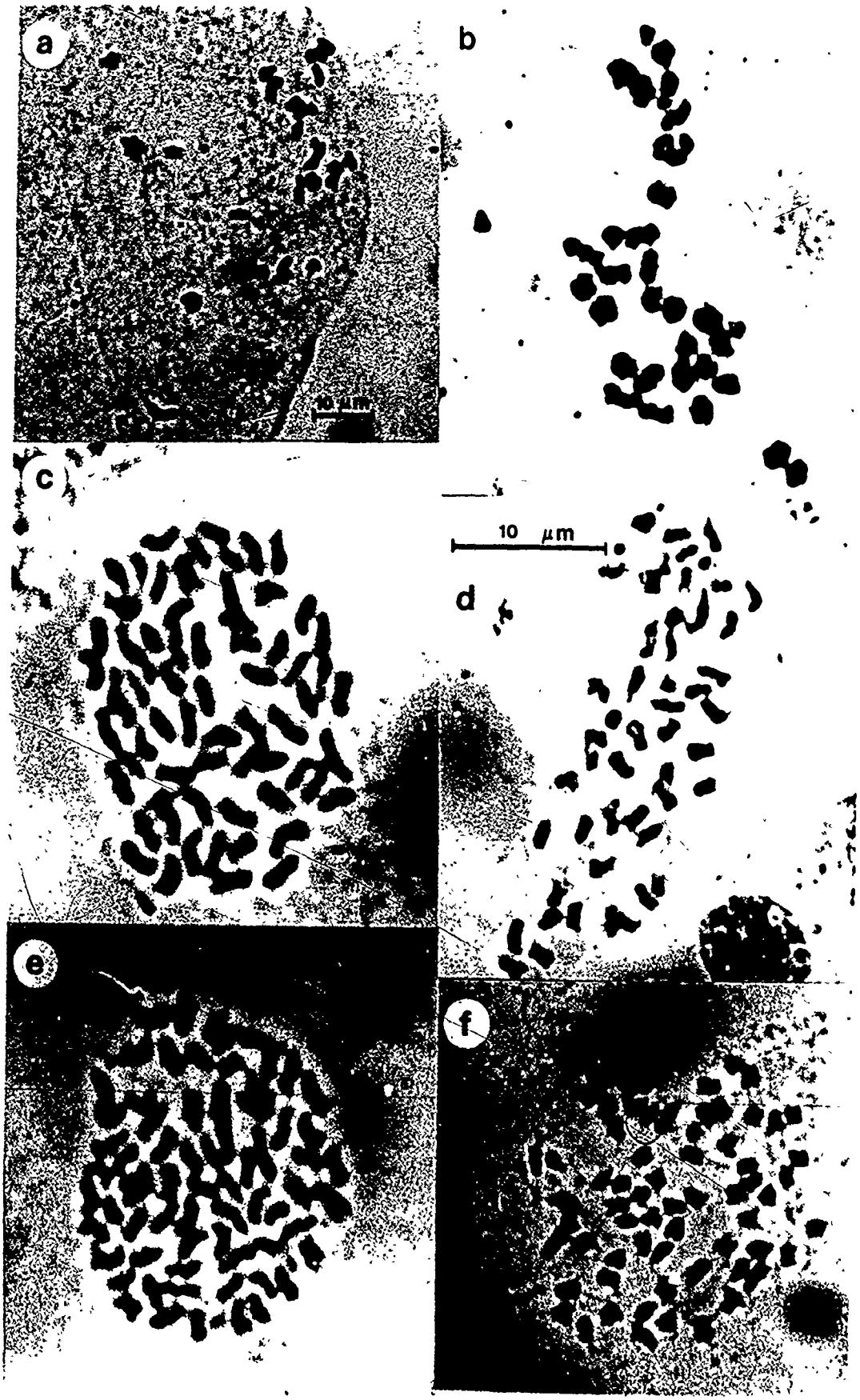
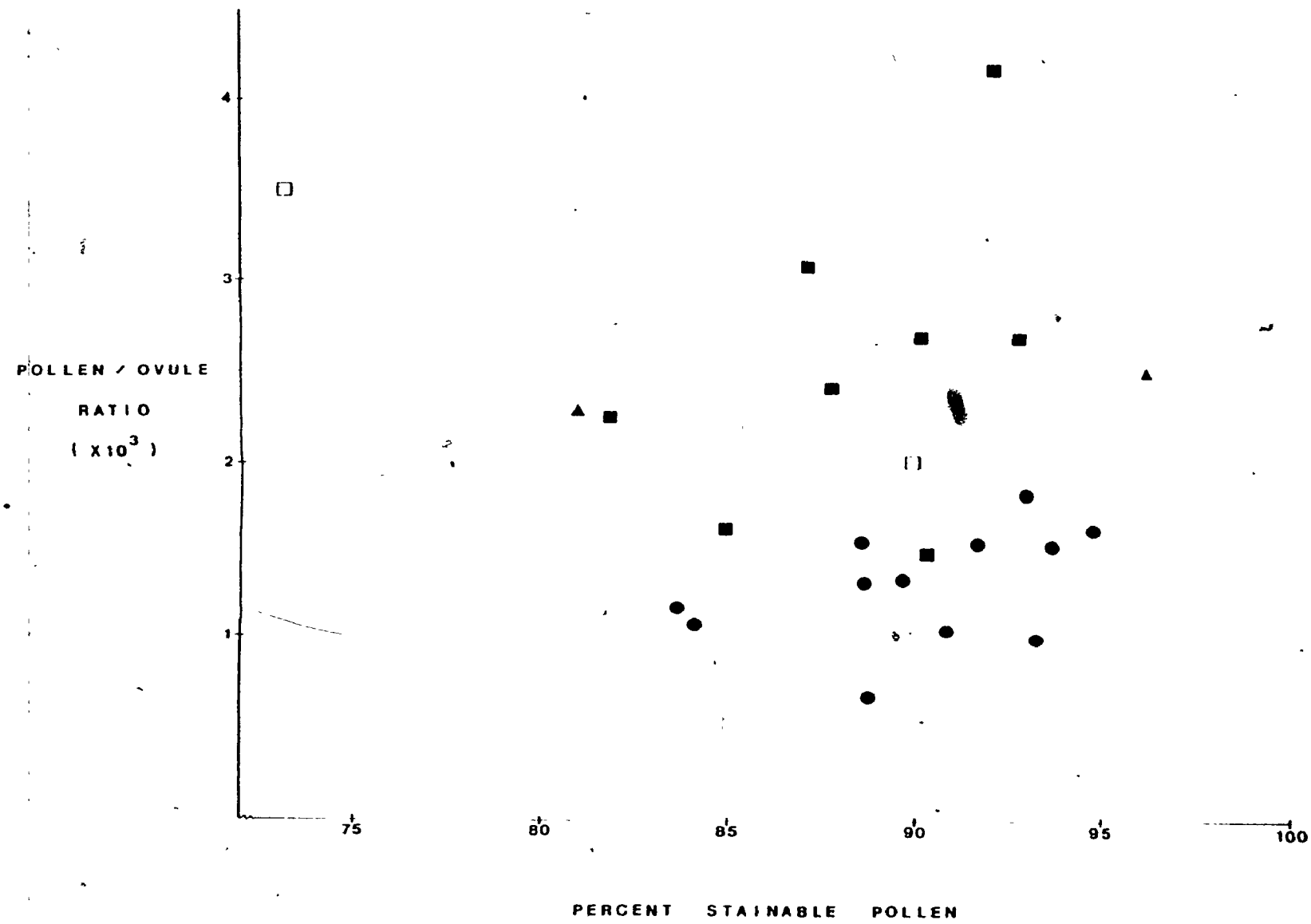


Figure 4.2 Percent pollen stainability and pollen / ovule (P/O) ratios for *Crataegus* series *Rotundifoliae* individuals.
Symbols:
Group 1 - closed squares
Group 2 - closed triangles
Group 3 - open squares
Group 4 - closed circles


POLLEN / OVULE
RATIO
(X10^3)
4
3
2
1
75
80
85
90
95
100
PERCENT STAINABLE POLLEN

ANOVA

GROUP 1

Variance Due To	D.F.	Sums of Squares	Mean Square	F-Ratio
Treatment	2	.1450	.0725	4.41*
Error	34	.5593	.0165	
Total	36	.7043		

GROUP 2

Variance Due To	D.F.	Sums of Squares	Mean Square	F-Ratio
Treatment	2	.0422	.0211	4.46*
Error	9	.0424	.0047	
Total	11	.0846		

GROUP 4

Variance Due to	D.F.	Sums of Squares	Mean Square	F-Ratio
Treatment	2	.9746	.4873	42.98*
Error	151	1.7121	.0113	
Total	153	2.6867		

Table 4.3 Analysis of Variance of transformed percentage conversion of styles to seed for each of the four groups examined in the pollination experiments. Asterisks indicate significant ($P<.05$) differences between means.

GROUP 1

Treatment	n	mean	t
I	13	.410	2.549*
II	12	.361	1.779
III	12	.513	3.953*

GROUP 2

Treatment	n	mean	t
I	4	.441	2.617*
II	3	.343	.998
III	5	.493	6.383*

GROUP 4

Treatment	n	mean	t
I	51	.520	11.205*
II	51	.329	2.239*
III	52	.461	7.629*

Table 4.4 Summary of univariate t-tests of the null hypothesis that (arc-sine square-root + .1 transformed) treatment means are equal to .3218 (transform of 0% conversion of styles to seed). n = sample size (# of inflorescences treated), * = significant (P< .05) difference indicated.

SITE 3

	GROUP 1			GROUP 3	
Date	Count	Fraction	Difference	Fraction	Count
May 30	0	0	0	0	0
31	0	0	0	0	0
June 1	0	0	0	0	0
3	0	0	.076	.076	4
5	3	.081	.221	.302	16
7	26	.703	.043	.660	35
10	37	1.00	0	1.00	53

	GROUP 1			GROUP 2	
Date	Count	Fraction	Difference	Fraction	Count
May 30	0	0	.080	.080	4
31	0	0	.140	.140	7
June 1	0	0	.240	.240	12
3	0	0	.840	.840	42
5	3	.081	*.919	1.00	50
7	26	.703	.297	1.00	50
10	37	1.00	0	1.00	50

	GROUP 2			GROUP 3	
Date	Count	Fraction	Difference	Fraction	Count
May 30	4	.080	.080	0	0
31	7	.140	.140	0	0
June 1	12	.240	.240	0	0
3	42	.840	*.764	.076	4
5	50	1.00	.698	.302	16
7	50	1.00	.340	.660	35
10	50	1.00	0	1.00	53

Table 4.5 Estabrook's Test of differences in flowering time between two groups. Count = # of flowers that had opened by the specified date, Fraction = Count expressed as a fraction of the number of flowers that eventually opened, Difference = absolute difference between Fractions in the two groups being examined (maximum Difference values are underlined for each test), asterisks indicate a significantly (P<.05) large Difference (see text for details).

SITE 4

	GROUP 1			GROUP 3	
Date	Count	Fraction	Difference	Fraction	Count
May 30	18	.180	.040	.140	17
31	29	.290	.030	.260	31
June 1	39	.390	.020	.410	48
3	91	.910	.020	.930	110
5	98	.98	.01	.99	117
7	100	1.00	0	1.00	118

Table 4.5 continued.

CHAPTER 5

TAXONOMIC CONSIDERATIONS

5.1 INTRODUCTION

The results presented in the two previous chapters have pertained only to four local populations of series *Rotundifoliae* in the vicinity of London, Ontario. In Chapter 3, the question of the correct name and taxonomic rank for the four groups discovered by cluster analysis was left unanswered. This question can be answered by linking morphometric variation in the groups with morphometric variation in a sample that includes exemplars of the species presumed to be present in the random sample. As stated in Chapter 2, the choice of the eight additional sampling sites was made in order to fulfil this goal. Thus, the present chapter will be devoted to taxonomic considerations, establishing the most appropriate taxonomic name for all of the different entities accepted here.

5.1.1 *Crataegus* series *Rotundifoliae* species in Ontario.

Phipps and Muniyamma (1980) provide the most recent (and most intensive) treatment of *Crataegus* in Ontario. As such, their revision represents a valuable starting point for consideration of the taxonomy of the Ontario *Rotundifoliae*. As previously stated, they recognize four fairly common to common species in the series and two rare ones. These six species are described (summarizing from Phipps and Muniyamma, 1980) below.

C. chrysocarpa Ashe (1901) is one of the most northern of the North American *Crataegi*, ranging from the foothills of the Rocky Mountains in Colorado, through the northern Great Plains and southern Priarie Provinces to the northern Great Lakes and Quebec to New England and all of the Atlantic Provinces of Canada. A collection has been made as far south as Wytheville, Virginia, in the Appalachians. *C. chrysocarpa* is distinguished from other Ontario *Rotundifoliae* by its smallish, rhombic to rhomb-ovate, serrate and sharply lobed leaves, usually pubescent corymbs, glandular-serrate calyx lobes and 10 stamens.

C. irrasa Sarg. (1903c) appears to be a 20 stamen variant of *C. chrysocarpa*, although it has not been noted in the western part of the range of *C. chrysocarpa*. Except for stamen number, Ontario *C. irrasa* individuals are morphologically very similar to *C. chrysocarpa*.

C. dodgei Ashe (1903) has an eastern to mid-western distribution, mostly south of the Great Lakes from Wisconsin and Illinois to Connecticut. The exact southern limit of its distribution is improperly understood. *C. dodgei* can be distinguished from other Ontario *Rotundifoliae* by its round-ovate

to ovate, slightly lobed leaves (the lobes generally not sharply acute as in *C. chrysocarpa*), glabrous corymbs and generally obtuse leaf apices. It is relatively scarce in Ontario.

C. margaretta Ashe (1900) is also distributed mostly south of the Great Lakes, from Iowa to western Pennsylvania, but is an inland, midwestern species reaching south to Missouri and perhaps northern Kentucky. *C. margaretta* possesses 20 stamens and, in Ontario at least, most nearly resembles *C. dodgei* in other morphological characters (except for the absence of thorns). However, it has very small anthers and material from outside Ontario often has elliptic leaves.

C. flavida Sarg. (1907b) is quite restricted in distribution, occurring primarily in southwestern Ontario, with extension into adjacent parts of Michigan. *C. flavida* can be distinguished by its ovate to obovate, more or less convex, somewhat sharply lobed leaves, long thorns, glabrous to sparsely long villous corymbs and entire calyx lobes.

C. aboriginum Sarg. (1903b) ranges from Wisconsin to New England. In general, it is sympatric with the southern edge of the range of *C. chrysocarpa* in its region of distribution. It can be distinguished by its rhombic to rhomb-ovate, irregularly lobed leaves, long-villous pedicels, glabrous hypanthia and glandular-serrate calyx lobes.

Figure 5.1 illustrates the differences between these six species.

In their revision, Phipps and Muniyamma (1980) observe that an imperfect morphological discontinuity exists between *C. chrysocarpa* and *C.*

dodgei. They make use of a rather broad definition of *C. aboriginum* to carve out a middle ground between these other two species. As well, both narrowly and widely varying species are recognized in the treatment. Thus, a species like *C. flavida*, with very narrow limits of morphological variation, is given the same taxonomic rank as *C. chrysocarpa*, which displays rather wide morphological variation. Also, they recognize apparent intermediates between *C. chrysocarpa* and *C. macrosperma* Ashe and/or *C. flabellata* (Bosc ex Spach) K. Koch (both in series *Tenuifoliae*) which they provisionally assign to *C. aboriginum*, thus generating a quite wide and polytypic range for *C. aborigunum*. They explicitly state that this species is a grab bag for intermediates.

5.1.2 Other Treatments

Previous revisions of *Crataegus*, specifically Palmer (1946) and Kruschke (1965), have dealt with the genus over a wider geographical range (Kruschke's treatment is essentially for Wisconsin with extrapolation elsewhere). In addition, Sargent revised *Crataegus* for Ontario in 1908 and his treatment effectively covers a wider area set beside his treatments for Michigan (1907b), New York (1903c) and Quebec (1901). Sargent's (and others') views are largely summarized in Palmer (1925). With respect to series *Rotundifoliae* taxa thought to occur in Ontario, all these treatments (Palmer 1925, 1946, Kruschke 1965, Phipps and Muniyamma 1980) vary to a greater or lesser extent from each other. These four different treatments are abstracted in Table 5.1. All of the series *Rotundifoliae* taxa that have been described for Ontario or are held to occur in Ontario are included in the table. Many others that

have been described for the series and do not occur in Ontario are ignored in this discussion, or else have been reduced to synonymy (given in parentheses) with one of the taxa given in Table 5.1 or been transferred to another series. Palmer's (1963b) rather terse treatment in Gleason and Cronquist's Manual of Vascular Plants is virtually identical to his 1946 treatment, with the exception of the absence of *C. rotundata*.

What is common to all the treatments presented in Table 5.1 is that two 10 stamen species (*C. chrysocarpa* and *C.dodgei*) and their 20 stamen relatives (*C. irrasa* and *C. margaretta* respectively) are recognized. If, as Phipps and Muniyamma (1980) suggest, more or less continuous variation exists between *C. chrysocarpa* and *C. dodgei*, then the differences between the treatments in Table 5.1 may simply constitute differences in taxonomic delineation of that intermediate variation.

Before pursuing the question of the various interpretations of the variation encountered in the *Rotundifoliae*, the nomenclature of the taxon referred to here as *C. chrysocarpa* Ashe will be dealt with. During much of the very early (pre-1901) description of the North American *Crataegi*, this species (along with a host of other, very different taxa) was referred to as *C. coccinea* L.. Sargent (1901), however, noted that a considerable degree of confusion surrounded the name *C. coccinea* and that the binomial could not be referred to a type specimen. He suggested that the name be abandoned. Eggleston (1908) also noted the great degree of confusion in the use of *C. coccinea*. However, it was not until Rehder, Palmer and Croizat (1938) that a formal proposal to reject *C. coccinea* as a *nomen ambiguum* was made. Kruschke (1956) provides a detailed recapitulation of the arguments for

rejecting the name *C. coccinea*.

The next available name that Sargent recognized after rejecting *C. coccinea* was *C. rotundifolia* Moench (not Lam.) (Kruschke 1956). However, as Rehder, Palmer and Croizat (1938) note, this binomial had been used for another species by Lamarck, prior to Moench's description, thus invalidating its use for this taxon. Cinovskis' (1971) use of *C. horrida* Medik. for this taxon is clearly mistaken, given that Medikus (1793) simply renamed *C. rotundifolia* Moench. as *C. horrida*, without providing a justification for the change or providing a description of the taxon.

Given the rejection of *C. rotundifolia*, the next available name for the taxon is either *C. columbiana* Howell or (as has been used here) *C. chrysocarpa* Ashe. Howell (1898) described *C. columbiana* from Washington State. Ashe (1901) described *C. chrysocarpa* from Colorado. The two species have long been considered sister species with distributions either east (*C. chrysocarpa*) or west (*C. columbiana*) of the Rocky Mountains. While a number of examples of sister species with similar distributions exist, it is possible that these two entities are conspecific subspecies. If further investigation reveals conspecificity, *C. columbiana* has priority as the specific epithet, given its earlier description. At present, however, cursory examination of specimens by J. B. Phipps (pers comm.) of *C. columbiana* suggest that it is at least varietally distinct. For the remainder of the thesis, conventional use of *C. chrysocarpa* for the taxon occurring east of the Rockies will be followed.

Given acceptance of the name *C. chrysocarpa* until the *C. columbiana* question can be resolved, the question of the interpretation of the variation existing between *C. chrysocarpa* and *C. dodgei* then arises. Ignoring the species in Table 5.1 that were not given recognition subsequent to 1925, five taxa can be associated (on the basis of morphological similarity) with *C. chrysocarpa*. These are *C. laurentiana* Sarg., *C. brunetiana* Sarg., *C. faxonii* Sarg. and *C. chrysocarpa* var. *phoenicea* Palmer. Four taxa can be similarly associated with *C. dodgei*. These are *C. lumaria* Ashe (*C. dodgei* var. *lumaria*), *C. flavida* Sarg., *C. rotundata* Sarg. and *C. delosii* Sarg..

Of the taxa associated with *C. chrysocarpa*, *C. laurentiana* can be can be distinguished by its more or less oblong leaves, darker twigs, sporadic occurrence and 10 pink anthers. Phipps (pers. comm.) suspects that *C. laurentiana* var. *laurentiana* is likely a *C. chrysocarpa* X *C.macracantha* hybrid. *C. faxonii* differs in its densely villous leaves and corymbs and more ovate leaves, but does not appear to be completely discontinuous from *C. chrysocarpa* in these respects. *C. brunetiana* (= *C. laurentiana* var. *brunetiana*) possesses dense pubescent corymbs and larger leaves. Here again, variation is not discontinuous with *C. chrysocarpa*. *C. aboriginum* and *C. chrysocarpa* var. *phoenicea* both possess glabrous hypanthia. *C. aboriginum* shows glabrous to long villous pedicels and large leaves, flowers and fruit. *C. chrysocarpa* var. *phoenicea* has glabrous pedicels. All of the taxa mentioned above share rhombic to isodiametric, sharply and often doubly serrate leaves and glandular-serrate calyx lobes.

C. lumaria differs from *C. dodgei* in the possession of pubescent corymbs. *C. delosii* differs in larger, ovate leaves, occasionally dentate calyx

lobes and smaller flowers. *C. flavida* is distinguished by its somewhat larger thorns and obovate, convex leaves. *C. rotundata* differs in its more broadly ovate, thicker leaves and larger flowers.

It is not within the scope of this thesis to provide a critical appraisal of the relative merits of each of the treatments presented in Table 5.1. Rather, it is pertinent here to attempt to associate each of the four groups with one or more of the 13 taxa discussed above. The question of the correct taxonomic rank for each of the groups can be discussed subsequently.

5.2 MATERIALS AND METHODS

The association of each group with a name from the list of 13 possible ones was carried out in two ways. First a cluster analysis was performed, using the additional (exemplar) collections discussed below. This cluster analysis was used to determine the affinity of each of the groups with exemplars of the species thought by Phipps and Muniyamma (1980) to occur in Ontario. Second, voucher specimens of all collections were compared with type descriptions and (where possible) type or other authentic specimens of the 13 taxa, to determine which description best fit each group. Type specimens and other material were examined both at UWO and at A (herbarium acronyms follow Holmgren and Keuken 1974).

5.2.1 Site Selection

As stated in Chapter 2, eight additional sampling sites (sites 5 to 12) were chosen for this part of the analysis. Of these, three (sites 5, 6 and

7) were chosen because they were thought to contain the broadly defined *C. aboriginum*. Site 8 is located on the southern extension of the Niagara Escarpment at the Royal Botanical Gardens in Hamilton, Ontario. It apparently represents the southernmost station for *C. chrysocarpa sens. str.* in Ontario. Representatives of *C. dodgei* also occur here. Sites 9 and 10 represent more or less pure *C. chrysocarpa* sites, within the normal geographical range of this species. Site 11 represents *C. dodgei* from near its type locality. Site 12 contains three individuals of *C. margaretta* from southern Michigan.

Additionally, five individuals from the major collection sites (sites 1 to 4) that were not part of the random sample, but that were used in the analyses presented in Chapter 4, were included in the cluster analysis.

5.2.2 Cluster Analysis

Cluster Analysis of the entire collection of OTUs (random plus additional collections) was carried out in the same way as the cluster analysis of the random sample presented in Chapter 3. The 73 additional OTUs were scored for the same 29 characters (table 3.1). Next, the scores for all 185 OTUs were ranged using Gower's (1971) method. The same four RM-CM combinations were used. A durchschnitt consensus tree (Neumann 1983) of the four phenograms thus produced was then computed. As well, a Principal Components Analysis, based on the descriptor correlation matrix, was also performed.

5.3 RESULTS

The four phenograms, representing results from application of the four RM-CM combinations, are presented in Figures 5.2 to 5.5. The durchschnitt consensus tree is presented in Figure 5.6 (the actual sequence of OTU numbers in tis tree is given in Appendix 4). The PCA plot, giving the position of each OTU in the space of the first two Principal Components, is given in Figure 5.7.

Several important trends are evident in both the consensus tree and the PCA plot. It is clear that both Group 3 and Group 4 (with the exception of one OTU) remain distinct. They cluster on the right-hand side of the consensus tree with two other groups. The first of these other groups consists of most of the *C. dodgei* OTUs from site 11 and the possible *C. aboriginum* population from site 7. The second is a mixture of *C. dodgei* OTUs from sites 8 and 11, one of the *C. margaretta* exemplars, the outlier from Group 4 and the *C. chrysocarpa* exemplar from site 10. All of the above-mentioned groups occur on the left-hand side of the PCA plot. There, Group 4 occupies the left-most position. Group 3 is above and to the right of Group 4, while the other two groups (and outliers) are below and to the right. A fair degree of overlap between these groups is apparent.

On the left-hand side of the consensus tree, Groups 1 and 2 appear to intermix with the *C. chrysocarpa* and *C. aboriginum* exemplar collections from sites 5, 7, 9 and 10. However, upon closer examination, two fairly large groups and several smaller, outlying groups are revealed, all of which fuse at the same level. The left-most of these clusters consists of the majority

of the Group 1 OTUs plus an additional OTU from site 4 (PS 144). The next cluster to the right is made up of all of the *C. chrysocarpa* exemplar collections, the majority of the Group 2 OTUs and the possible *C. aboriginum* collections from site 6. The outlier clusters include several Group 1 OTUs, one Group 2 OTU, the possible *C. aboriginum* collections from site 5 and the other two *C. margaretta* collections from site 12. In the PCA plot, the outliers identified above are indicated by a semi-circle under the symbol identifying their position. These outliers are located generally to the left of the two main clusters identified above. The left-most cluster in the consensus tree (mostly Group 1) is in the upper right of the PCA plot. The other cluster (Group 2 and *C. chrysocarpa* exemplars) occurs in the lower right of the plot. Some overlap of these two clusters is apparent, at least in the space of the first two Principal Components.

In summary, the four groups identified in Chapter 3 are retained to a large extent in this analysis. Groups 3 and 4 are almost completely unmodified. Group 2, although it clusters with the *C. chrysocarpa* exemplars, is not split up. In Group 1, several OTUs were moved to outlying positions as a result of including the additional collections. Exemplars of *C. dodgei* clustered separately, while *C. chrysocarpa* exemplars clustered with the Group 2 OTUs. The possible *C. aboriginum* collections clustered with different groups; individuals from site 5 occupied an outlying position, those from site 7 clustered with *C. dodgei* exemplars and those from site 6 clustered with *C. chrysocarpa* exemplars. The three *C. margaretta* OTUs were also split up, with one clustering with the *C. dodgei* exemplars and the other two becoming outliers.

5.4 DISCUSSION

Upon examination of voucher specimens of Group 2 and comparison of them with type descriptions, it was determined that members of this group represent *C. chrysocarpa* Ashe *sensu stricto*. Two of the exemplars that clustered with Group 2 (PS284, PS285) were found to be tetraploids (see Table 4.2). The collections from additional site 6 (that also clustered in this group) were tentatively identified as *C. aboriginum*, although some aspects of the leaf shape of members of this population suggest influence (possibly via hybridization) from series *Tenuifoliae*. Outliers associated with the *C. chrysocarpa* core, specifically PS175 (from site 3, a tetraploid -see Table 4.2) and PS180 and PS181 (site 5) were identified as *C. aboriginum*. Within the *C. dodgei* collections, two subgroups and a few outliers were present. The first of the subgroups represents more or less pure *C. dodgei* (sites 7 and 11). The second subgroup (from site 8) is morphologically very similar to the first and thus was tentatively identified as such but may, upon further investigation prove to be *C. delosii*. One of the outliers (PS292, from site 10) located on the far right of the consensus tree appears to be the only collection of *C. chrysocarpa* var. *phoenicea* in the sample.

Groups 3 and 4 fall best (almost exactly) within the type description of *C. flavida*. The outlier from Group 4 that clustered with the *C. dodgei* exemplars (PS276, a triploid -see Table 4.2), clearly falls under the description of *C. dodgei*. Groups 3 and 4 can be most easily distinguished from each other on the basis of fruit color. Members of Group 3 posess consistently

more green-spotted fruit. Other, less obvious differences also occur (see Chapter 3).

C. flavida has been made a synonym of *C. dodgei* by both Palmer (1946) and Kruschke (1965). While there certainly exists close morphological similarity between these two entities, the consistency of the morphological differences between them that have been displayed in this study seem to warrant some taxonomic recognition. It may be that *C. flavida* could best be treated as a variety of *C. dodgei*. However, a formal proposal to that end cannot be justified until more of the variability within *C. dodgei* is understood. For the remainder of the thesis, then, Phipps and Muniyamma's (1980) interpretation of *C. flavida* as a distinct species (albeit a local one of narrow morphological variation) will be followed.

Random sample Group 1 remains to be discussed. This Group occupies a truly intermediate position between *C. chrysocarpa* and *C. dodgei*, in terms of leaf shape, corymb pubescence and calyx margination (e.g. the major characters used to differentiate the two). The corymbs of this group are sparsely long-villous to pubescent with occasionally mostly glabrous hypanthia. In these characters it most resembles *C. chrysocarpa sensu lato*. The calyx lobes, however, are mostly entire and eglandular (occasionally somewhat dentate), which is indicative of *C. dodgei*. Leaf shape variability is wide within the group, with short shoot leaves usually more or less long-ovate but ranging between orbicular and rhombic. Nomenclaturally, some of the material in this group comes within the type description of *C. lumaria*. A few other individuals, if viewed in isolation, could be identified as *C. chrysocarpa s.l.*. However, to break up the group along these lines would

hide more than it discloses. It is not useful to widen the descriptions of two otherwise rather different species in order to divide an internally consistent (if extremely local) third group. On the other hand, to provide a new name for this group, and perpetuate yet another *Crataegus* binomial, on the basis of a regional (rather than a continental) understanding of the series would be irresponsible. Let it stand that, as in the case of *C. flavida*, the group probably deserves taxonomic recognition, but at what level that recognition should be remains unclear. Because some name would be useful for referring to the group, and because the type description of *C. lumaria* encompasses a fair portion of the variation in the group it will henceforth be referred to as *C. ?lumaria*. The rather obvious speculation that this group reprsents hybridization between the more northerly *C. chrysocarpa* and the more southern *C. dodgei* will be discussed in Chapter 6.

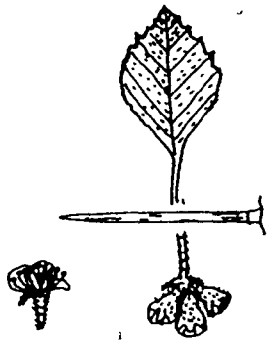
Use of the name *C. aboriginum* in this thesis remains to be clarified. In 1920, Sargent re-evaluated the status of *C. aboriginum* and made it a variety of *C. rotundifolia*, the name he was then applying to material later referred to as *C. chrysocarpa*. Kruschke (1965) formally made *C. aboriginum* a variety of *C. chrysocarpa*. As previously stated, Phipps and Muniyamma (1980) made use of the name *C. aboriginum* as a sort of catch-all for *chrysocarpa-dodgei* intermediates. It is suggested here that many of the individuals placed by Phipps and Muniyamma into *C. aboriginum* are better considered under *C. ?lumaria*, because these individuals fall more closely under the type description of that taxon. *C. aboriginum*, then, becomes more closely allied with *C. chrysocarpa*, in that those individuals identified here

as *C. aboriginum* possess larger, more serrate, generally pubescent and more-or-less isodiametric leaves, long villous pedicels and glabrous hypanthia. This treatment of *C. aboriginum* reflects Sargent's and Kruschke's association of *C. aboriginum* with *C. chrysocarpa*, as a variety of it.

Although only one *C. dodgei* OTU (PS276) was included in the random sample, it seems clear that this species is locally fairly common both east (site 7) and west (site 11) of the study area.

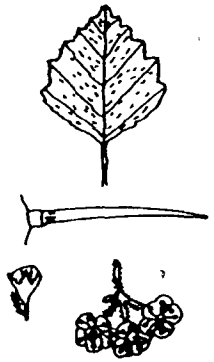
In summary, the four groups revealed in Chapter 3 can be identified as follows; Group 1 - *C. ?lumaria*, Group 2 - *C. chrysocarpa* (with a few *C. aboriginum* outliers), Group 3 and 4 - *C. flavida* (with one *C. dodgei* from site 1). A list of OTU numbers and collection site for members of each of these species is given in Appendix 2. Voucher specimens of all OTUs used in the analysis here and in Chapters 3 and 4 are deposited at UWO.

Figure 5.1 Leaves, thorns and flowers illustrating the differences between the six *Crataegus* series *Rotundifoliae* species occurring in Ontario (based on Phipps and Muniyamma 1980).

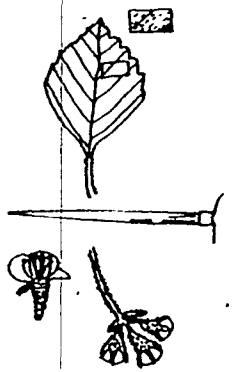


C. irrasa

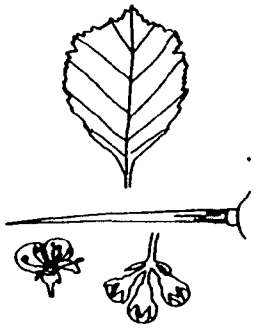
C. margaretta



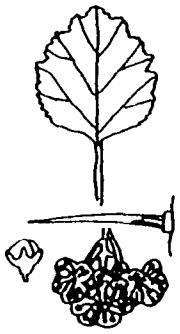
C. aboriginum



C. chrysocarpa



C. flavida



C. dodgei

CRATAEGUS § ROTUNDIFOLIAE TAXA

IN ONTARIO

PALMER (1925)	PALMER (1946)	KRUSCHKE (1965)	PHIPPS and MUNIYAMMA (1980)
C. blanchardi	*(C. irrasa)*	*C. irrasa var. blanchardi*	
C. brunetiana	*C. brunetiana*	*C. laurentiana var brunetiana*	*(C. chrysocarpa)*
C. chrysocarpa	*C. chrysocarpa*	*C. chrysocarpa*	*C. chrysocarpa*
	C. chrysocarpa var. phoenicea	*C. chrysocarpa var. phoenicea*	*C. chrysocarpa var. phoenicea*
C. coccinata	*(C. chrysocarpa)*		
C. crassifolia	*(C. dodgei)*	*(C. dodgei)*	*(C. dodgei)*
C. delosii	*(C. chrysocarpa)*		*C. delosii*
C. dodgei	*C. dodgei*	*C. dodgei*	*C. dodgei*
C. dodgei var. lumaria		*C. dodgei var. lumaria*	
C. faxonii	*C. faxonii*	*C. faxonii*	*(C. chrysocarpa)*
C. flavida	*(C. dodgei)*	*(C. dodgei)*	*C. flavida*
C. inaudita			
C. irrasa	*C. irrasa*	*C. irrasa*	*C. irrasa*
C. laurentiana		*C. laurentiana*	
C. margaretta	*C. margaretta*	*C. margaretta*	*C. margaretta*
C. minutiflora	*(C. chrysocarpa)*	*(C. chrysocarpa)*	*(C. dodgei)*
C. praecoqua		*C. faxoni var. praecoqua*	
C. rotundata	*C. rotundata*	*C. dodgei var. rotundata*	*(C. dodgei)*
C. rotundifolia var. aboriginum	*(C. chrysocarpa)*	*C. chrysocarpa var. aboriginum*	*C. aboriginum*

Table 5.1 Names applied to species and varieties of Ontario members of *C.* series *Rotundifoliae*, comparing the treatments of different authors. Synonmy is indicated by parentheses.

Figure 5.2 Phenogram resulting from Cluster Analysis of 185 OTUs (random plus exemplar collections) based on 29 ranged descriptors. The RM-CM combination used was Sums of Squares Agglomeration of Euclidean Distances. Scale at left represents the resemblance level (Euclidean Distance) at which groups of OTUs fused.

Symbols:
Group 1 - closed squares
Group 2 - closed triangles
Group 3 - open squares
Group 4 - closed circles
C. chrysocarpa exemplars from sites 8, 9 and 10 - C
C. dodgei exemplars from sites 8 and 11 - D
C. margaretta exemplars from site 12 - M
C. aboriginum exemplars from site 5 - A5
C. aboriginum exemplars from site 6 - A6
C. aboriginum exemplars from site 7 - A7

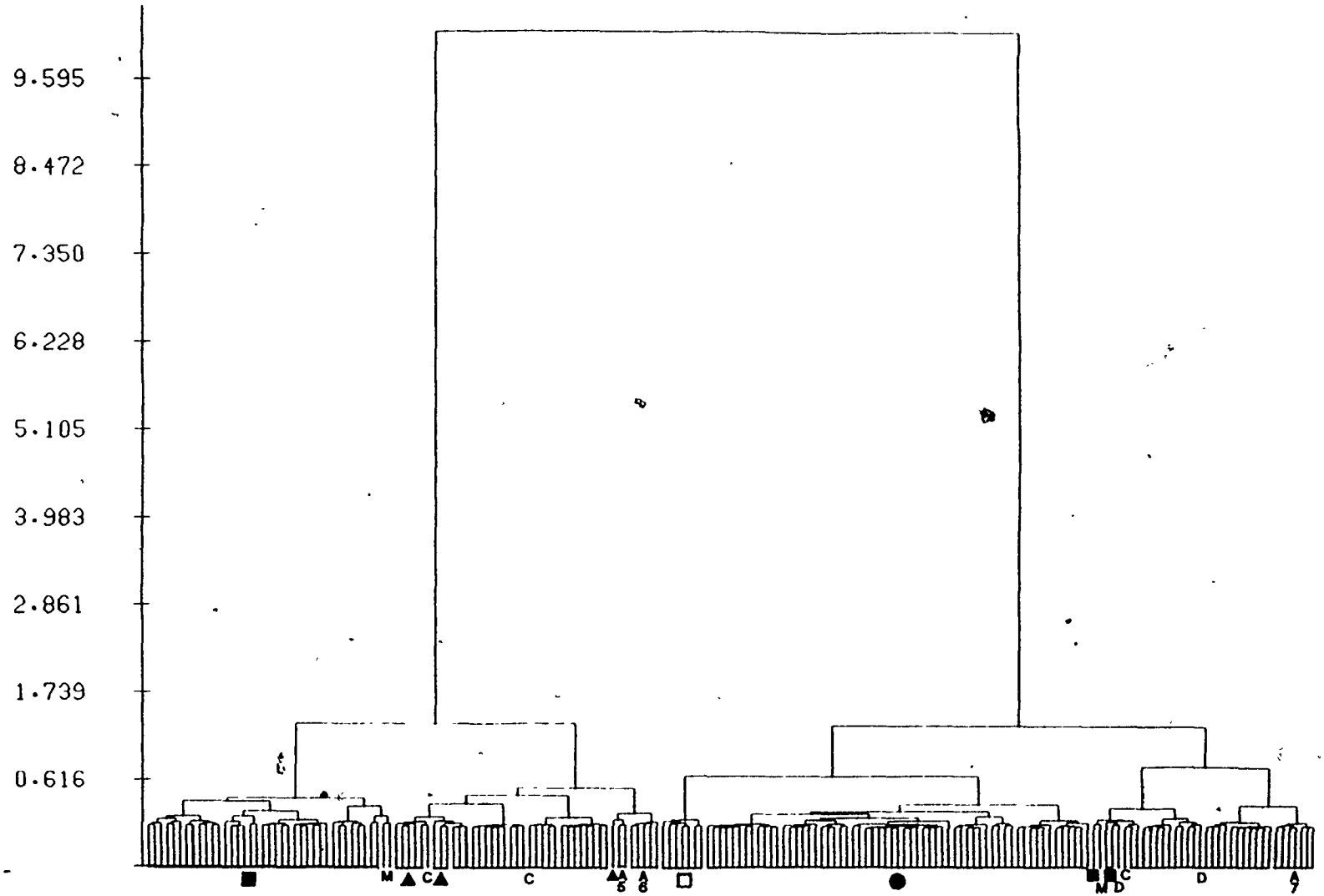

9.595
8.472
7.350
6.228
5.105
3.983
2.861
1.739
0.616
M
C
C
C
D
D

Figure 5.3 Phenogram resulting from Cluster Analysis of 185 OTUs (random plus exemplar collections) based on 29 ranged descriptors. The RM-CM combination used was Unweighted Pairs Grouped by Mathematical Averages of Euclidean Distances. Scale at left represents the resemblance level (Euclidean Distance) at which groups of OTUs fused.

Symbols:
Group 1 - closed squares
Group 2 - closed triangles
Group 3 - open squares
Group 4 - closed circles
C. chrysocarpa exemplars from sites 8, 9 and 10 - C
C. dodgei exemplars from sites 8 and 11 - D
C. margaretta exemplars from site 12 - M
C. aboriginum exemplars from site 5 - A5
C. aboriginum exemplars from site 6 - A6
C. aboriginum exemplars from site 7 - A7

Gustafsson, A. 1943. The genesis of the European blackberry flora. Acta Univ. Lund. n. s. 39(6):1-200.

Gustafsson, A. 1946-47. Apomixis in higher plants. Acta Univ. Lund. n. s. 42-43.

Harris, R. J. 1975. A Primer of Multivariate Statistics. Academic Press, New York.

Harris, R. J. 1976. The invalidity of partitioned-U tests in Canonical Correlation and Multivariate Analysis of Variance. Multivar. Behav. Res. 2:353-365.

Hauser, J. P. & J. H. Morrison. 1964. The cytochemical reduction of nitro blue tetrazolium as an index of pollen viability. Amer. J. Bot. 51:748-752.

Hiesey, W. M. & M. A. Nobs. 1982. Experimental Studies on the Nature of Species VI. Carnegie Inst. Washington Publ. No. 636.

Holmgren, P. K. & W. Keuken. 1974. Index Herbariorum. Oosthoek, Scheltema & Holkema, Utrecht, Netherlands.

Howell, T. J. 1898. A Flora of Northwest America. Portland, Ore.

Hutchinson, R. 1964. The Genera of Flowering Plants, Angiospermae. Clarendon Press, Oxford.

Jardine, N. & R. Sibson. 1971. Mathematical Taxonomy. John Wiley & Sons Ltd., London.

Knox, R. B. & J. Heslop-Harrison. 1963. Experimental control of aposporous apomixis in a grass of the Andropogoneae. Botaniska Notiser 116:127-141.

Kruschke, E. P. 1965. Contributions to the taxonomy of *Crataegus*. Milwaukee Public Mus. Publ. No. 3.

Longley, A. E. 1924. Cytological studies in the genus *Crataegus*. Amer. J. Bot. 11:295-317.

Loudon, J. C. 1838. Arboretum et Fruticetum Brittanicum. Longman, Orme, Brown, Green & Longmans, London.

Love, R. & M. Feigen. 1979. Interspecific hybridization between native and naturalized *Crataegus* in Oregon. Madrono 25:211-217.

Marie-Victorin, Frere. 1938. Phytogeographical problems of eastern Canada. Amer. Midl. Naturalist. 19:489-558.

Marshall, D. R. & A. D. H. Brown. 1981. The evolution of apomixis. Heredity 47:1-15.

Marshall, D. R. & B. S. Weir. 1979. Maintenance of genetic variation in apomictic plant populations. I. Single locus models. Heredity 42:159-172.

McMorris, F. R. & D. A. Neumann. 1983. Consensus n-trees. Bull. Math. Biosci. 43:239-244.

Medikus, F. C. 1793. Geschichte der Botanik Unserer Zeiten. Schwan und Gotz, Mannheim.

Moffett, A. A. 1931. A preliminary account of chromosome behavior in the Pomoideae. J. Pomol. Hort. Sci. 9:100-110.

Muniyamma, M. & J. B. Phipps. 1979a. Cytological proof of apomixis in *Crataegus* L.. Amer. J. Bot. 66:149-155.

Muniyamma, M. & J. B. Phipps. 1979b. Meiosis and polyploidy in Ontario species of *Crataegus* in relation to their systematics. Canad. J. Genet. Cytol. 21:231-241.

Muniyamma, M. & J. B. Phipps. 1984a. Studies in *Crataegus*. X. A note on the occurrence of diplospory in *Crataegus dissona* Sarg. (Maloideae, Rosaceae). Canad. J. Genet. Cytol. 26:249-252.

Muniyamma, M. & J. B. Phipps. 1984b. Studies in *Crataegus*. XI. Further cytological evidence for the occurrence of apomixis in North American hawthorns. Canad. J. Bot. 62:2316-2324.

Neumann, D. A. 1983. Faithful consensus methods for n-trees. Math Biosci. 63:271-287.

Nygren, A. 1967. Apomixis in Angiosperms. pages 551-596 in W. Ruhland [ed.], Encyclopedia of Plant Physiology, Vol. 18. Springer Verlag, Berlin.

Orlóci, L. 1967. An agglomerative method for classification of plant communities. J. Ecology 55:193.

Orlóci, L. 1975. Measurement of redundancy in species collections. Vegetatio 31:65-67.

Orlóci, L. 1978. Multivariate Analysis in Vegetation Research, 2nd ed. W. Junk, The Hague.

Palmer, E. J. 1925. Synopsis of North American *Crataegi*. J. Arnold Arbor. 6:5-128.

Palmer, E. J. 1932. The *Crataegus* problem. J. Arn. Arbor. 13:342-362.

Palmer, E. J. 1946. *Crataegus* in northeastern and central United States and adjacent Canada. Brittonia 5:471-490.

Palmer, E. J. 1963a. *Crataegus*. pages 338-375 in H. A. Gleason & A. Cronquist, The New Britten and Brown Illustrated Flora of the Northeastern United States and Adjacent Canada, Vol. 2. Hafner Publishing Co., for the New York Botanical Garden, New York.

Palmer, E. J. 1963b. *Crataegus*. pages 389-393 in H. A. Gleason & A. Cronquist, Manual of the Vascular Plants of Northeastern United States and Adjacent Canada, D. Van Nostrand Co. Inc., Princeton, New Jersey.

Phipps, J. B. 1983. Biogeographic, taxonomic and cladistic relationships between east asiatic and North American *Crataegus*. Ann. Missouri Bot. Gard. 70:667-700.

Phipps, J. B. & M. Muniyamma. 1980. A taxonomic revision of *Crataegus* (Rosaceae) in Ontario. Canad. J. Bot. 58:1621-1699.

Pimental, R. 1979. Morphometrics - the Multivariate Analysis of Biological Data. Kendall Hunt, Dubuque, Iowa.

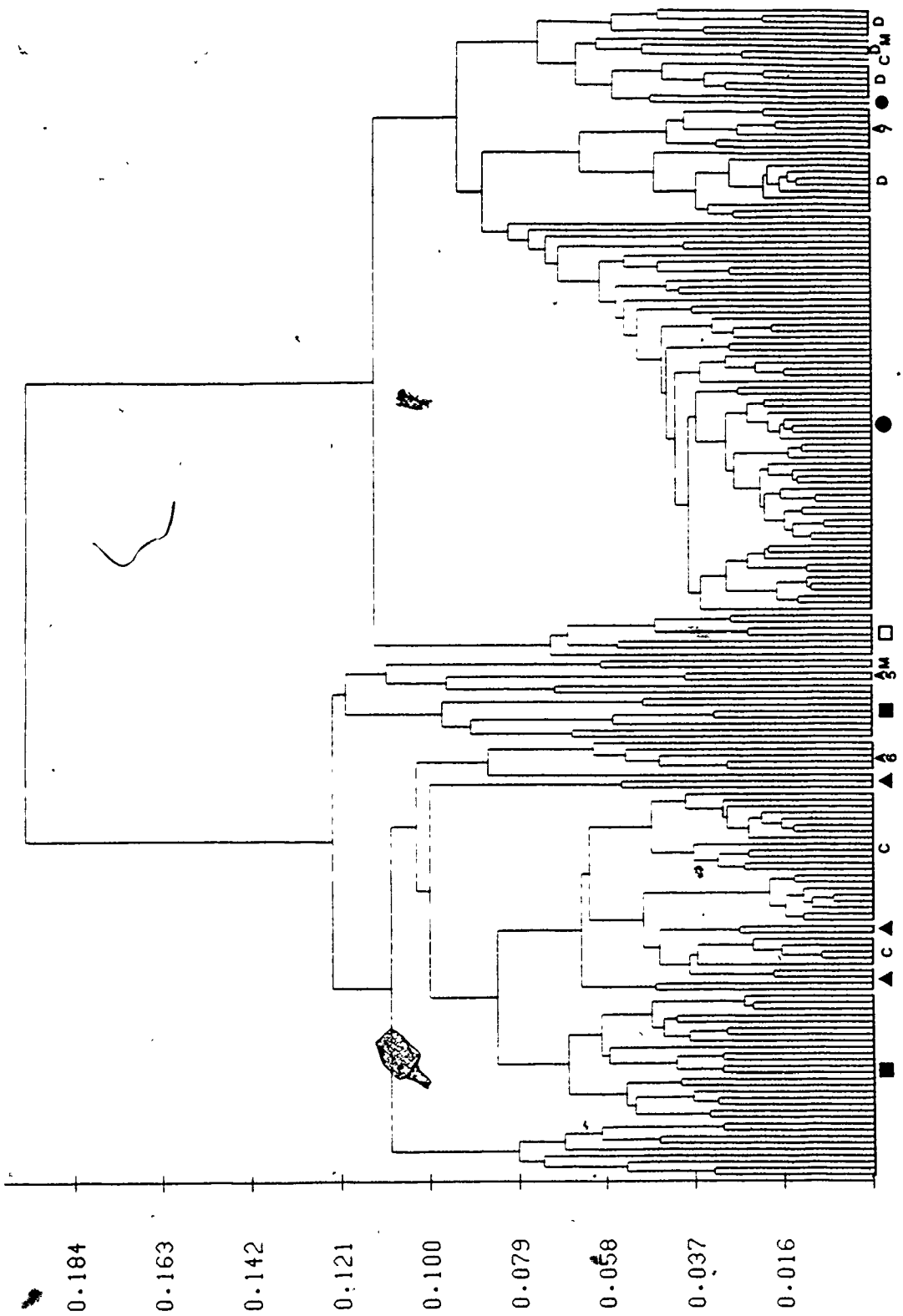

0.184
0.163
0.142
0.121
0.100
0.079
0.058
0.037
0.016

Figure 5.4 Phenogram resulting from Cluster Analysis of 185 OTUs (random plus exemplar collections) based on 29 ranged descriptors. The RM-CM combination used was Single Linkage of Euclidean Distances. Scale at left represents the resemblance level (Euclidean Distance) at which groups of OTUs fused.

Symbols:
Group 1 - closed squares
Group 2 - closed triangles
Group 3 - open squares
Group 4 - closed circles
C. chrysocarpa exemplars from sites 8, 9 and 10 - C
C. dodgei exemplars from sites 8 and 11 - D
C. margaretta exemplars from site 12 - M
C. aboriginum exemplars from site 5 - A5
C. aboriginum exemplars from site 6 - A6
C. aboriginum exemplars from site 7 - A7

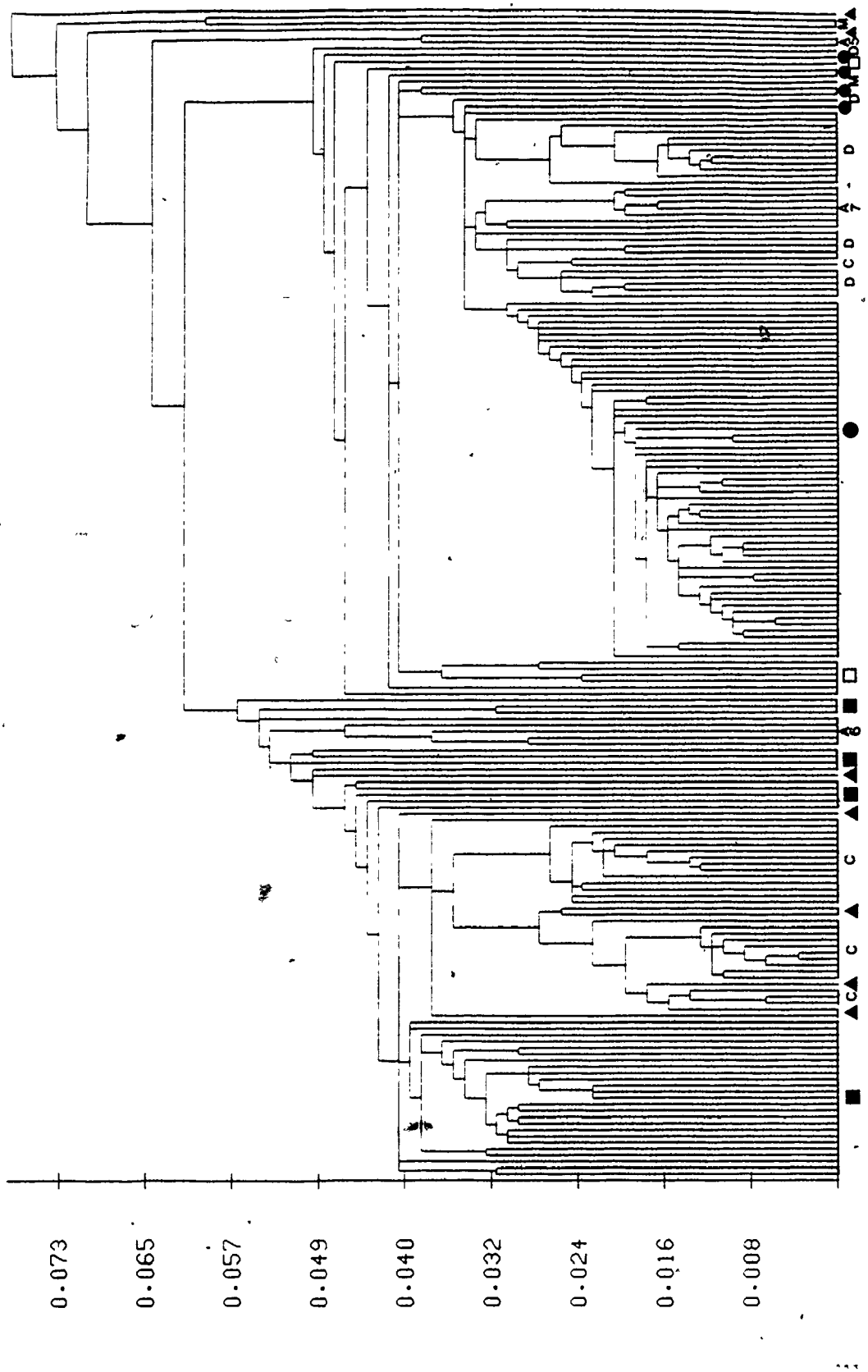

0.073
0.065
0.057
0.049
0.040
0.032
0.024
0.016
0.008

Figure 5.5 Phenogram resulting from Cluster Analysis of 185 OTUs (random plus exemplar collections) based on 29 ranged descriptors. The RM-CM combination used was Unweighted Pairs Grouped by Mathematical Averages of Centered Cosines. Scale at left represents the resemblance level (Centered Cosines) at which groups of OTUs fused.

Symbols:
Group 1 - closed squares
Group 2 - closed triangles
Group 3 - open squares
Group 4 - closed circles
C. chrysocarpa exemplars from sites 8, 9 and 10 - C
C. dodgei exemplars from sites 8 and 11 - D
C. margaretta exemplars from site 12 - M
C. aboriginum exemplars from site 5 - A5
C. aboriginum exemplars from site 6 - A6
C. aboriginum exemplars from site 7 - A7

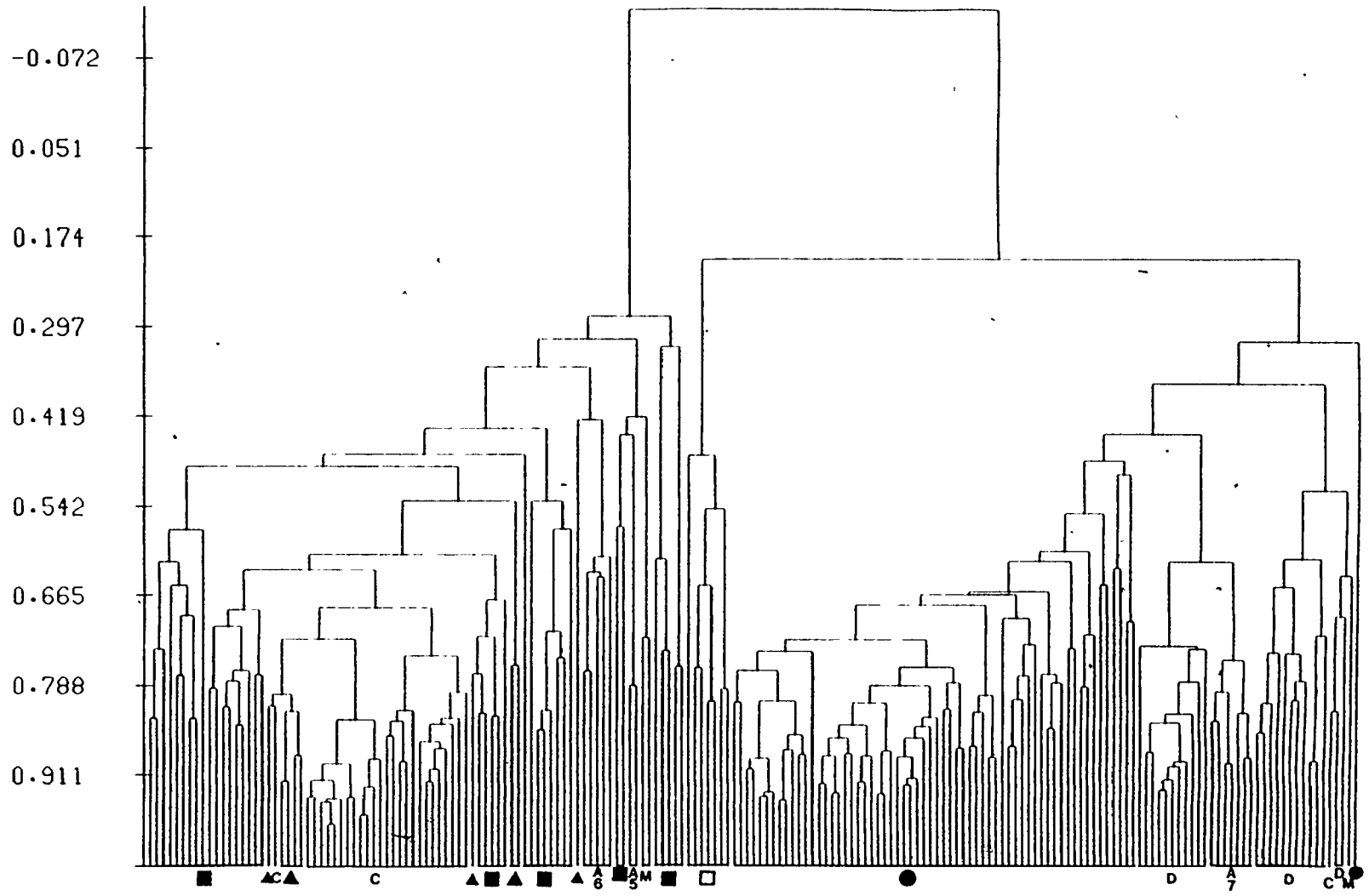

-0.072
0.051
0.174
0.297
0.419
0.542
0.665
0.788
0.911
C
A
6
A
5
M
D
A
7
D
C
D
M

Figure 5.6 Durchschnitt consensus tree, illustrating the information about group structure shared between the four previous phenograms. Scale at left represents the fusion level at which consensus clusters were formed (see text and Smith and Phipps 1984 for details). The actual sequence of OTU numbers in this figure is given in Appendix 4.

Symbols:
Group 1 - closed squares
Group 2 - closed triangles
Group 3 - open squares
Group 4 - closed circles
C. chrysocarpa exemplars from sites 8, 9 and 10 - C
C. dodgei exemplars from sites 8 and 11 - D
C. margaretta exemplars from site 12 - M
C. aboriginum exemplars from site 5 - A5
C. aboriginum exemplars from site 6 - A6
C. aboriginum exemplars from site 7 - A7

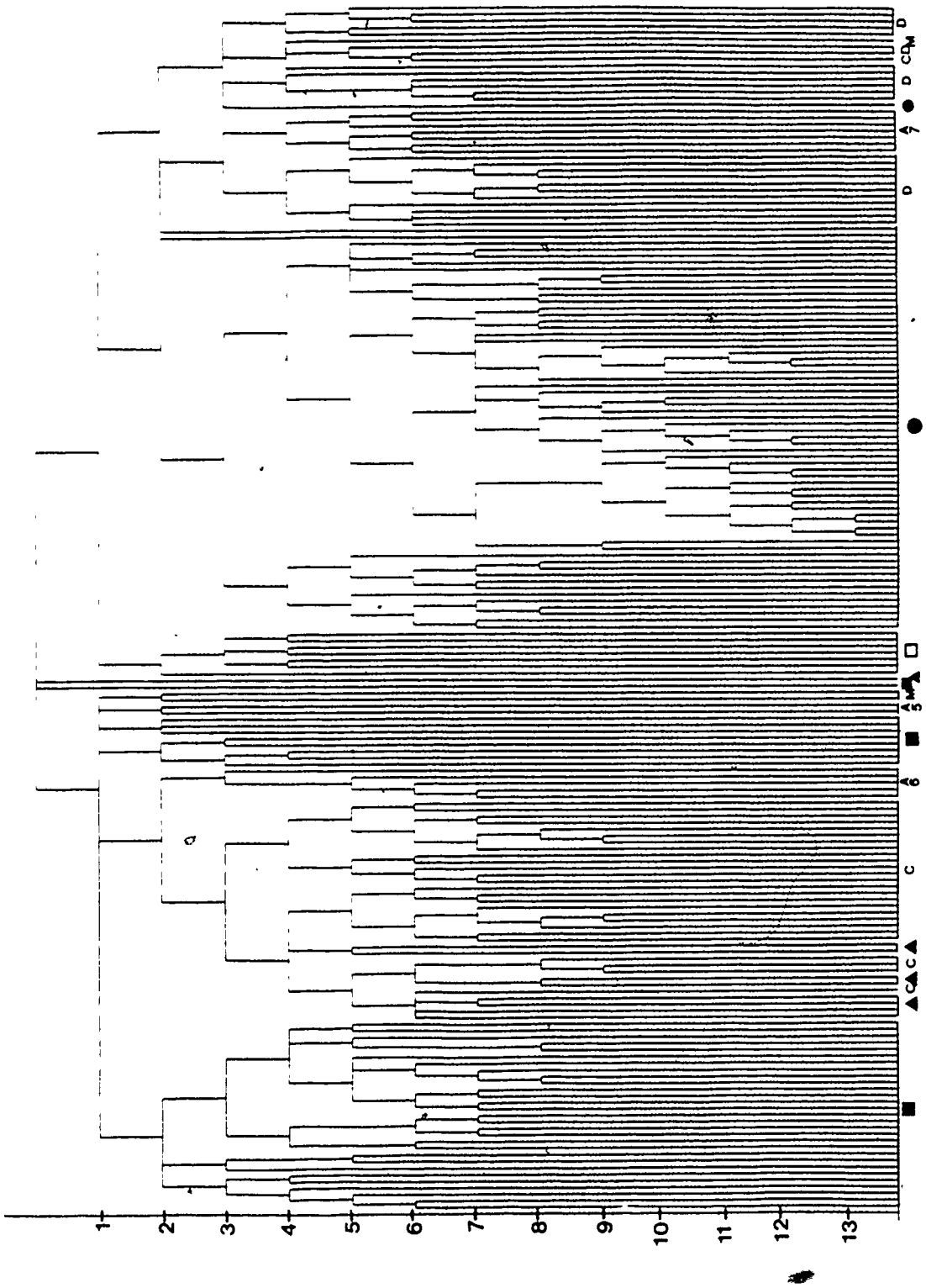
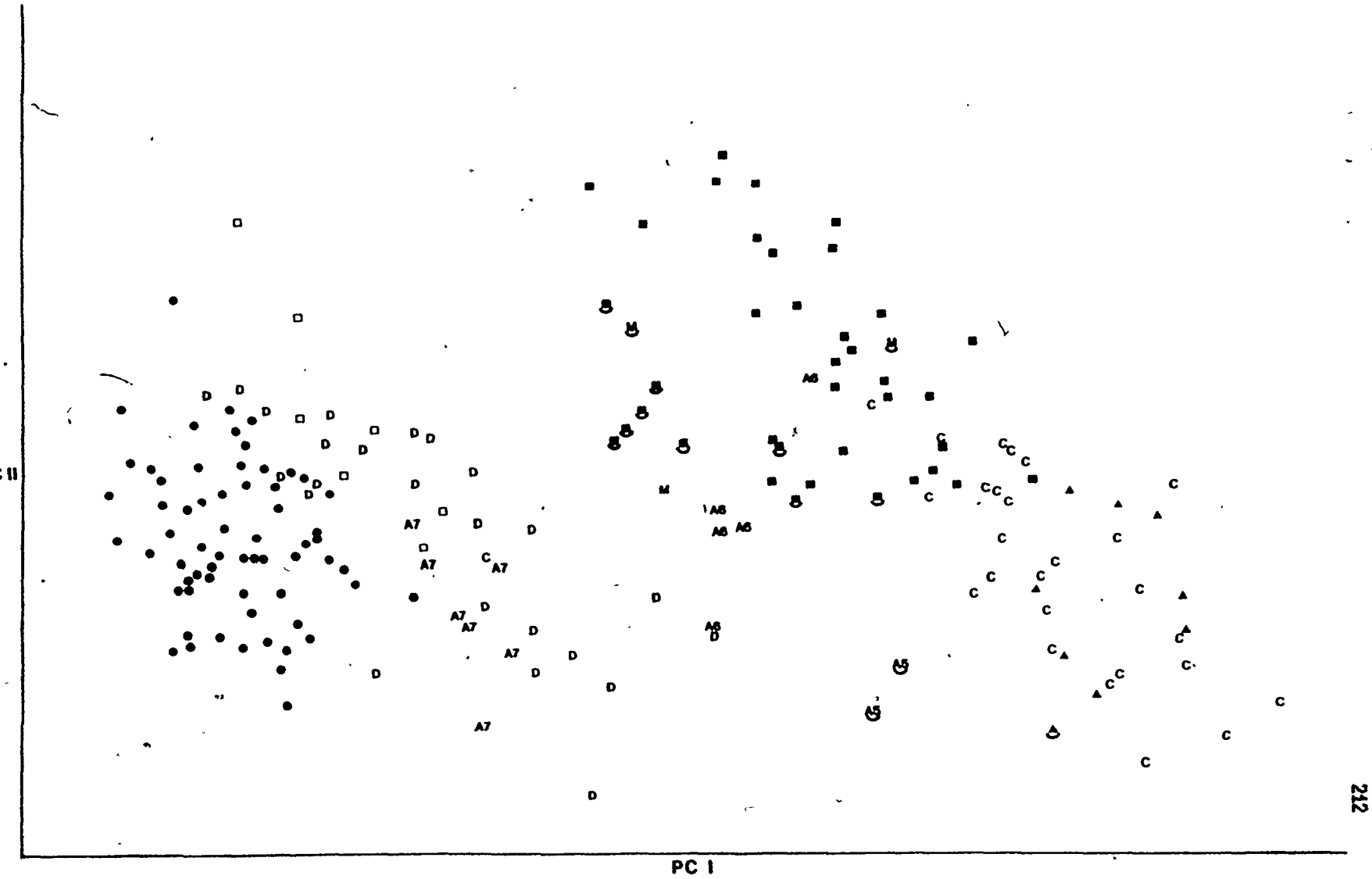

1
2
3
4
5
6
7
8
9
10
11
12
13
a
c
D
D
D

Figure 5.7 Scatter diagram of the position of each individual of the random plus exemplar collections (185 OTUs) in the space of the first two Principal Components (PC I and II) of a Principal Components Analysis of the correlation matrix (29 ranged descriptors). Semi-circles under symbols indicate outliers (see text for details).

Symbols:
Group 1 - closed squares
Group 2 - closed triangles
Group 3 - open squares
Group 4 - closed circles
C. chrysocarpa exemplars from sites 8, 9 and 10 - C
C. dodgei exemplars from sites 8 and 11 - D
C. margaretta exemplars from site 12 - M
C. aboriginum exemplars from site 5 - A5
C. aboriginum exemplars from site 6 - A6
C. aboriginum exemplars from site 7 - A7


PC II
PC I

CHAPTER 6

GENERAL DISCUSSION

6.1 INTRODUCTION

In this chapter, the evidence presented in Chapters 3, 4 and 5 will be brought together and discussed. First, a short summary of results will be presented. Next, these results will be synthesized and used to defend the thesis. Specifically, it will be shown that individuals of *Crataegus* series *Rotundifoliae* in Ontario are members of an agamic complex (first statement of the thesis) and that the patterns of variation exhibited by the individuals can be best understood in the light of their membership in the complex (second statement of the thesis). Following this, the results will be examined with reference to current knowledge of agamic complexes, regarding their structure and evolution. In this section, plausible hypotheses as to the origin and distribution of the entities studied here, will be presented. Suggestions for future research on series *Rotundifoliae* and on *Crataegus* will also be made.

6.2 SUMMARY OF RESULTS

In Chapter 3, patterns of morphometric variation were examined at the level of the whole random sample, at the level of major groups within the random sample and at the level of populations within the major groups. Results indicated that very strong group structure exists in the random sample. The effect of this group structure was to cause the descriptors most important in differentiating the major groups to predominate in the Analysis of Dispersion Structure, in Principal Components Analysis and in Cluster Analysis. Of the four major groups that were identified, three consisted of individuals (OTUs) located at more than one collection site, while the fourth group was made up of individuals from the same site.

Within the major groups, differing levels of univariate (standard deviations) and multivariate (dispersion determinants) descriptor variability were revealed. As well, different patterns of descriptor covariance and correlation were revealed in the Analysis of Dispersion Structure and in the PCA. Groups 1, 2 and 3 showed fairly high to extremely high levels of Dispersion Structure (many high multiple correlation coefficients) while Group 4 showed fairly low levels. In PCA, Groups 1, 3 and 4 displayed similar patterns of variation, with descriptors PELEN (petiole length), LFLEN (leaf length) and R3412 (width at 3/4 length / width at 1/2 length) loading heavily on, or showing high correlation with, the first two Principal Components. In Group 2, PTLEN (petal length) and PELEN loaded heavily on the first two Principal Components.

At the level of local populations, results were similar to those at the level of major groups, in that similar differences in amounts of univariate and multivariate variability were displayed. In PCA, however, very different patterns of descriptor loading were displayed for all populations.

In Chapter 4, results indicated that all four groups are probably mostly pseudogamous, facultative, aposporous agamosperms. All are apparently self-compatible facultative outcrossers with high pollen viabilities. Polyploid (predominantly tetraploid, with a few triploid) chromosome numbers were discovered in all individuals for which counts could be made. However, Group 4 did display some differences in chromosome number (the only firm triploid counts were found in this group) and in slightly lower pollen/ovule ratios. With respect to flowering phenology, Group 2 displayed significant differences in peak flowering from Groups 1 and 3, but some overlap in flowering was observed among all three of these groups.

In Chapter 5, it was revealed that members of Groups 3 and 4 (with one exception) belonged to the same local species, *C. flavida*. Individuals in Group 2 were identified as belonging to *C. chrysocarpa*, although one outlier of this group was determined to have closer affinity with *C. aboriginum*. Members of Group 1 were determined to be intermediate in morphology between *C. chrysocarpa* and *C. dodgei*. Many of the individuals in this group came within the type description of *C. lumaria*, although others did not. The group is referred to here as *C. ?lumaria*. The possibility that the individuals in this group represent the result of hybridization between *C. chrysocarpa* and *C. dodgei* was also raised.

6.3 DEFENCE OF THE THESIS

6.3.1 The Agamic Complex

Babcock and Stebbins (1938, p.55) formally define a heteroploid complex as

> "a group of species, containing forms with different chromosome numbers, of which those with the lowest number (ie. the diploids) are more or less distinct from one another morphologically, and are usually isolated from one another by sterility barriers, but in which some of the aneuploid or polyploid types are intermediate between the diploids or show different recombinations of their characteristics"

They further define (p. 58) agamic heteroploid complexes (more simply, agamic complexes) as

> "...those [heteroploid complexes] in which the normal sexual reproduction of the polyploid derivatives (or rarely some diploid forms...) is partially or wholly replaced by some sort of asexual propagation for dispersal... Reproduction by rhizomes or stolons is not included in this category, since the clone is not naturally dispersed over a wide geographical range by these methods."

Thus, an agamic complex consists of diploid, usually sexual progenitor species that are reproductively isolated from each other and polyploid derivative species that show agamospermous reproduction. Hybridization between derivative species and sexual progenitors and within derivatives is also of relatively frequent occurrence in agamic complexes (Grant 1981).

The results presented in Chapter 4 establish that the individuals studied are part of an agamic complex. Evidence for reproduction via apospory, of polyploidy and of the potential for hybridization through overlap in flowering times all serve to justify this conclusion. With the exception of sexual diploidy, all of the other elements defining an agamic complex are present. Sexual diploidy is presumed to be present elsewhere in the series or to have existed previously.

The size of the agamic complex, in terms of the number of species and their geographical range, cannot be inferred from the evidence at hand. However, it is certain that the complex extends beyond the taxa studied here, given the absence of diploid progenitor species in the study. Moreover, it is not likely that the diploid progenitors are extinct, given the diploid counts for the series (Longley 1924, Muniyamma and Phipps 1979b) discussed in Chapter 4. As well, it has been suggested (Phipps and Muniyamma 1980 and pers. comm.) that interseries hybrids exist that connect series *Rotundifoliae* with series *Tenuifoliae* and (possibly) with series *Macracanthae*. By extension, then, the complex may include, through similar interseries hybrids, all of what Phipps (1984, based on El-Gazzar's subgenus *Americanae*) refers to as section '*Americanae*' (most of the northeastern North American species) or even a majority of the North American *Crataegi*. Only a small number of very distinct North American *Crataegi*, such as *C. phaenopyrum* (L. f.) Medic. appear to have no putative natural interseries hybrids.

6.3.2 Patterns of Morphometric Variation

That complex patterns of morphometric variation are intrinsically associated with agamic complexes is virtually a rule in botanical systematics. To this extent, proof of the second statement of the thesis is implied by proof that the individuals under study are part of an agamic complex. However, as stated in Chapter 1, the important point here is to show how the particular patterns of variation exhibited by members of series *Rotundifoliae* can be biologically interpreted and fitted into a practical classification scheme. In other words, while the second statement of the thesis may be generally true, it is the purpose of this work to show how it is specifically true.

Before proceeding with a specific analysis of how agamospermous reproduction has affected morphometric variation in the series, it is worthwhile to examine what general expectations exist with respect to morphometric variation in an agamospermous group.

The essential difference between sexual and agamospermous reproduction is that agamospermy precludes recombination. Thus, offspring produced apomictically belong to the same genetic clone. In the simplest case, then, of an obligately apomict population arising from a single parent, morphometric variation would be expected to be extremely low. What variation is present would be perceived to be caused by differing environmental effects, chance mutation of the apomictic clone or generation of closely similar apomictic clones by further hybrid events between different phenotypes of the progenitor species. Similarly low levels of

variation in a sexual population would not be expected unless the sexual population were strictly autogamous.

Gustafsson (1946-7) has noted the possibility of what he refers to as autosegregation (recombination or chromosome aberration during meiosis in the megaspore mother cell) in diplosporous agamosperms. While autosegregation may generate some low level of variability in the few plant groups where it exists, its significance is difficult to assess because its frequency is not known. Clearly, mutation may also contribute to variability in apomictic populations, although the effect may be reduced in polyploids, due to the buffering effect of multiple genomes.

However, the level of variability exhibited by an agamospermous population is also crucially dependent upon factors of dispersal. If the founders of such a population were derived from two or more genetically different stocks (different biotypes), then variability in the population would be a function of the differences between biotypes. Thus, variability exhibited by agamospermous populations is at least partially a function of the variability contained in other populations of the same species that are within dispersal distance. This factor becomes important in explaining geographical distributions of members of agamic complexes, and their variability in various parts of their range.

Obviously, variability in an agamospermous population also depends upon the degree to which agamospermy is obligate. Facultative agamosperms, by virtue of the presence of at least some sexual reproduction (and thus recombination) would be expected to show greater levels of variability than

obligate agamosperms. In fact, facultative agamosperms appear to have the potential to generate vast amounts of variability through interspecific hybridization (sexual reproduction) followed by stabilization of the hybrid derivatives through apomixis. Questions remain, however, about the extent to which this potential is realized. This line of thought will be developed in a subsequent section.

With two or more agamospermous populations of the same species, a further expectation arises. Each population would be expected to show low variability and, again depending upon dispersal factors, each could be expected to show minute (or possibly larger) but consistent morphological differences from the others. Two sexual populations of the same species would not normally be expected to display such differences except where founder effects occur.

From the preceeding discussion, expectations with respect to agamospermous populations can be more specifically expressed in terms of the measures of morphometric variability used in the analyses of Chapter 3. In a population of obligate agamosperms, arising from one biotype, one would expect low descriptor variances, relatively little covariance or correlation and a low dispersion determinant. The amount of correlation displayed would, of course, also depend upon general factors such as allometric correlations (eg. length and width correlations arising from general growth patterns) as well as what could be termed taxonomic factors (correlations unique to a taxon that arise as a result of linkage between adaptive and non-adaptive characters). In a population of obligate agamosperms arising from different biotypes, variance and correlation would

be increased, specifically in those descriptors which differentiate the biotypes. In a facultative agamosperm, variances would also be expected to be higher but correlations, in the case of different founder biotypes, would be expected to be decreased, both as a result of recombination.

The strong group structure displayed in clustering at the level of the random sample may, therefore, be interpreted in the following way. The relative uniformity of the groups, especially in *C. flavida* (Groups 3 and 4), is most readily explained by an effect of facultative aposporous reproduction. In a cluster analysis of several aposporous species, apospory would serve to restrict within-groups variability and so emphasize between groups differences and thereby yield strong group structure. However, strong group structure also depends upon the number of biotypes present and their degree of differentiation. If a large number of slightly different biotypes occur in a sample, a cline (rather than strong grouping) may be revealed when the sample is subjected to clustering. It is suggested that the biotypes present in this study are few and reasonably well differentiated. The heavy loading of the pubescence descriptors in ADS and PCA is probably a combined effect of their large individual variances, correlations between them, and their importance in discriminating groups.

The separation of the two *C. flavida* groups (Groups 3 and 4) in cluster analysis can also be explained by reproductive factors. As previously mentioned, two agamospermous populations may or may not display morphometric differences between populations, depending upon dispersal patterns (how many biotypes were dispersed into the site during colonization). Here, the small but consistent differences between Group 3

(sites 3 and 4, north of London) and Group 4 (sites 1 and 2, south of London) are interpreted to be a result of small but consistent differences between colonizers of sites 1 and 2 versus those at sites 3 and 4. These differences are thought to have been perpetuated by aposporous reproduction. The lack of differences between the two populations making up Group 1 (*C. ?lumaria*) may be the reverse effect. It is suggested that the founders of these two populations were from the same stock. The origin of differences between the populations will be discussed in a subsequent section.

The differing levels of variability, in both descriptor variances and in dispersion determinants, between the groups can be interpreted as a result of several factors, discussed below. First, however, the levels of variability displayed in this study of apparently facultative aposporous species can be contrasted with levels of variability displayed by the sexual diploid *C. punctata* and with levels of variability displayed by the facultative apopsorous *C. crus-galli s.l.* Both of these taxa were investigated by Dickinson (1983) and Dickinson and Phipps (1985 in press). Although the suite of descriptors used in that study was somewhat different from the one used here, the methodology was similar enough to provide a useful comparison. Specifically of interest are the dispersion determinants. In Dickinson's analysis, a number of different OTU sets and desciptor sets were used, and a number of different dispersion determinants were calculated. The combination of OTU and descriptor sets in Dickinson's study that is most comparable to that used in the analyses of Chapter 3 is a set of four continuous flower and fruit descriptors assessed for 111 OTUs from several *C. crus-galli* populations and one *C. punctata* population. The *C. punctata*

population displayed a dispersion determinant of 2.86 x 10^{-10}. The *C. crus-galli* populations displayed determinants ranging from .001 x 10^{-10} to .149 x 10^{-10}. As is apparent from Table 3.34, the *C. crus-galli* determinants fall within the range of those displayed by *C. flavida*. The *C. punctata* determinant is slighlty larger than that of *C. ?lumaria* (population 1).

Given that all three species in the present study showed more or less similar levels of production of aposporous embryo sacs it might be expected that they would show generally similar levels of variability within groups. This, however, was not the case. One factor in contributing to differences in variability between groups has to do with variability in *C. ?lumaria*. In order to explain the high variances and large dispersion determinant in this taxon, hybridization is invoked. It was suggested in Chapter 5 that the morphological intermediacy of this group indicated the possibility that the group was a product of introgressive hybridization between *C. chrysocarpa* and *C. dodgei*. While no hard evidence for this exists (e.g. through direct crossing experiments or through flavonoid analysis), indirect evidence, such as morphological intermediacy (Chapter 5) and the overlap in flowering times (Chapter 4), does lend some weight to the hypothesis. As well, the occurrence of *C. ?lumaria* is in the region of overlap of ranges of the two putative parents. If this group is a product of hybridization, it is perceived that the individuals studied here represent F_2 or subsequent generations. In these generations, variability generated by sexual F_1's is being fixed, via apospory, into new aposporous lines. Gustafsson (1943) and Clausen et al (1949) give examples of similar patterns of segregation. Current knowledge of agamic complexes (Grant 1981) suggests that it is exactly such

hybridization events that contribute to the formation of agamic complexes.

The extremely low dispersion determinant displayed by Group 3 has previously been explained as, at least partially, an artifact of small sample size. Presumably, had there been more individuals in this group, the dispersion determinant would have been larger. Certainly the univariate standard deviations (except PELEN, table 3.14) indicate levels of variability similar to those displayed by the other *C. flavida* group.

The differences between *C. chrysocarpa* (Group 2) and *C. flavida* (Group 4) cannot be as readily explained. *C. chrysocarpa* is clearly more variable. However, it should be noted that *C. chrysocarpa* is far more widely distributed than *C. flavida*. It may be that the pool of variability from which the local *C. chrysocarpa* population was drawn is far more diverse than that of *C. flavida*. As well, even though aposporous embryo sacs are produced in both taxa, the relative numbers of progeny produced through sexual versus aposporous means is still unknown. Perhaps the *C. chrysocarpa* population is simply producing more progeny sexually.

The different patterns of descriptor covariance and correlation displayed at the group level are probably not so much a function of reproductive behavior as they are a function of genetic differences between taxa (eg. the taxonomic correlations discussed earlier). *C. ?lumaria* and *C. flavida* both showed similar strong influences, in ADS and PCA, by descriptors LFLEN, PELEN and R3412. This pattern was not evident in *C. chrysocarpa* because of the low variance of LFLEN and its low correlation with PELEN. With respect to overall dispersion structure, the low R^2_h values

displayed by Group 4 are a result of its overall uniformity (low descriptor variances).

The levels of variability displayed at the population level are explained in the same way as differences displayed at the group level. It is notable, however, that in both *C. ?lumaria* and *C. flavida* (Group 4) the dispersion determinants were reduced when these groups were split into populations. As well, in both groups the two population centroids were determined (via MANOVA) to be significantly different from each other. These results would not be expected if these groups were reproducing sexually.

The different patterns of PCA descriptor loading and R^2_h values displayed at the population level can be explained as a result of the reduction in correlation between LFLEN and PELEN when groups were split into populations. Reference to Figure 3.11 shows that both of these descriptors display somewhat separated mean values in population 1 versus 2 and in population 5 versus 6. Thus, the importance of these descriptors at the group level is decreased when the more homogeneous populations are formed and other descriptors assume importance.

6.4 THE STRUCTURE AND EVOLUTION OF AGAMIC COMPLEXES

6.4.1 Structure

Current understanding of the structure and evolution of agamic complexes has been derived primarily from extensive studies of very few genera, such as *Crepis* (Babcock and Stebbins 1938), *Rubus* (Gustafsson 1942,

1943, 1946-47), *Cotoneaster* (Sax 1954), and others. Stebbins (1950) and Grant (1981) provide excellent treatments of the phenomenon of agamic complexes in the context of modern biosystematic theory. All of the evidence thus far collected indicates the presence of a fairly common structure for these complexes. As indicated in the definition given previously, the classical model of agamic complexes dictates that they consist of reproductively isolated, sexual diploid progenitor species and a superstructure of polyploid apomictic derivative species. These derivative species represent the products of hybridization and segregation and/or polyploidy. It is characteristic of these derivative species that they display either the same morphological features as one of the diploids or a combination of features of two or more diploids. Hybridization between derivative species may serve to confound this pattern. In *Crepis*, the geographical distribution of the apomictic forms far surpasses the range of the diploids. This has been attributed to the production of new gene combinations in apomictic derivative species that are adaptively superior in new environments. As well, it appears that the number of different apomictic derivatives is far greater within the range of the diploids. This is seen as a result of the continuing process of throwing off new apomictic forms from the sexual diploids (Babcock and Stebbins 1938). Outside of the range of the diploids apparently few or no new forms are being generated. However, Stebbins (1950) suggests that in agamic complexes with facultative agamospermy, the range in apomictic forms is spread more evenly over the range of the complex.

The results presented in this thesis can now be interpreted in the light of the classical model of agamic complexes outlined above. Departures

from the classical model, displayed by the entities studied here (specifically *C. flavida*), will be noted.

In order to generate hypotheses about the origin of the entities studied here, some assumption must be made about the nature of those entities that originally gave rise to the complex. If the complex is to be seen as following the classical pattern, then one or more sexual diploid progenitor species must be postulated. Certainly, diploidy must have been present in the progenitor species in order to account for the presence of both diploidy and polyploidy in the extant species (unless widespread dihaploidization is invoked -see below). However, purely sexual diploid progenitors are not an essential prerequisite, given the widespread presence of apomixis in the *Rosaceae*. The potential for apomixis may well have been present in the ancestors of the entire genus. Suffice it to say that the progenitors of the agamic complex studied here were probably diploid and displayed some level of sexual reproduction. These will be referred to as diploid progenitors.

It seems likely that the diploid progenitors of this part of the complex have a form similar to *C. chrysocarpa* and *C. dodgei*, because these two species encompass the main range of morphological variability encountered in the series (with the exception of anther color and stamen number). Within *Crataegus*, anther colour and stamen number appear to be quite labile characters in that closely related species may differ only in one of these descriptors, while within a species, these two descriptors are not usually variable.

Given *chrysocarpa*-like and *dodgei*-like diploid progenitors, plausible hypotheses about the origin and distribution of the entities studied here can be erected. In general, other taxa within the series can be classed as derivatives of one or the other of these two species or as combination of the two.

It is suggested that extant *C. chrysocarpa* populations have arisen as a result of autopolyploidy in a *C. chrysocarpa*-like diploid progenitor. Auto-triploidy could have arisen as a result of fertilization of an unreduced female gametophyte by reduced pollen. An occurrence such as this is plausible if a propensity towards formation of unreduced gametes (e.g. a start in the direction of apomixis) is stipulated. Once triploidy is attained, tetraploidy can also arise via similar fertilization of unreduced (here triploid) gametes. A direct route to tetraploidy is also possible, via fusion of two unreduced gametes. It appears that *C. chrysocarpa* has stabilized at the tetraploid level, given the strictly tetraploid counts found here, and by Moffett (1931). Stabilization at the tetraploid level may also have allowed for the re-establishment of more normal meiotic division (diploidization) and hence, some level of sexuality in the tetraploids. It is interesting to note that the absence of any 5x or higher polyploid counts for this species and for the genus as a whole implies that 3x or 4x (unreduced) ESs are never fertilized. This leads to an interesting line of thought regarding the intrinsic limits to higher polyploidization in this agamic group. The relatively large range of morphological variability displayed by this species in general (and specifically by the population studied here) is consonant with recognizable levels of sexual reproduction. The extremely wide

distribution of *C. chrysocarpa* (Colorado to New England) suggests a reasonably old origin for this species, which is emphasized by the south trending spurs at both the eastern and western end of its distribution. *C. chrysocarpa* appears to be one of the coldest-dwelling of the *Crataegi*.

One individual of *C. aboriginum* (PS175), determined to be tetraploid, was included in the random sample. Muniyamma and Phipps (1979b) have also found tetraploidy in *C. aboriginum*. Morphologically this species is very similar to *C. chrysocarpa*. Two possible explanations are postulated for its origin. Either, as Phipps (pers. comm.) suggests, this entity represents a tetraploid segregate of a *C. chrysocarpa* X *C. flabellata* (series *Teniufoliae*) interseries hybrid, or it is a result of an extreme recombinational variant arising from sexual reproduction within tetraploid *C. chrysocarpa*. Its distribution, mostly sympatric with the southern portion of the range of *C. chrysocarpa* east of the prairies, is consonant with either of the two suggested modes of origin. It is predicted that this entity is a strictly tetraploid facultative apomict.

An origin similar to that suggested for *C. chrysocarpa* is proposed for *C. dodgei*. In this species, however, diploid, triploid and tetraploid individuals have been found within the study area. Only one OTU of *C. dodgei* was included in the random sample and thus, levels of population variability were not studied. However, subjective assessment of morphological variation in the two populations of this species that were included in the exemplar collections indicates levels of variability similar to that displayed by *C. chrysocarpa*. The wide range of *C. dodgei* (Illinois to Connecticut) also indicates a reasonably long age for this species.

C. flavida presents the greatest difficulty in interpreting the complex studied here in terms of the classical model of agamic complexes. In morphology (extremely little variation) and distribution (restricted) it appears to be a local (southwestern Ontario and adjacent parts of Michigan) variant of *C. dodgei*. If only polyploid individuals were present in the species, then it could easily be interpreted as a polyploid derivative. However, diploid (as well as triploid and tetraploid) individuals have been found within the study area. A possible explanation of this situation is that this species has arisen in a manner similar to that described for *C. chrysocarpa* and *C. dodgei*. *C. flavida* would then be considered to be another progenitor species. However, this explanation is not consonant with the low levels of variability displayed by the two populations studied here. One would expect greater variability in populations of a species with diploid (and thus sexual, according to the classical model) individuals so close by. An alternate explanation is that *C. flavida* arose as a diploid apomictic variant from diploid *C. dodgei*, and that auto-triploids and -tetraploids have arisen subsequently. The two different biotypes of this entity (Group 4, sites 1 and 2, Group 3, sites 3 and 4) may then have arisen as the result of occasional sexual crosses in this predominantly apomictic species. A third possibility is that the diploid *C. flavida* are the result of dihaploidization (autonomous development of a reduced egg cell in a tetraploid). Dihaploidization has, as yet, only been found in apomictic *Gramineae* (De Wet and Harlan 1970, Savadin and Pernes 1982)

It has already been suggested that *C. ?lumaria* has arisen as a result of introgressive hybridization between *C. chrysocarpa* and *C. dodgei*. The

wide variability displayed by the two populations studied here has been interpreted as the result of the continuing process of segregation wherein new apomictic lines are being formed. Thus, a relatively recent origin for this entity is suggested. Both triploidy and tetraploidy were found in *C. ?lumaria* individuals. The original hybrid(s) could, then, have arisen either at the triploid (4x *C. chrysocarpa* X 2x *C. dodgei*) or tetraploid (both parents 4x) level. In the former case, backcrossing between unreduced megaspores of the hybrid and reduced pollen from diploid *C. dodgei* would explain the presence of the tetraploids. In the latter case, triploids could have arisen from crossing between reduced gametes of both the hybrid and *C. dodgei*.

C. margaretta was largely unstudied in this work. The three OTUs of this species that were included in the exemplar collections showed wide morphological variability, in that two of the OTUs clustered separately from the third. However, state and local floras within the distributional boundaries of *C. margaretta* indicate that a number of local variants, morphologically similar to *C. margaretta*, exist. Moreover, this species appears to have very small anthers, which has been associated (Muniyamma and Phipps 1979b, Dickinson 1983) with extremely low pollen viability and thus with obligate apomixis. It may be that *C. margaretta s.l.* consists of a number of different obligate apomictic variants, two of which were included in the exemplar collections.

6.4.2 Evolution of Agamic Complexes

Gametophytic apomixis arises as a result of at least two mutations (Asker 1979). The first of these involves the circumvention of meiosis in

the megaspore mother cell, and development of an unreduced embryo sac. The second allows for autonomous development (commonly via parthenogenesis) of a functional embryo. Marshall and Brown (1981) have argued that the relative scarcity of agamospermy in the higher plants may be a result of the requirement for two mutations, either one of which, alone, can cause severe selective disadvantage. The genetic control of apomixis varies between plant groups but, in general, involves complex inheritance (Grant 1981). There appears to be a strong association between strict outcrossing and the development of apomictic behavior, in that the vast majority of the sexual relatives of apomicts are self-sterile (Gustafsson 1946-7). This may be (Charlesworth 1980) a result of a reduced selective advantage for gametophytic apomixis in taxa already possessing self-fertility. Once apomixis is achieved, self-sterility barriers apparently break down. The association between hybridization, polyploidy and apomixis may also be explained in terms of the development of apomictic behavior. In sterile hybrids, no further selective disadvantage is accrued in a mutation involving the breakdown of meiosis in the female.

Once gametophytic apomixis arises in a plant group, the development of polyploidy, through fertilization of unreduced gametes, is further enhanced (Asker 1979). Moreover, it has been suggested (Darlington 1939, Stebbins 1950) that the primary advantage of apomixis is in allowing the (often sterile) products of hybridization to be perpetuated. Thus, selection against hybridization (and associated allopolyploidy) is circumvented, and the frequency of polyploids and hybrids increases. Later stabilization of these hybrid forms can then occur via selection.

The hypothesis that apomixis represents only an 'escape from sterility' (Darlington 1939) also bears on the question of the amount of variability generated within agamic complexes. Two schools of thought exist. The first (Darlington 1939, Stebbins 1950, De Wet and Stalker 1974) view apomixis as an evolutionary dead-end. They see the apomictic mode as so severely reducing recombination that variability is limited to the diploid progenitors. As these diploids become extinct (e.g. through competition with apomictic derivatives) the possibility for variation becomes less and less. In this scenario, the apparently relictual apomict *Houttuynia cordata* is seen as the last vestige of a once wide ranging agamic complex (Babcock and Stebbins 1938).

However, more recent evidence supports the alternate view that apomictic complexes can and do perpetuate reasonable levels of variability. Clausen (1954) proposed what is now referred to as the 'Model-T hypothesis', wherein a relatively low level of sexual reproduction in facultative apomicts is sufficient to generate new, highly adaptive forms. These forms can then then be mass-produced (analogous to the mass production of the highly popular Model-T Ford) via apomixis. Thus facultative apomicts are seen to have the best of both sexual and asexual reproduction, especially in colonizing situations, where a particularly fit genotype can be replicated by a single colonizing individual. Such a genotype would be broken up by recombination in sexual colonizers. Marshall and Weir (1979) have presented calculations which establish that variation can be maintained in apomictic populations even when extremely small amounts of sexual reproduction exist. As well, they show that the maintainance, through selection, of balanced

polymorphism is enhanced in facultative apomicts that practice autogamy or mixed selfing and outcrossing over those that practice strict outcrossing. This may be an important factor in the breakdown of self-sterility barriers in groups in which apomixis has arisen. Gustafsson (1942) and, more recently, Hiesey and Nobs (1982) have experimentally shown that alternation between apomixis and sexuality (arising in F_1's of facultative apomict X facultative apomict interspecific crosses) does occur in *Rubus* and *Panicum*.

Whether or not the species making up an agamic complex are less long lived on the evolutionary scale than their sexual counterparts remains a moot point. It does seem that, if the interpretation presented in the previous section is correct, significant morphometric variation in series *Rotundifoliae* is limited to the presumed diploid progenitors and their autopolyploid derivatives. However, within that limit, facultative apospory appears to be sufficient to provide for continuing production of new forms (e.g. *C. ?lumaria*). As well, the adaptability of *C. chrysocarpa* to a very cold northern climate is indicative of a strong adaptive potential in the series.

As discussed in Chapter 1, it has been suggested (Phipps 1983) that the ancestors of the '*Americanae*' section of the genus (including series *Rotundifoliae*) migrated from Asia at one or more times during the mid-Tertiary. Later Pleistocene glacations must certainly have had an enormous effect upon evolution in the section. Glaciations would have served to cause wide changes in species distributions and abundances, caused possibly vast amounts of extinction and have provided new possibilities for hybridization between different progenitor parents. As well, wide new areas were made

available for colonization during glacial retreat and thus, competition between both sexual and apomictic forms was presumably taking place. According to Gustafsson (in Grant 1981), the diploid progenitors of the agamic complex in European *Rubi* are restricted to the Mediterranean region as a result of glaciation and the competitive superiority of the apomictic derivatives in subsequent recolonization. Sinnott (1979) and Dickinson (1983) have suggested that the effect in *Crataegus* has been one of wide mixing of types alternating with stabilization during alternating glacial and interglacial periods. Stabilization of the section, in post-glacial times may well have been hampered by more recent agricultural practices of indigenous peoples and of European settlers (Marie-Victorin 1938, Ulf Hansen 1985).

6.5 SUMMARY AND FUTURE WORK

In sumary, it has been shown that most individuals of *C.* series *Rotundifoliae* in Ontario belong to one of five species (excluding *C. margaretta* and *C. irrasa*) which form part of an agamic heteroploid complex. Some of these species have restricted, local distributions (e.g. *C. flavida*) while others are more widely spread (*C chrysocarpa*). The portion of the complex which these species form may be connected (via interseries hybrids) to both series *Tenuifoliae* and series *Macracanthae*. The patterns of variation displayed by these species has been related to their origin within the complex, local distribution patterns and to the degree of commitment to an aposporous mode of reproduction.

It has also been shown that the species studied here all appear to be facultative apomicts. This is perhaps the most interesting result arising

from the study, in that it demonstrates the presence of a set of related but nontheless distinct species all participating in an agamic mode of reproduction without the strong population differentiation displayed in a previous study of the *C. crus-galli* complex. As well, it seems clear that the evolutionary potential of series *Rotundifoliae* is by no means restricted by apomictic reproduction, given evidence for the ongoing production of new forms via hybridization and segregation (*C. ?lumaria*) and the adaptibility of extant members of the series (e.g. cold adapted *C. chrysocarpa*).

In order to develop a more complete understanding of *Crataegus* in North America, systematic investigations must proceed both at the level of the entire genus and at the level of intensive local studies, such as the one presented here. With respect to the entire genus, taxonomic relationships within and between sections and series must be established in order to provide, among other things, a reasonable general taxonomic scheme upon which further investigations can be based. Specifically, the necessary cytological and histological surveys cannot be usefully executed until identification of study specimens can be performed with surety and consistency. Mid-century approximations of this taxonomic work already exist (Palmer 1946, 1963a, Kruschke 1965) for North America, while more critical revisions are in progress (Phipps, Vascular Flora S.E.U.S.). The next step should be a survey of the frequency of different ploidy levels extant in the more common species of the genus. Difficulties still exist in obtaining countable material in quantity (see Chapter 4 and Muniyamma and Phipps 1979b). However, a cytological survey on this scale appears to be the shortest

route to establishing general trends in reproductive behavior in the genus, given the association between polyploidy and agamospermy exhibited here and in other studies (Muniyamma and Phipps 1979b, Dickinson 1983). Furthermore, knowledge of the range of ploidy levels in a given series can aid in interpretation of results obtained from intensive local studies. For example, results obtained in the present study would have been interpretatively far richer had it been known if diploidy exists elsewhere (taxonomically and geographically) in the series.

In addition, knowledge of the broad cladistic lines in the genus (if it is believed that these can be reliably estimated) would be of great value in assessing the ability of agamic complexes to maintain variability and thus long-term evolutionary potential.

At the local level, a moderate number of further studies, similar to the one presented here, need to be performed in order to establish trends in local variability. Of specific interest is the question of whether or not the presence of apospory is associated with (and therefore can be reliably predicted from) low levels of morphometric variability. Certainly the results obtained thus far indicate the likelihood of such an association. Local studies are also invaluable in helping to validate taxonomic conclusions based on more regional investigations, as was the case here.

APPENDIX 1

SITE DESCRIPTIONS

A1.1 Site 1, St. Thomas

Elgin Co., Yarmouth Twp.

42° 44′ N. Lat., 81° 01′ W. Long.

Elevation 215 m. a.s.l.

This site is an abandoned pasture, covered, for the most part, in a thick stand of many individuals of several different *Crataegus* species. Individuals of *Malus* and *Populus* are scattered throughout the field and *Pinus* plantings are situated along one side. The site is on an east facing slope of about 5 ° and is bordered by pasture on the north and east sides, mature forest on the south and a steep (ca. 30°) slope leading to a small floodplain of Kettle Creek on the west side. Fox Sandy Loam, Eroded and Bottomland soils occur here. The presence of many large series *Rotundifoliae* individuals at this site indicates relative maturity. The site is reasonably well drained although small pockets of standing water were observed here, in early spring. One individual of *C. dodgei* (PS276) and 19 individuals of *C. flavida* were collected from this site.

A1.2 Site 2, Walker's Ponds

Middlesex Co., London Twp., City of London

42° 56′ N. Lat., 81° 14′ W. Long.

Elevation 275 m. a.s.l.

This site is located off Wellington Road, near the intersection of Southdale Rd., in the City of London Ontario. It is an abandoned field, partially bordering one of the Wesminster or Walker's Ponds. The remainder of the site is bordered by a mature forest stand. The site consists of a slightly elevated area surrounding a small central drainage basin. The soil is Huron Clay Loam. Most of the *Crataegus* individuals occurring at the site are located in the raised, more drained area, although some occur along the borders of the woodland. According to a local resident, the field has been abandoned for at least 30 years. Colonization by *Crataegus* appears to be originating from the eastern end, as the shrubs become less mature from east to west. 43 individuals of *C. flavida* were collected from this site.

A1.3 Site 3, McGillivray

Middlesex Co., Mcgillivray Twp.

43° 13′ N. Lat., 81° 30′ W. Long.

Elevation 260 m. a.s.l.

This site is situated north of County Road 24, about 200 m. east of the intersection of sideroad 10 and Concession Road 4. It covers a mixed area of floodplain of a tributary of the Ausble River and pasture edge. Evidence of considerable flooding was observed in the lower portions of the site, during early spring. Very large willows occur along the banks of the stream. The soil is Huron Clay Loam and Bottomland. Cattle have been observed throughout the site. Individuals of *C. chrysocarpa* (8 OTUs), *C. flavida* (2 OTUs) and *C. ?lumaria* (22 OTUs) were collected from this site.

A1.4 Site 4, Ausable-Bayfield

Middlesex Co., West Williams Twp.

43° 06′ N. Lat., 81° 49′ W. Long.

Elevation 230 m. a.s.l.

Located at the Ausable-Bayfield Conservation Authority Forest, north of Arkona, on the Middlesex-Lambton County line, this site has been widely planted with *Pinus*, which are now mature. *Crataegus* shrubs occur along the edges of fire roads and footpaths cut through the plantation. The site is situated atop a deep gorge of the Ausable River and is reasonably well drained. The soil type is Huron Clay Loam and Eroded soils. *C. flavida* (5 OTUs) and *C. ?lumaria* (14 OTUs) were found here.

Additional Sites

A1.5 Site 5, Elginfield

Middlesex Co., London Twp.

43° 10′ N. Lat., 81° 22′ W. Long.

Elevation 280 m. a.s.l.

Located approximately 500 m. south of the southeast corner of the intersection of Ontario Highways 4 and 7, this site contains several hundred *Crataegus* individuals of various species. These are located around the edges of a cornfield and along a short roadway leading to the field, off of Highway 4. Two individuals of *C. aboriginum* (PS181, PS182) were collected here.

A1.6 Site 6, Elma

Perth Co., Elma Twp.

43° 41′ N. Lat., 81° 01′ W. Long.

Elevation 380 m. a.s.l.

This site is located on the north side of Elma Township Concession Road 3-4, approximately 100 m. east of Highway 23. It is bounded on the east by railroad tracks, on the north and west by a small stream and on the south by Concession Road 3-4. The site consists of a thicket of two to three hundred *Crataegus* individuals surrounding a cultivated cornfield. Five individuals, all identified as *C. dodgei* were collected here.

A1.7 Site 7, Halton

Regional Municipality of Halton, Nassagawea Twp.

43° 30′ N. Lat., 80° 01′ W. Long.

Elevation 305 m. a.s.l.

Located at the Burns Conservation Area, north of the junction of Number 10 Sideroad and the First Line, northeast of Campbellville, Ontario, this site covers an area of about 2 hectares, surrounding a pond near the entrance to the conservation area. Collected individuals were located along a short roadway leading to the pond and in an abandoned pasture northeast of the pond. Seven individuals of *C. dodgei* were collected at this site.

A1.8 Site 8, R. B. G.

Regional Municipality of Hamilton - Wentworth,
East Flamborough Twp.

43° 18′ N. Lat., 79° 52′ W. Long.

Elevation 105 m. a.s.l.

This site is located on the grounds of the Royal Botanic Gardens in Hamilton, Ontario. *C.* series *Rotundifoliae* individuals were collected from an old field surrounding the junction of Pinetum Trail and Homestead Walk. Ten individuals of *C. dodgei* and five of *C. chrysocarpa* were collected at this site.

A1.9 Site 9, Carden

Victoria Co., Carden Twp.

44° 37′ N. Lat., 79° 04′ W. Long.

Elevation 245 m. a.s.l.

Located approximately 3 Km. north of Rohallion, Ontario, on the east side of highway 503, about 100 m. north of a sharp northward bend in the highway, the site is an apparently frequently used, sparsely vegetated cattle pasture. Individuals of *C. chrysocarpa* and *Rhus typhina* are scattered throughout the pasture. Eight individuals of *C. chrysocarpa* were collected at this site.

A1.10 Site 10, Manitoulin

Manitoulin Island,

Allen Twp., 45° 53′ N. Lat., 82° 22′ W. Long., Elevation 230 m. a.s.l.

Gordon Twp., 45° 51′ N. Lat., 82° 32′ W. Long., Elevation 220 m. a.s.l.

Howland Twp., 45° 55′ N. Lat. 81° 56′ W. Long., Elevation 215 m. a.s.l.

Sandfield Twp., 45° 47′ N. Lat. 82° 03′ W. Long., Elevation 235 m. a.s.l.

A number of different collection sites, all on Manitoulin Island, are grouped under this site number. All 14 individuals collected on the Island were identified as *C. chrysocarpa*. Nine Individuals were collected from Sandfield Township. Eight of these were found in an abandoned pasture located between Highway 542 and a short road leading to the Township Dump, about 4 Km. north of the Town of Sandfield. The other individual was collected from a waste area on the northeast corner of the junction of Highway 542 and Sandfield Concession Road 4, just north of Sandfield. Two individuals were collected in Howland Township, along Highway 68 (now Highway 6), in an abandoned pasture about 2 Km. north of Sheguiandah, where the old roadbed of Highway 68 rejoins Highway 6. Two other individuals were collected in Gordon Township, along the south side of Concession Road 6, 200 m. east of Highway 540. These individuals were located

along the edge of a well-grazed cattle and horse pasture. Finally, one individual was collected in a pasture on the east side of Ice Lake, in Allen Township, approximately 400 m. northwest of the intersection of Sideroad 20 and Concession Road 6.

A.1.11 Site 11, Gratiot

St. Clair Co., Michigan

42° 53′ N. Lat. 82° 32′ W. Long.

Elevation 200 m.

This site consists of a large abandoned pasture containing several hundred *Crataegus* individuals of various species. Thirteen of the 14 series *Rotundifoliae* individuals collected here were identified as *C. dodgei*. The other was identified as *C. chrysocarpa* var. *phoenicea*. The site is located on the north side of Gratiot Road, approximately 1 Km. west of the junction of Wadhams and Gratiot Roads, southwest of Port Huron, Michigan.

A1.12 Site 12, Morenci

Lenawee Co., Michigan

41° 43′ N. Lat. 84° 12′ W. Long.

Elevation 185 m.

Part of this site is located about 2 Km. west of the junction of Lime Creek Road and Rt. 156, 4.5 Km. north of Morenci, Michigan. Two individuals were collected along Lime Creek Rd., the first 1.6 Km west of Rt. 156, in a waste area along the roadway, opposite a golf course. No other *Crataegi* were found here. The second individual was collected 1 Km. further west on the northwest corner of Lime Creek and Bothwell Roads, along the fencline of a cultivated field. The third individual was located along the fenceline of an abandoned pasture, bordering Rt. 34, where is veers sharply north, 5 Km. west of its junction with Rt. 156. All three individuals were identified as *C. margaretta*.

APPENDIX 2

ACCESSION NUMBER, SITE OF COLLECTION AND SPECIES AFFILIATION FOR EACH INDIVIDUAL USED IN THE STUDY.

C. aboriginum

Site 3 PS175

Site 5 PS280, PS281

Site 6 PS339, PS340, PS343, PS344, PS345

C. chrysocarpa

Site 3 PS158, PS159, PS160, PS176, PS181, PS182, PS196, PS214

Site 8 PS146, PS147, PS148, PS149, PS156

Site 9 PS346, PS347, PS348, PS349, PS350

Site 10 PS284, PS285, PS287, PS288, PS289, PS290, PS291, PS292 (var. *phoenicea*), PS294, PS346, PS347, PS348, PS349, PS350

C. dodgei

Site 1 PS276

Site 7 PS315, PS316, PS319, PS321, PS322, PS323, PS324

Site 8 PS133, PS134, PS135, PS137, PS150, PS151, PS152, PS153, PS154, PS155

Site 11 PS299, PS300, PS301, PS302, PS305, PS306, PS307, PS308, PS309, PS310, PS311, PS312, PS313, PS314

C. flavida

Site 1 PS144, PS259, PS261, PS262, PS263, PS264, PS265, PS266, PS267, PS268, PS269, PS270, PS271, PS272, PS273, PS274, PS275, PS277, PS278

Site 2 PS108, PS111, PS113, PS218, PS219, PS220, PS221, PS222, PS223, PS224, PS25, PS226, PS227, PS228, PS229, PS230, PS231, PS232, PS233, PS235, PS236, PS237, PS238, PS239, PS240, PS241, PS242, PS243, PS244, PS245, PS246, PS247, PS248, PS249, PS250, PS251, PS252, PS253, PS254, PS255, PS256, PS257, PS258

Site 3 PS207, PS216

Site 4 PS104, PS141, PS168, PS190, PS192

C. ?lumaria

Site 3 PS115, PS119, PS172, PS174, PS177, PS178, PS179, PS195, PS197, PS198, PS200, PS201, PS203, PS204, PS205, PS206, PS207, PS208, PS209, PS210, PS211, PS215

Site 4 PS101, PS103, PS161, PS162, PS163, PS164, PS165, PS167, PS169, PS183, PS187, PS188, PS189, PS191

C. margarettd

Site 12 PS295, PS296, PS297

APPENDIX 3

SEQUENCE OF OTU/ACCESSION NUMBERS IN THE 112-OTU DURCHSCHNITT CONSENSUS TREE (Fig. 3.6).

(Left to right sequence of accession numbers in Figure 3.6 is given, from top to bottom, left to right. Spaces indicate differentiation of symbols in Figure 3.6. PS prefixes are omitted from accession numbers.)

101	183		249	263
119	164	104	223	229
115	165	216	239	246
103	209	207	245	252
201	215	192	242	247
161	179	141	240	227
162	206	190	267	221
167	210	168	251	230
174	187		259	219
169	189	218	256	233
170	163	235	257	273
172	208	248	253	269
204	211	272	228	265
200		241	274	254
202	158	271	266	255
203	159	275	237	226
197	160	243	232	258
198	214	244	278	225
177	176	224	261	270
178	196	238	220	276
195	181	250	264	231
205	182	222	268	262
191	175	236	277	

APPENDIX 4

SEQUENCE OF OTU/ACCESSION NUMBERS IN THE 185-OTU DURCHSCHNITT CONSENSUS TREE (Fig. 5.6).

(Left to right sequence of accession numbers in Figure 5.6 is given, from top to bottom, left to right. Spaces indicate differentiation of symbols in Figure 5.6. PS prefixes are omitted from accession numbers.)

101	346	339	258	256	314
119		343	266	218	312
115	176	344	265	235	
140	196	345	273	245	315
169		340	226	251	319
170	347		255	267	321
163	348	187	252	248	322
183	350	188	254	261	316
191		189	269	278	323
103	160	209	229	238	324
201	214	215	220	259	
161		165	223	242	276
162	325	208	222	272	
164	327	211	236	239	133
167	328		249	263	135
172	330	280	271	264	150
200	334	281	240	268	151
198	331		243	277	154
204	329	296	228	246	134
174	333	297	253	262	
205	349		274	108	292
202	146	210	241	111	
203	149		257	113	300
197	156	175	244	144	301
177	147		275	231	
178	148	104	219		295
179	284	141	221	302	
195	285	192	227	305	137
206	288	168	247	313	299
	287	190	230	308	152
158	291	207	233	306	153
159	294	216	224	307	155
181	289		232	311	
182	290	225	237	309	
		270	250	310	

REFERENCES

Adams, E. N. III. 1972. Consensus techniques and the comparison of taxonomic trees. Syst. Zool. 21:390-397.

Alexander, M. P. 1969. Differential staining of aborted and non-aborted pollen. Stain Technol. 44:122.

Asker, S. 1979. Progress in apomixis research. Hereditas 91:231-240.

Ashe, W. W. 1900. New Eastern American Species of *Crataegus*. J. Elisha Mitchell Sci. Soc. 16:70-79.

Ashe, W. W. 1901. Some new species of the genus *Crataegus* and notes on some dichotomous *panicums*. Bull. N. C. Exp. Sta. 175:109-116.

Ashe, W. W. 1903. New North American thorns. J. Elisha Mitchell Sci. Soc. 19:10-31.

Babcock, E. B. & G. L. Stebbins Jr. 1938. The American species of *Crepis*. Carnegie Inst. Washington Publ. No. 504.

Battaglia, E. 1963. Apomixis. in Recent advances in the embryology of Angiosperms. pp. 221-264.

Berlyn G. P. & J. P. Miksche. 1976. Botanical Microtechnique and

Cytochemistry. Iowa State University Press, Ames, Iowa.

Box, G. E. P. 1949. A general distribution theory for a class of likelihood criteria. Biometrika 36:317-346.

Brown, H. B. 1910. The genus *Crataegus* with some theories of the origin of species. Bull. Torrey Bot. Club 37:251-260.

Byatt, J. I. 1975. Hybirdization between *Crataegus monogyna* Jacq. and *C. laevigata* (Poiret) DC. in southeastern England. Watsonia 10:253-264.

Byatt, J. I. 1976. The structure of some *Crataegus* populations in northeastern France and southeastern Belgium. Watsonia 11:105-115.

Camp, W. H. 1942a. The *Crataegus* problem. Castanea 7:51-55.

Camp, W. H. 1942b. Ecological problems and species concepts in *Crataegus*. Ecology 23:368-369.

Chapman, A. W. 1897. Flora of the Southeastern United States, 3rd ed. New York.

Charlesworth, B. 1980. The cost of sex in relation to mating system. J. Theor. Biol. 84:655-671.

Cinovskis, R. E. 1971. *Crataegi Baltica*. Riga, Latvian SSR.

Clark, G. [ed.] 1973. Staining Procedures used by the Biological Staining Commission, 3rd ed. Williams & Wilkins Co., Baltimore, Maryland.

Clausen, J. 1954. Partial apomixis as an equilibrium system in evolution. Caryologia 6(Suppl.)469-479.

Clausen, J., Keck, D. D., Hiesey, W. M. & P. Grun. 1949. Experimental Taxonomy. Carnegie Inst. Washington Yearbook 48:95-106.

Cooley, W. W. & P. R. Lohnes. 1971. Multivariate Data Analysis. John Wiley & Sons, Inc., New York.

Cruden, R. W. 1977. Pollen-ovule ratios: a conservative indicator of breeding system in flowering plants. Evolution 31:32-46.

Cruden, R. W. & S. Miller-Ward. 1981. Pollen-ovule ratio, pollen size, and the ratio of stigmatic area to the pollen-bearing area of the pollinator: an hypothesis. Evolution 35:964-974.

Darlington, C. D. 1939. The Evolution of Genetic Systems. Cambridge University Press, Cambridge.

DeWet, J. M. J. & J. R. Harlan. 1970. Apomixis, polyploidy and speciation in *Dichanthium*. Evolution 24:270-277.

DeWet, J. M. J. & H. T. Stalker. 1974. Gametophytic apomixis and evolution in plants. Taxon 23:689-697.

Dickinson, T. A. 1983. *Crataegus crus-galli* L. *sensu lato* in Southern Ontario: Phenotypic Variation and Variability in Relation to Reproductive Behavior. Ph. D. Thesis. University of Western Ontario.

Dickinson, T. A. & J. B. Phipps. 1984. Studies in *Crataegus* IX. Short shoot leaf heteroblasty in *Crataegus crus-galli* s. l. Canad. J. Bot. 62:1775-1780.

Dickinson, T. A. & J. B. Phipps. 1985. Studies in *Crataegus* XIII. Variation and variability in *Crataegus* Section *crus-galli* in Ontario. Syst. Bot. in press.

Eggleston, W. W. 1908. The Crataegi of the northeastern United States and adjacent Canada. Rhodora 10:73-84.

El-Gazzar, A. 1980. The taxonomic significance of leaf morphology in *Crataegus* (Rosaceae). Bot. Jahrb. Syst. 101:457-469.

Estabrook, G. F., Winsor, J. A., Stephenson, A. G. & H. F. Howe. When are two phenological patterns different? Bot. Gaz. 143:374-378.

Gower, J. C. 1971. A general coefficient of similarity and some of its properties. Biometrics 27:857-874.

Grant, V. 1981. Plant Speciation, 2nd ed. Columbia University Press. New York.

Gray, A. 1867. A Manual of Botany of the Northern United States, 5th ed. Cambridge, Boston.

Green, P. E. 1978. Analyzing Multivariate Data. Dryden Press, Hinsdale, Ill.

Gustafsson, A. 1942. The origin and properties of the European blackberry flora. Hereditas 28:249-277.

Rao, C. R. 1952. Advanced Statistical Methods in Biometric Research. John Wiley & Sons Inc., New York.

Reheder, A., Palmer, E. J. & L. Croizat. 1938. Seven binomials proposed as nomina ambigua. J. Arnold Arbor. 19:282-290.

Rickett, H. W. 1936. Forms of *Crataegus prunosa*. Bot. Gaz. 97:788-793.

Rickett, H. W. 1937. Forms of *Crataegus crus-galli*. Bot. Gaz. 98:609-616.

Rohlf, F. J. 1977. A note on the measurement of redundancy. Vegetatio 34:63-64.

Rohlf, F. J. 1982. Consensus indices for comparing classifications. Math. Biosci. 59:131-144.

Ryan, T. A., Joiner, B. L. & R. F. Ryan. 1981. Minitab Reference Manual. Minitab Project, The Pennsylvania State University, University Park, Pa.

Sargent, C. S. 1901. *Crataegus* in the Champlain Valley. Rhodora 3:19-31.

Sargent, C. S. 1903a. Recently recognized species of *Crataegus*. Rhodora 5:52-66.

Sargent, C. S. 1903b. Recently recognized species of *Crataegus* in eastern Canada and New England. Rhodora 5:159-168.

Sargent, C. S. 1903c. *Crataegus* in Rochester, New York. Proc. Rochester Acad. Sci. 4:93-136.

Sargent, C. S. 1907a. The genus *Crataegus* in North America. J. Bot. 45:289-292.

Sargent, C. S. 1907b. Rep. Geol. Survey Michigan 1906:554.

Sargent, C. S. 1908. *Crataegus* in Ontario. Ontario Nat. Sci. Bull. 4:11-98.

Sargent, C. S. 1920. Notes on North American trees. X. J. Arnold Arbor. 3:196-197.

Savadin, Y. & J. Pernes. 1982. Diploid - tetraploid - dihaploid cycles and the evolution of *Panicum maximum* Jacq. Evolution 36:396-400.

Sax, H. J. 1954. Polyploidy and apomixis in *Cotoneaster*. J. Arnold Arbor. 35:334-365.

Sinnott, Q. P. 1978. A Taxonomic Study of series *Pruinosae* (*Crataegus*, Rosaceae). M. Sc. Thesis. University of Western Ontario.

Sinnott, Q. P. & J. B. Phipps. 1983. Variation patterns in *Crataegus* series *Pruinosae* (Rosaceae) in southern Ontario. Syst. Bot. 8:59-70.

Smith, P. G. & J. B. Phipps. 1984. Consensus trees in phenetic analyses. Taxon 33:586-594.

Sneath, P. H. A. & R. R. Sokal. 1973. Numerical Taxonomy. W. H. Freeman & Co., San Francisco.

Sokal, R. R. & F. J. Rohlf. 1962. The comparison of dendrograms by objective methods. Taxon 11:33-40.

Sokal, R. R. & F. J. Rohlf. 1973. Introduction to Biostatistics. W. H. Freeman & Co., San Francisco.

Sokal, R. R. & F. J. Rohlf. 1981. Biometry, 2nd ed. W. H. Freeman & Co., San Francisco.

Sokal, R. R. & P. H. A. Sneath. 1963. Principles of of Numerical Taxonomy. W. H. Freeman & Co., San Francisco.

Standish, L. M. 1916. What is happening to the Hawthorns? J. Heredity 7:266–279.

Stebbins, G. L. Jr. 1950. Variation and Evolution in Plants. Columbia University Press, New York.

Ulf-Hansen, P. F. 1985. An examination of large-scale population changes in eastern North American hawthorns (*Crataegus*, Rosaceae). M. Sc. Thesis, University of Western Ontario.

Ward, J. H. Jr. 1963. Hierarchical grouping to optimize an objective function. J. Amer. Statist. Assoc. 58:236–244.

Wishart, D. 1978. Clustan User Manual, 3rd ed. Rep. No. 47, Interuniversity Research Councils Series, Edinburgh.

END

12/11/85

FIN