Abstract

Using Leaf Architectural Data for Phenetic Ecological Comparison
of Modern and Fossil Forest Stands
Walton Atwater Green
2007

Without reference to the fossil record, our ability to predict future consequences of ecological change is limited by the length of the time over which we can perform ecological experiments or at most the period of time over which ecologists have collected historical data. The longest continuous records go back only hundreds of years, supplemented by anecdotal information from earlier historical times. In order to describe the past or make predictions about the future beyond that range, we must compare ecological data derived from modern and fossil ecosystems. This dissertation represents an attempt to bridge the gap between terrestrial plant ecology and paleoecology by providing a simple method for numerically representing and characterizing forests (both fossil and modern). This allows identification of large scale patterns in time and space: temporal dynamics happening on an order of a thousand years or longer and spatial patterning ranging from an order of a kilometer up to biogeographical (continental) scales. For practical reasons, the examination is restricted to forest ecosystems, which have been some of the most important ecosystems on the earth since the Late Cretaceous and which constitute the majority of the plant fossil record since that time.

Specifically, in this dissertation I propose and test the thesis that semi-quantitative descriptions of the leaf architectural characteristics of forest floras allow direct comparison of fossil and modern forests on functional grounds that bypass taxonomic ambiguity in the fossil record. A corollary is that such descriptions provide a way of identifying large scale patterns and trends in the past and therefore making testable predictions about the future evolution of plant ecosystems.

Discounting anthropogenic changes, the landscapes in which we live are largely created by plant ecosystems. Modern ecologists can only see a brief clip of the moving picture of landscape patterns through time; improved diachronic paleobotanical data would permit a showing of the entire film reel, albeit at lower resolution. The eco-morphological ordination of ancient floras in a modern framework has the potential to provide a general method of paleoecological analysis that is independent of and therefore comparable with paleoecological inferences made from sedimentological or phylogenetic data.

Using Leaf Architectural Data for Phenetic Ecological Comparison of Modern and Fossil Forest Stands

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

Walton Atwater Green

Dissertation Director: Leo J. Hickey

May, 2007

Using Leaf Architectural Data for Phenetic Ecological Comparison of Modern and Fossil Forest Stands

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

Walton Atwater Green

Dissertation Director: Leo J. Hickey

Final version typeset April 30, 2007

For F. L. R., who unfortunately did not live to see it,
and
for K. E. B., who fortunately did.

Acknowledgements

Thanks first to my advisory committee, headed by Leo Hickey and consisting, in addition, of Graeme Berlyn, Dolph Seilacher, Elisabeth Vrba, and Scott Wing. For logistical reasons, Dolph Seilacher was unfortunately unable to serve as a reader as well, but suggestions of his improved chapter 5 in particular, as well as providing general inspiration. *Gern bin ich gedolpht worden.* Discussions with the following people contributed to my understanding of specific parts of this dissertation: Krister Smith on chapter 2, Ian Miller on chapter 3, Christopher Traiser on chapter 4, Stefan Little on chapter 6, John Hartigan on chapter 7. Thanks to Jay Emerson for his role in getting me hooked on *R,* and to Daniel Mason and Anna Martinetz who (respectively in Central Park and on a train between Delhi and Bombay) forced me to articulate more clearly both what I was working on and why I was working on it. Without the support, advice, and assistance of Kaspar Hauser, and the other members of *Gedanken Donuts*, and *Crania et Femora,* I would never have finished this dissertation. +CAKE *is* the primitive condition. Friends and family both nearby and far distant who cannot be enumerated also provided intellectual and emotional support out of all proportion to the time I spent with them in the last seven years. Thanks particularly to David Grewal, Adam Haslett, Julia Prest, and Ben Zaitchik for helping make New Haven more than merely endurable, and to Katy Black for making it less so. For specific comments on various draft manuscripts that have been incorporated into this dissertation, thanks to Katy Black, Jason Downs, Nancy Green, Leo Hickey, Ian Miller, Dana Royer, Dolph Seilacher, Christopher Traiser, Elisabeth Vrba, Scott Wing, and several anonymous reviewers. Thanks also to Emily Schenker, Amanda Baker, Melissa McGrath and several other employees of the Yale Peabody Museum who helped collect much of the raw data examined here and to Jonathan Adams and Yangjian Zhang for their collaboration on chapter 6. Over the past 60 years, the *Compendium Index of North American Mesozoic and Cenozoic Type Fossil Plants* has been maintained financially by Princeton University and Yale University and practically by the work of many volunteers and museum employees including the current paleobotany collections manager at the Peabody museum, Linda Klise, and many others whose names I do not know. I am nevertheless grateful to them for their contributions to a multi-generational project of which we are now reaping the benefit. In addition to my graduate fellowship from Yale University, this work was supported by National Science Foundation grant DEB–0235736. Finally, if I had tried to write this dissertation with commercial software or perform any of the analyses described in it, I would long since have given up in despair and frustration. It could not have been written without LaTeXand *R*, so thanks to the open-source community that has made such powerful and elegant tools available.

Author's Preface

The formal structure of a dissertation has prevented the addition of some comments of a more general and perhaps unscientific nature. If my experiences are typical, doctoral students are often 'fair and young when in hope they began that long journey'. This dissertation is neither as complete nor as conclusive as I originally intended it to be. Nevertheless, I hope that it provides enough thought and data to be either interesting or useful—possibly even both—as a starting point for further examination of an interesting set of problems.

One of the problems with which I have been faced is that there are too many data to show. Thus I have been caught between the Scylla of cherry-picking and the Charybdis of confusing the reader by presenting too much information represented in too many different ways. This superfluity of data seems to me a symptom of changes recently induced in the field of paleontology. I have tried to treat it as a challenge to which it is possible to rise.

Although it is not explicitly discussed, an important aspect of this dissertation is its reliance on data analysis using a computer scripting language, in this case *R*. If Charlemagne is apocryphally supposed to have equated having a second language with possessing a second soul, perhaps learning a computer programming language is like being possessed by a demon. In the process of analyzing the data for this dissertation I have become thoroughly convinced of the value of dynamical data analysis via scripting. The power and flexibility of analyzing data using a scripting language can also, however, be defects: they allow complex manipulation and elaborate graphical displays that can be very difficult for readers to understand at first glance. I have tried to explain my methods fully, but like so many such things, the only way for the reader to understand the results may ultimately be to replicate them. One advantage of data analysis using a scripting language is that given the data in appendix A and the scripts in appendix B, virtually all the results discussed in the text can in theory be replicated by the reader. Unfortunately, as is frequently the case, it would not be easy in practice to perform this simple theoretical exercise. Most of the blame for that is mine: since I began from scratch six years ago and have learned to program as I went along, many of my earlier results went undocumented. Nevertheless in the interests of full disclosure I have attached the data files and scripts as appendices in all their complexity.

As will be apparent, I have used the first person singular freely when it seemed to be warranted.

Some of the figures benefit greatly from being printed at larger sizes, so in certain cases I have replaced the in-text figure with a larger fold-out plate at the end of the dissertation.

LAX ET VANITAS

Yale University motto, approximately

Contents

List of Figures

1 Introduction

Modern ecology has focused primarily on dynamics of competition and physiological change at short temporal scales. Therefore our ability to predict future consequences of ecological change is limited by the length of the time-series collected by modern ecologists. The fossil record, on the other hand, contains data about large scale differences in the worldwide distribution of plants that can potentially put modern ecological issues into temporal perspective, a perspective that is increasingly recognized as important (Rees et al. 2001). Progress in describing ecological change through time has been made for marine invertebrate ecosystems: initially as early as the 1950s (Ladd 1957), most influentially in the work of the Chicago School (Sepkoski et al. 1981) and most recently in the emerging field of 'evolutionary paleoecology' (Allmon and Bottjer 2001). The advances in marine paleoecology may be due partially to the generally better preservation of animals with mineralized skeletons in marine environments: a shell bed in the fossil record provides a relatively good record of the living molluscan community that produced it. Terrestrial ecosystems, however, are generally more biased by taphonomic processes and selective preservation. Terrestrial plant paleoecology, therefore, has been largely descriptive (Behrensmeyer et al. 1992); that is, there has been little success in using models of ecological process to guide description of ancient communities. In broad terms, this dissertation attempts to link terrestrial plant ecology with paleoecology by providing a method for semi-quantitative description of fossil and modern angiosperm forests allowing us to identify large scale patterns in time and space—temporal dynamics happening on an order of a millennium or longer and spatial patterning up to continental scales. The average accumulation rate of a stratigraphic bed in terrestrial environments provides our smallest general time unit and the forest stand or fossil collection locality provides the smallest unit of spatial analysis. Shorter time scales and smaller areas generally can not be resolved in the geological record and are therefore not the subject of this research. For practical reasons it is restricted to angiosperm forest ecosystems. These have been some of the most important ecosystems on the earth since the Late Cretaceous and constitute the vast majority of the plant fossil record since that time.

As one way to bridge the gap between plant ecology and paleoecology, this dissertation offers a new method for the ecological analysis of fossil plant ecosystems, which I am provisionally calling 'leaf ecophenetics' because it depend on the phenetic or algorithmic classification of plant ecosystems based on quantitative descriptions of their leaves. Chapter 2 explains and develops this methodology in some detail. In this introductory chapter, I provide a brief summary of existing methods of plant paleoecology, introduce leaf ecophenetics in section 1.2, and summarize the structure and content of the following chapters.

An explicitly hypothetico-deductive framework does not seem obviously applicable in a dissertation like this one, focussed on developing a new method rather than applying a known method to an unresolved question. Methodological development, however, can also be seen as testing: in this case, testing whether semi-quantitative descriptions of the leaf architectural characteristics of forest floras (such as are described in chapter 2) allow direct comparison of fossil and modern forests on functional grounds that bypass taxonomic ambiguity in the fossil record. This thesis, I believe, can easily be demonstrated for certain cases as is done in chapter 2. The real difficulty is in showing that such a comparison is useful and ecologically informative. Therefore the bulk of the dissertation (chapters 3–6) is devoted to illustrating broad applications of the method as the best evidence for its utility.

1.1 Existing Methods of Plant Paleoecology

Before describing the method being developed in this dissertation, I should provide a brief survey of the existing alternatives. Further detail can be found in Krassilov (1975), DiMichele and Wing (1988), and Dodd and Stanton (1990). Geochemical methods and modeling are included for the sake of completeness, but most biogeochemical models and isotopic proxies are paleoclimatological rather than paleoecological in general intent. This is an important distinction that will reappear in chapter 2: paleoclimatology is concerned with biological processes only insofar as they reveal what characteristics of fossils will make good proxies for ancient climates; paleoecology is relatively uninterested in the historical description of ancient climates *per se*, instead being concerned with the biological processes that tend to occur under given situations. To the ecologist, paleoclimates are interesting because they reveal regularities in biological response to extrinsic forces; not simply to document, for instance, when in the past it was hot and when it was cold. The following hierarchical list gives a classification (in rough chronological order) of the methods of paleoecological analysis that have been widely applied to fossil plant assemblages.

1. Nearest living relative approach ‘NLR approach’ (see discussion below)
    (a) Single taxon
        i. Weak phylogenetic inference (de Queirioz and Gautier 1992)
        ii. Strong (justified) phylogenetic inference or ‘bracketing’ (Witmer 1995)
    (b) Multiple taxon or ‘coexistence approach’ (Mosbrugger and Utescher 1997)
        i. Weak phylogenetic inference
        ii. Strong phylogenetic inference
2. Sedimentology (see discussion below)

    (a) Facies analysis (Crosby 1972; Gall 1983; Miall 1996)
    (b) Paleosol interpretation (Retallack and Germán-Heins 1994)
        i. Pedogenic carbonate horizon analysis (Royer 1999)
3. Ecomorphology or Functional Ecology (see discussion below)
    (a) Leaf Physiognomy Bailey and Sinnott (1915)
        i. Univariate models (Leaf Margin Analysis, LMA) (Wilf 1997)
        ii. Multivariate models
            A. Multiple Linear Regresssion (Wing and Greenwood 1993; Wiemann et al. 1998a)
            B. Climate Leaf Analysis Multivariate Program (CLAMP) (Wolfe 1993)
            C. Digital Leaf Physiognomy (Huff et al. 2003)
    (b) Wood Physiognomy (Wheeler and Baas 1993)
    (c) Leaf Ecophenetics (Green and Hickey 2005; Green 2006)
4. Geochemistry, *e.g.* light stable isotopes from soil carbonates
5. Modeling, *e.g.* global circulation models, weathering models (Berner and Kothavala 2001)

1.1.1 'Nearest living relative' and related approaches

The development of methods for paleoecological analysis began implicitly early in paleobotany with what is now called the nearest living relative approach or NLR. In its current form, the nearest living relative approach consists of a more-or-less rigorous phylogenetic analysis of each plant fossil from a particular collection locality. The nearest living relative of each fossil taxon is recorded along with the range of environments that it currently inhabits. It is then assumed that the fossil taxon has inhabited an environment within the range in which its nearest living relative is observed. When a number of fossil taxa are found autochthonously deposited in the same locality, the environmental ranges given by each of their nearest living relatives can be intersected in order to increase precision. This is called the *coexistence approach* by Mosbrugger and Utescher (1997). In the list above, I distinguish strong and weak phylogenetic inference. This is my distinction, based on the difference illustrated in figure 1.1 between *justified* and *unjustified phylogenetic inference* (de Queirioz and Gautier 1992) or *bracketing* (Witmer 1995).

If taxon A, the nearest living relative of fossil taxon X, has character 1 then *ceteris paribus* it is reasonable to assume that X also shares character 1. This, however, is the weakest possible type of phylogenetic inference because it depends both on the fact that A and X really are sister taxa *and*

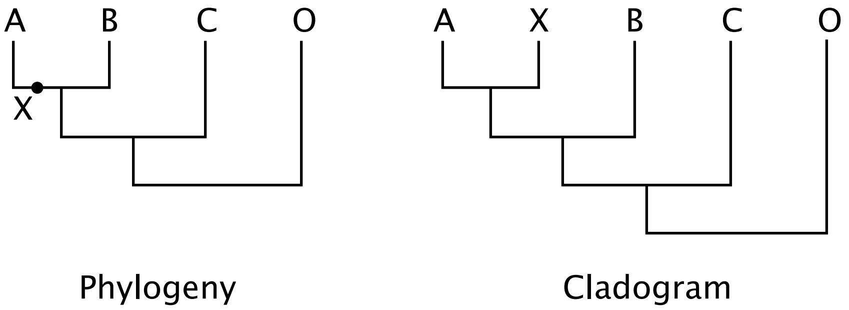


Figure 1.1: Theoretical phylogeny (historical relationships) and cladogram (abstracted branching order) of living taxa A, B, C, outgroup O and fossil taxon X to illustrate levels of phylogenetic inference.

on the assumption that 1 is a derived character. If B also has character 1, then a bracket is present and the phylogenetic inference can be considered justified, *i.e.* stronger evidence, because if the phylogeny is correct it is parsimonious to consider the character derived. If C also shares character 1, then the evidence that X had the character is still stronger because even if the phylogeny is slightly wrong, it is still parsimonious to infer that X shared the character. The more close relatives that X has all possessing character 1, the more certain we are that X must have had it as well, again assuming the approximate accuracy of the phylogenetic reconstruction.

Paleoecological inferences based on genetic relationships are frequently useful, but are naturally limited by the necessity of performing a complete and accurate phylogenetic analysis, which is always laborious and sometimes impossible with plant fossil assemblages.

1.1.2 Sedimentology

Stratigraphic description of fossil plant collections had been standard since the 19th century because it was needed for dating purposes. During the late 1970s and 80s some paleobotanists (e.g. Hickey 1977; La Pasha and Miller 1984) began to collect sedimentological information about the rocks in which fossil floras were found. This information included small-scale measured sections, grain-size and texture descriptions, many closely spaced and quantitatively sampled collection localities, bed geometry, and sometimes chemical analyses of sediments. In addition to contemporary work on analysis of paleosols (Retallack 1983; Krauss and Bown 1988; Retallack and Germán-Heins 1994) and facies analysis of the fluvial environments in which plant fossils are generally found (Allen 1965; Crosby 1972; Schumm 1972; Gall 1983; Bridge 1993; Miall 1995, 1996), these data allow certain inferences about the local environmental conditions. For instance, laterites and calcic or gypsic hard pans were shown to indicate tropical and dry subtropical conditions respectively; aggrading fining upward sequences implied proximity to a river channel, and carbonaceous shales distal or backswamp conditions. Useful as these sedimentological interpretations frequently are, they generally provide only qualitative information about climatic and environmental conditions. Quantitative estimates based on these have often been found problematic, as in the case of using calcic hardpan depth to estimate paleoprecipitation (Royer 1999).

1.1.3 Ecomorphology or Functional Ecology

Since early in the 20th century (Bailey and Sinnott 1915, 1916) there have been various attempts to obtain environmental information directly from plant morphology, of which the most successful and widely applied has been called *leaf physiognomy* (Dolph 1976, 1978, 1979; Wolfe 1979; Dolph and Dilcher 1980a,b; Williamson 1981; Williamson et al. 1983; Dolph 1984). In the early 1990s a new leaf physiognomic approach called the Climate Leaf Analysis Multivariate Program (CLAMP) was introduced by Wolfe (1993), and subsequent debate over the merits of such an approach has led to the division of leaf physiognomy into two methodological schools. The older approach, based on linear regression of a single variable (percentage of entire-margined leaves) against mean annual temperature, came to be known as *leaf margin analysis* or LMA, while CLAMP acquired a specific group of adherents, developers, and critics (Kovach and Spicer 1995; Wolfe 1995; Herman and Spicer 1996; Stranks 1996; Kennedy 1998; Jacobs 1999; Spicer 1999; Wolfe and Spicer 1999; Gregory-Wodzicki 2000; Jacobs 2002; Kennedy et al. 2002; Kowalski 2002; Spicer et al. 2004; Traiser 2004; Spicer et al. 2005; Traiser et al. 2005; Green 2006). In addition, many other leaf physiognomic analyses performed in the 1990s relied on multivariate data without following the CLAMP protocol or compared different methodologies (Wing and Greenwood 1993; Jacobs and Deino 1996; Jordan 1997; Wilf 1997; Wilf et al. 1998; Wiemann et al. 1998a, 2001; Huff et al. 2003; Kowalski and Dilcher 2003; Greenwood et al. 2004; Royer et al. 2005; Miller et al. 2006; Royer and Wilf 2006). Though many of these studies employed both leaf margin analysis and CLAMP, there have remained differences in approach between supporters and opponents of: (1) complex analyses of multivariate data as opposed to simple univariate regression and (2) multivariate ordination techniques like principle components analysis, as opposed to simple linear regression.

Relatively fewer studies have looked at wood anatomy Baas (1986); Wheeler and Baas (1993); Woodcock and Ignas (1994); Wiemann et al. (1998b, 2001) with similar physiognomic ends in mind, and somewhat less work (Raunkiaer 1934; Richards 1939; Givnish 1978) has been done to relate plant growth form to environmental parameters quantitatively.

1.2 Leaf Ecophenetics: A New Method for Plant Paleoecology

Leaf ecophenetics is a methodology for producing semi-quantitative and ecologically meaningful descriptions with which to compare fossil and modern forest floras. It springs directly from leaf physiognomy: the data are collected exactly the same way as it is for CLAMP or another multivariate leaf physiognomic method, but the intent behind the investigation and the methods of data analysis are very different. Instead of seeing the analysis of these multivariate data on the distribution of leaf architectural data as a problem of *estimation*, leaf ecophenetics sees it as a problem of *classification*. The intent is ecological as opposed to paleoclimatological and the data analysis is exploratory rather than being based on regression models. For these reasons, I list leaf ecophenetics in the table in the previous section as part of ecomorphology or functional paleoecology, and not as a subsection of leaf physiognomy, which has since its inception been primarily focussed on climate.

Ecophenetic descriptions of floras rely on data from leaf architecture as a proxy metric for aspects of the ecosystem because (1) inter-comparable leaf-architectural data are easily obtainable from imperfectly preserved or poorly described fossil and modern floras, and (2) a strong prima facie case has been made for the presence of an ecological signal in leaf morphology (Givnish 1986). The method that is being proposed to extract this signal is by coding floras (assemblages of leaves collected from the same locality) in such a way as to produce for each flora an n-dimensional vector that describes the leaf architectures found in the flora. There are two main published methods for producing this vector, the first based on Compendium Index Categories or CICs (Ash et al. 1999; Green and Hickey 2005) and the second on the Climate Leaf Analysis Multivariate Program or CLAMP (Wolfe 1993, 1995). In either case, the leaves from a fossil bed or a forest stand are sorted into categories like ‘Toothed’ and ‘Entire’ or ‘Pinnately veined’ and ‘Palmately veined’. Then the proportion of leaves in each category is determined and these proportions are arranged in a vector that describes the composition of the entire flora. There are numerous variations on the exact method of coding (*e.g.* CLAMP has 31 and CICs 64 variables); chapter 2 includes an experiment comparing the two published methods. Analysis of these floral vectors by hierarchical cluster analysis and multivariate graphical analysis confirms expectations that regardless of the coding method leaf shape reflects environment (among other variables) and that the tools of exploratory data analysis provide a means of deconvoluting these variables.

The semi-quantitative descriptions of forest floras that are produced by this methodology allow direct comparison of fossil and modern forests on functional grounds that bypass taxonomic ambiguity in the fossil record. Exploratory data analysis allows such data to be presented in such a way that a large amount of detail that is significant to inter-stand variation can be quantitatively described and visually appreciated. This connects the microscale of site census data with the macroscale of diversity curves through geological time by emphasizing the variation in the data that is significant to ecological characterization at large scales.

This methodology is made particularly useful by the presence in the paleobotanical literature of a large body of descriptive information on fossil floras, which provides a rich source of floras to code. Moreover, the *The Compendium Index of North American Mesozoic and Cenozoic Type Fossil Plants*, a bibliographic and taxonomic database housed at the Yale Peabody Museum, provides about 250 fossil floras that have already been coded by CIC. This dissertation provides a large enough body of data on modern floras to support a multi-dimensional framework within which fossil floras can be ordinated.

1.3 Sources of Data and Structure of the Dissertation

This section provides a summary of the data that are dealt with in the body of the dissertation. Chapter 2 validates ecophenetics as a methodology. Chapters 3–6 are relatively stand-alone applications of different aspects of the general method, applied to different data sets and testing different hypotheses. Chapter 7 provides a more theoretical examination of the analytical problem of classifying vegetation on the basis of architectural scores; and chapter 8 concludes the dissertation and provides a program for continued research.

1.3.1 Validation of Leaf Ecophenetics Using Newly Coded Data

Chapter 2 describes in more detail the new paleoecological methodology introduced briefly above in section 1.2. The data discussed consist of multivariate morphological ‘scores’ for 162 modern and 8 fossil floras obtained using one of the two existing scoring methods (CLAMP and CICs). In addition to exploratory analyses of these data, chapter 3 provides a basic validation of the methodology and defense of the thesis that semi-quantitative descriptions of the leaf architectural characteristics of forest floras allow direct comparison of fossil and modern forests on functional grounds that bypass taxonomic ambiguity in the fossil record.

1.3.2 Variation Through Time: Data from the Compendium Index

Chapter 3 deals with diachronic variation—patterns though time—with reference specifically to data from *The Compendium Index of North American Mesozoic and Cenozoic Type Fossil Plants*, a card index of fossil plant occurrences held in the Yale Peabody Museum. In particular I focus on the

Cretaceous/Tertiary boundary, and show how our appreciation of its influence on plant evolution depends on the scale of analysis. An article has already appeared in print (Green and Hickey 2005) from which most of the material in chapter 3 is taken directly, and an electronic version of the *Compendium Index* has been released as Hickey et al. (2006).

1.3.3 The Problem of Inhomogenous Spatial Sampling: Meta-analysis of CLAMP

Chapter 4 is a detailed reanalysis of data from four studies (Wolfe 1993; Jacobs 1999, 2002; Gregory-Wodzicki 2000; Kowalski 2002) that have published multivariate data on modern forests obtained using the CLAMP coding method. All of these data have appeared before in print, so the raw numbers are not provided again here; the electronic version of the data and the scripts used for data processing are provided in the appendices. A version of chapter 4 has already been published (Green 2006).

1.3.4 Dealing with Spatial Variation using Synthetic Floras

Chapter 5 raises the issue of spatial variation and the potential for irregular sampling to bias conclusions about relationships between leaf architectural variables and environmental conditions. It contains a reanalysis of native North American tree distribution maps drawn by Elbert Little (Critchfield and Little 1966; Little Jr. 1971, 1976, 1977, 1978) and digitized by Thompson et al. (1999). The methodology employed is similar to that in Traiser (2004) and Traiser et al. (2005), and the results reported here are based on a collaboration with Jonathan Adams and Yangjian Zhang of Rutgers University. An article reporting our conclusions is currently in preparation (Adams et al. in prep).

1.3.5 Taxonomic Variation: Analysis of data from the National Cleared Leaf Collection

Temporal variation is treated in chapter 3; spatial in chapters 4 and 5. Chapter 6 is concerned with taxonomic variation. The data examined consist of the CIC codes for the 6767 specimens of the *National Cleared Leaf Collection* held at the Yale Peabody Museum.

Chapter 7 is a more theoretical treatment of some of the classification problems introduced in chapter 2. In it, I employ data drawn from chapters 2 as well as some synthetic data to show the operation of a new method for clustering floras based on *structured factors.*

1.4 Broader Significance

Chapter 8 discusses the broader significance of this project: the potential for paleoecology to provide time depth for ecological studies with which to extend predictions about future ecological change.

The extent to which organisms and communities of organisms can create and modify their environments can best be seen at long time scales.

From the perspective of an ecologist or evolutionary biologist, it is more important to classify and understand variation in ecosystems than it is to be able to use them to predict or estimate climatic parameters. The landscapes we see around us are created not by the climate alone but by complex interactions between extrinsic (physical and geological) forces and evolving biological communities. Climatology and paleoclimatology have their own importance and significance, but do not explain biological evolution.

Therefore a paleoecological method like ecophenetics that is focussed on classification may have more potential for elucidating ecological dynamics than methods—even if they provide better paleoclimatic proxies—that treat biological processes only as a convenient way to record the inorganic environment.

2 Leaf ecophenetics: A new semi-quantitative method of vegetation classification

In this chapter, I suggest a new semi-quantitative method of vegetation classification based on the architecture or physiognomy of leaves. This method, which I am calling *leaf ecophenetics,* suffers less from the sort of subjective and taxonomic biases that plague current methods and seems to produce hierarchical classifications that are in reasonable accord with existing classifications. Leaf ecophenetics is a method for the classification and ordination of plant communities that is applicable to modern ecosystems and to fossil plant assemblages. It differs primarily from other methods of ecological or paleoecological analysis in being (1) ecomorphological rather than systematic, (2) semi-quantitative rather than qualitative, (3) focussed on classification of communities or ecosystems rather than estimation of environmental parameters.

Most existing methods for classifying vegetation rely on the presence or relative abundance of identifiable species in a flora, and often depend on subjective estimates of coverage or abundance. Leaf ecophenetics depends instead on the distribution of leaves in architectural categories. I use two different published methods of locating floras in n-dimensional morphospace based on leaf architectural variables and then apply hierarchical clustering algorithms and graphical analysis in order to display these scatters of ordered n-tuples as classifications. The classifications obtained are broadly consistent with each other and with prior knowledge about the floras, indicating that the procedure can produce reasonable classifications of small, homogeneously collected data sets.

2.1 Historical Context

Early on (von Humboldt 1807) plant ecology was not distinguished from plant geography, but during the second half of the nineteenth century, the fields of study that had been known loosely as natural history were broken up and reconstituted under the rubric of Darwinian evolution as biology and its various subdisciplines. The name ecology for the subdiscipline dealing with organisms in their environments was coined some time in the 1850s. Haeckel defined the term in 1870 (Kormondy 1969; McIntosh 1985), though ecology does not seem to have become self-conscious as a discipline until the very end of the 19th century. Moreover, Haeckel's definition was firmly zoological, while the appearance of ecology as a discipline in the 1890s was largely as a development of plant geography, the roots of which go back well into the 18th century. Within ecology, the history and methodology of vegetation classification have been fully described by Whittaker (1962), Whittaker (1973), and Shimwell (1971), and a sourcebook of foundational papers is available edited by McIntosh (1978). Unlike systematics, however, particularly phylogenetic systematics or cladistics, vegetation classification has not been substantively affected by the introduction of the personal computer in the 1980s and 90s. Instead, the entire field of community classification, once considered central to ecology, is scarcely mentioned in current introductory courses or textbooks. The reason vegetation classification has recently been neglected seems not to be its lack of importance, but frustration with its persistent subjectivity, and difficulty in integrating it into the increasingly mathematical framework of population ecology.

The intent of this chapter is to suggest a method of vegetation classification that meets current criteria for operator repeatability while pursuing the same traditional goal of classifying plant ecosystems in biologically useful ways. Applications of this method are provided in subsequent chapters; here it is intended only to show that the methods chosen for producing semi-quantitative descriptions of floras encode information on variables of biological interest like biogeography and canopy differentiation, as well as noise due to the coding process.

2.2 Methods of Data Collection

In order to address the problem of vegetation classification in the face of distrust of traditional methods, I have developed a method of classification that seems to be comparatively insensitive to subjective bias, produces quantitative or at least semi-quantitative descriptions of plant communities, and encapsulates or displays a reasonably large amount of ecologically interesting information. As a method it is not intended to replace traditional methods; instead it has it's own sphere of validity, strengths and weaknesses. Its particular strengths are its ability to support strictly hierarchical, fuzzy and non-exclusive classifications, its simplicity, and the ease with which it can be interpreted or the original data recovered from the classification. Weaknesses include its lack of generality (it only applies to communities dominated by woody dicots, i.e. forests) and its reliance on a measurement system designed for easy application not for ecological significance. It also spreads variation that should probably be measured by a smaller number (about 10) quantitative variables out over a larger group (31 or 56) ordinal variables. This makes data analysis and application of statistical methods needlessly difficult and complex.

In the following analyses, the units being classified are all plant communities *sensu* the group of plants growing in a well-defined geographical area during a historical duration of time. These communities can be treated as individuals (Clementsian superorganisms) or they can be considered communities (sets of populations). I will remain agnostic on this point, though the discipline of ecology has been driven in the individualistic direction for purely historical reasons. Classification is needed in either case and in either case the attributes that make a good classification are the same. A good classification should be:

1. Simple to make from easily collected data
2. Invariant with respect to the person who makes it, the time and place it is made
3. Capable of revealing or displaying biologically interesting variables
4. Stable under application to other data
5. As little removed as possible from the raw data used to produce it
6. Flexible enough that it can be adapted for disparate purposes

This is not an exhaustive list, and obviously some of these desirable attributes are mutually exclusive, *e.g.* a totally unambiguous classification is liable to be ecologically uninformative.

Many characteristics could have been chosen to describe vegetation, but for ecological purposes there is a *prima facie* rational for choosing elements of the vegetative body of the plant that are particularly susceptible to modification under different environmental conditions, i.e. that are most ecologically plastic. With the possible exception of overall growth form (plant architecture), the leaf is notably the most environmentally plastic plant organ. An additional reason for focussing on the architectural elements of leaves is their applicability to fossil ecosystems from which only leaf impressions are preserved. This choice unfortunately limits the method to dicot forest ecosystems, a limitation that is less restrictive because the majority of preserved fossil material since the middle Cretaceous is from such ecosystems.

There are a number of ways to describe a dicot leaf assemblage or floras by a numerical vector that reflects the shape and venation of the leaves in the flora. Two of these have been extensively applied. The first and most broadly applied is the Climate Leaf Analysis Multivariate Program (CLAMP) of Wolfe (1993). The second is the set of Compendium Index Categories (CICs) used to organize the *Compendium Index of North American Mesozoic and Cenozoic Type Fossil Plants* at Yale University (Hickey et al. 2006). The *Compendium Index* was the first systematic catalogue of fossil plants, started by Dorf (1940) as an aid to taxonomic identification. It is cross-indexed taxonomically, stratigraphically, and alphabetically, and organized according to the CICs, simple categories of leaf shape and venation defined in Ash et al. (1999) and Green and Hickey (2005). Definition of these categories is based on petiole attachment, primary venation pattern, leaf-shape, and tooth type; the combinations of these variables produce 56 numerically labeled categories into which reasonably well-preserved broad-leaved dicot leaves can be sorted. Though the *Compendium Index* was primarily intended for cataloguing North American fossils, similar or derivative schemes have been applied to other modern floras over the past two decades, including Australian (Christophel and Hyland 1993) and Chinese (Yu and Chen 1991) floras. These applications of the same principles of classification to very different floras show the robustness of the method and indicate that the current *Compendium Index* base can easily be extended to incorporate floras worldwide. Furthermore, because the *Compendium Index* was designed for classification of fossils, the characteristics needed in order to assign a CIC to a specimen are easily recognized in fossil material or from imperfect illustrations in old publications. For convenience, I have used the CIC numbers (100–164) as variable names, but these numbers are not themselves significant: they are merely bins into which leaves can be sorted based on their morphology and structure.

I coded five floras (11 if canopy and understory components of the same flora are split apart) using both CLAMP variables and CICs in order to compare two different procedures for producing quantitative descriptions of leaf floras. As discussed above, the coding procedure used should be applicable to fragmentary fossil material and inadequate publications, easy and quick to perform and repeatable by different researchers, comparable when applied to modern and fossil assemblages, and ecologically informative. Naturally neither of these procedures fulfils all these desiderata: the CICs were designed to allow easy (taxonomic) identification of fossil leaves, therefore by construction they supply the first two desiderata mentioned above: wide applicability and ease of application. CLAMP variables take much longer to code, but were chosen for their functional importance so they are more likely to provide ecologically meaningful descriptions. Below I describe the evidence provided by a sample data set that the ecological signals I am trying to extract are so strong that it does not matter which set of variables is ultimately chosen for the analysis.

In addition to the eight floras coded by both methods, 162 other modern floras were coded by CIC to provide a broad framework within which fossil floras could be ordinated. Floral lists and leaf descriptions were obtained from Braun (1950); Maycock (1994); Smith (1954); Berry (1916); MacGinitie (1962); Sargent (1905); Gentry (1993); Grimm (1957); Preston (1961); Gleason and Cronquist (1963); Little Jr. (1971); Lee (1935); Wang (1961); Wu and Raven (1994–); Krüssmann (1985); Little Jr. et al. (1964, 1974); Leon and Alain (1946–1962); Richards (1939); Proctor (1984); Espinal and Montenegro (1963); Stoffers (1956); Dansereau and Buell (1966); Liogier (1995); Britton and Wilson (1923–1930); Stoffers (1962–1984); Hooker (1872–1897); Hutchinson and Dalziel (1954–1968); Kong and Watts (1993); Lind and Morrison (1927); Sarlin (1954); Champion and Seth (1968); Champion et al. (1965); Rodwell (1991); Mueller-Dombois and Fosberg (1998); Wardle (1991); Groves (1994); Beadle (1981); Armstrong (1993); Tamm (1989); Hickey (1999); Johnson (1985); Williams (1988); Tutin et al. (1964–1983); Stace (1997); Allen (1961); Wagner et al. (1999); Smith (1979–1981); McMullan (1999); Hickey (1987–1992); Axelrod (1958, 1966, 1962); Chaney et al. (1938); MacGinitie (1969); Chaney et al. (1944); Hickey (1977); Becker (1961); Lakhanpal (1958); MacGinitie (1974) and MacGinitie (1953). (Note that these sources are listed in the order of a temporary reference number used in the original data scoring sheets.)

Figure 2.1 is a flowchart that shows the process of data acquisition and analysis schematically;

the following sections provides analysis of various subsets of the data in order to demonstrate the reliability of the coding method.

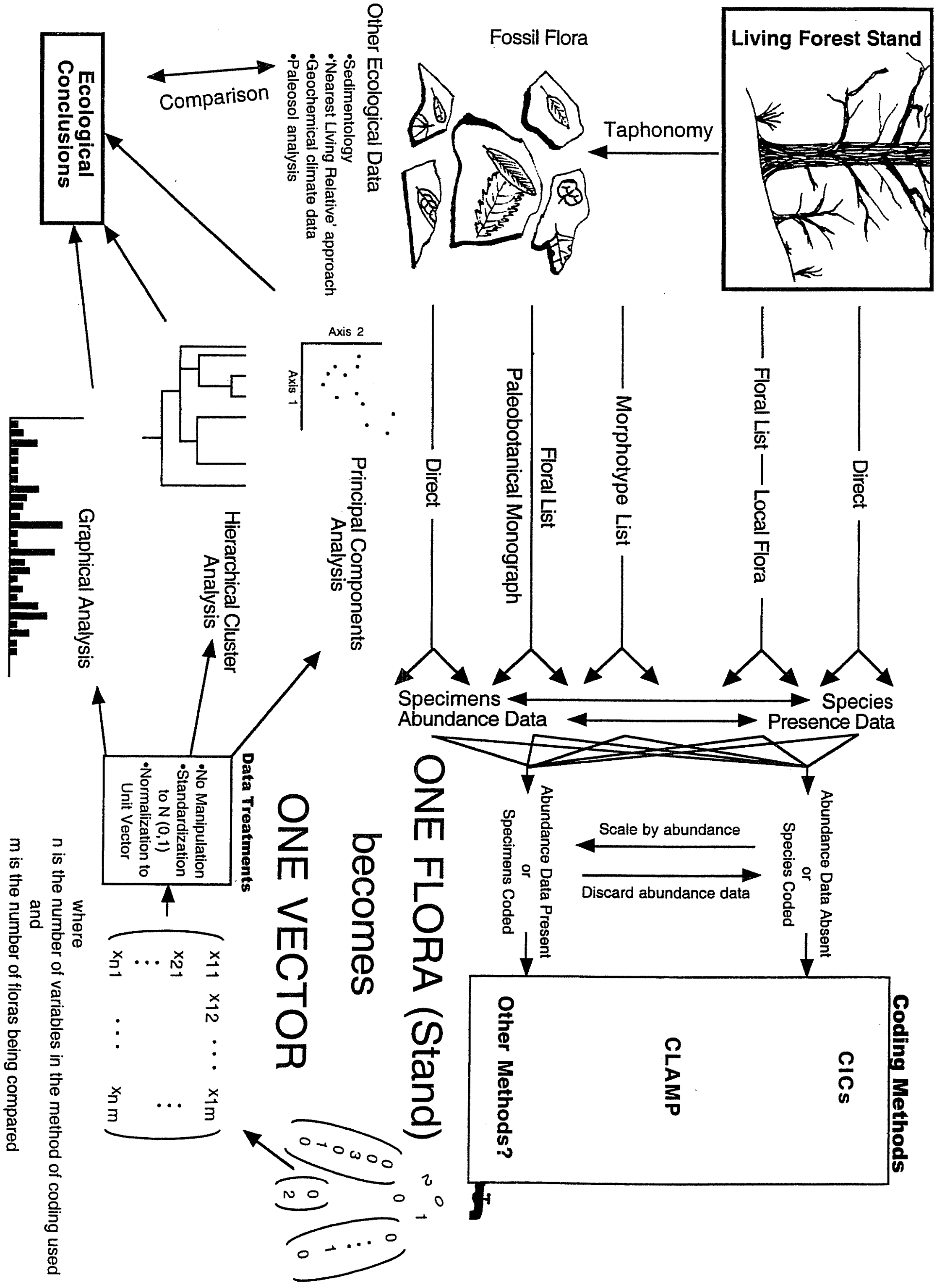


Figure 2.1: Flowchart showing procedure for data acquisition and analysis.

2.3 Data Analysis

Validation of this methodology rests on two bases: first, on the concordance between the two methods of coding that we have applied and second on the way in which traditional classifications are in general terms substantiated by the new technique.

Figure 2.2 shows how a similar clustering structure can be recovered from morphological codes of a simple, homogeneously collected set of floras. Note that in this case CICs do a better job than CLAMP variables at recovering the expected relationships. (Here and below, the term ‘relationship’ does not, of course, imply genetic relatedness, but merely similarity of leaf architecture.) The floras considered are all derived from Braun (1950) and consist of two deciduous floras (one upland, one lowland) from the Duke University experimental forest in North Carolina, in the ‘Oak-Pine Region’; two floras from Kentucky from the ‘Mixed Mesophytic Region’, and one broadleaf evergreen flora from Louisiana, in the ‘Southeastern Evergreen Region’. The suffix ‘.c’ refers to the *canopy* component of each flora; the ‘.u’ suffix the *understory.*

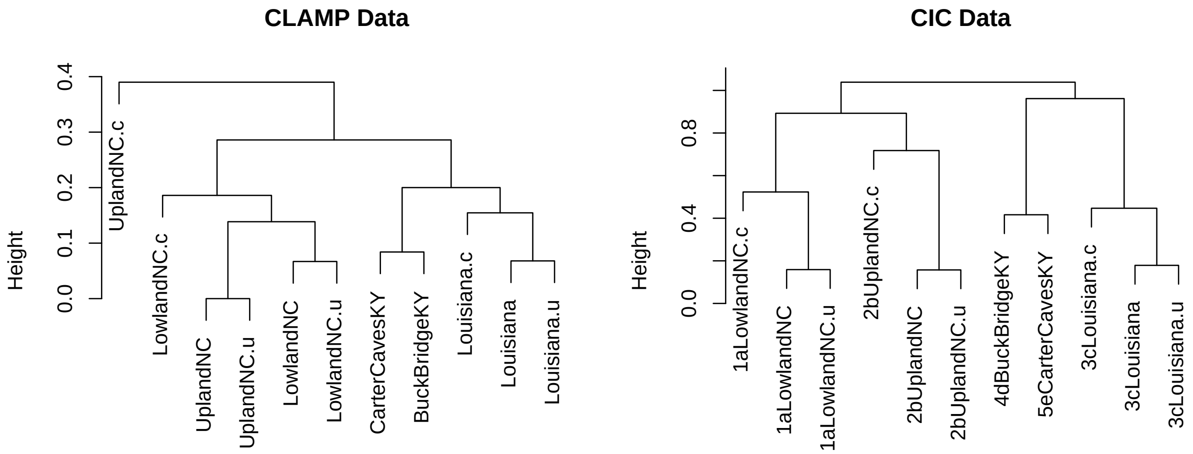


Figure 2.2: Comparison of clusters obtained using the CICs and CLAMP variables.

Figure 2.3 shows three of the same five floras as are clustered in figure 2.2; here, however, the actual spectra for each flora are shown as well as the clusters produced by a hierarchical clustering algorithm. All of the floral lists from which these were coded come from the same source (Braun

1950), and all the actual coding was done by the same person (the author). The form of presentation is called a ‘heatmap’. It consists of a matrix of shaded blocks showing the values or weights that each variable (CIC) takes for each flora under consideration. One-way clustering of floras is performed; i.e. the variables remain in a fixed order, but the floras are hierarchically clustered and reordered according to their distances from each other. Vertically oriented lines of dark blocks show variables (leaf forms) that are shared by floras; horizontally oriented lines indicate floras that have unique leaf forms (among the floras being considered). The remainder of the figures in this chapter consist of heatmaps of different subsets of the data showing how different aspects of the coding process affect the results. The colored blocks to the left of the heatmap in each figure are used to show a single categorical variable such as the person coding the flora or the region from which the flora came.

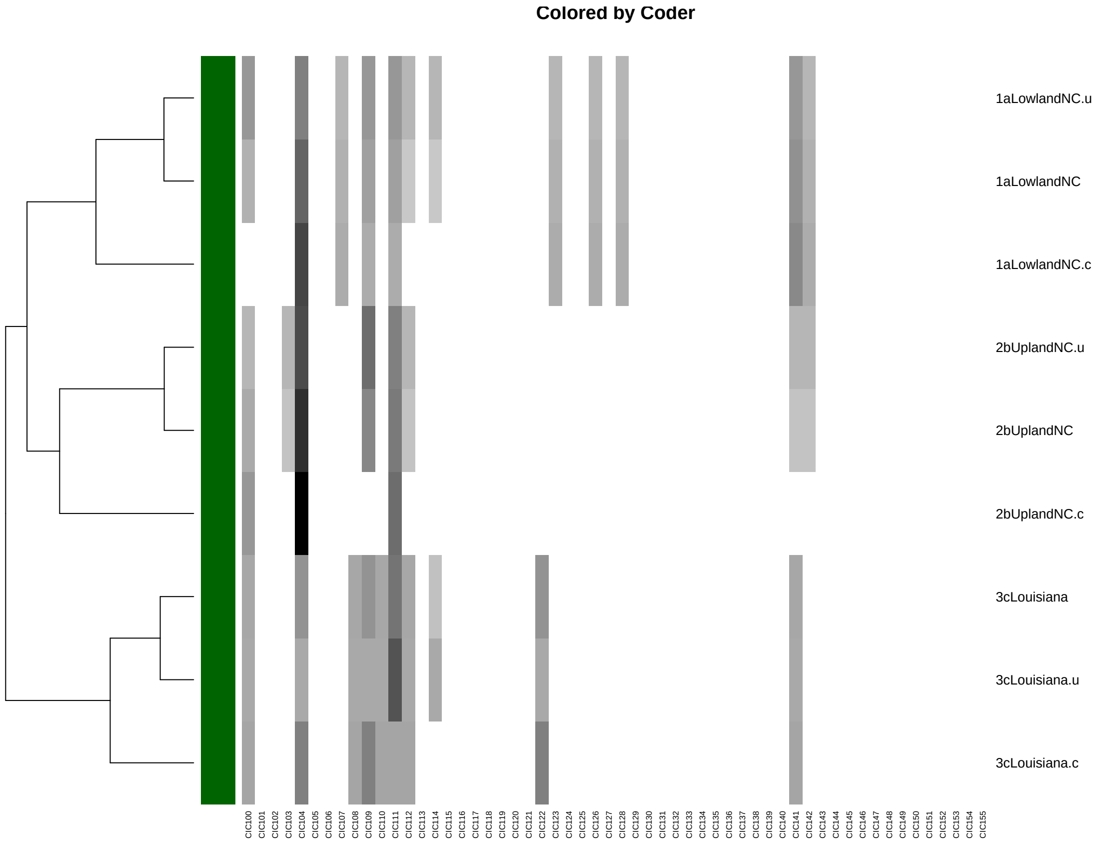


Figure 2.3: Heatmap of three floras (split into nine floras and subfloras) with hierarchical clustering based on CIC scores. All floras were coded by the same person (the author).

In figure 2.4, recoded versions of the same three floras by two other people are added to those shown in figure 2.3. Note the noise induced by the imperfect coding process: floras no longer cluster together as they are supposed to, although their relationships are still approximately correct. The colored bar at the left side shows the three different coders whose systematic biases influence the quantitative descriptions produced: light green is for E. A. Schenker, medium green for A. L. Baker, dark green for the author.

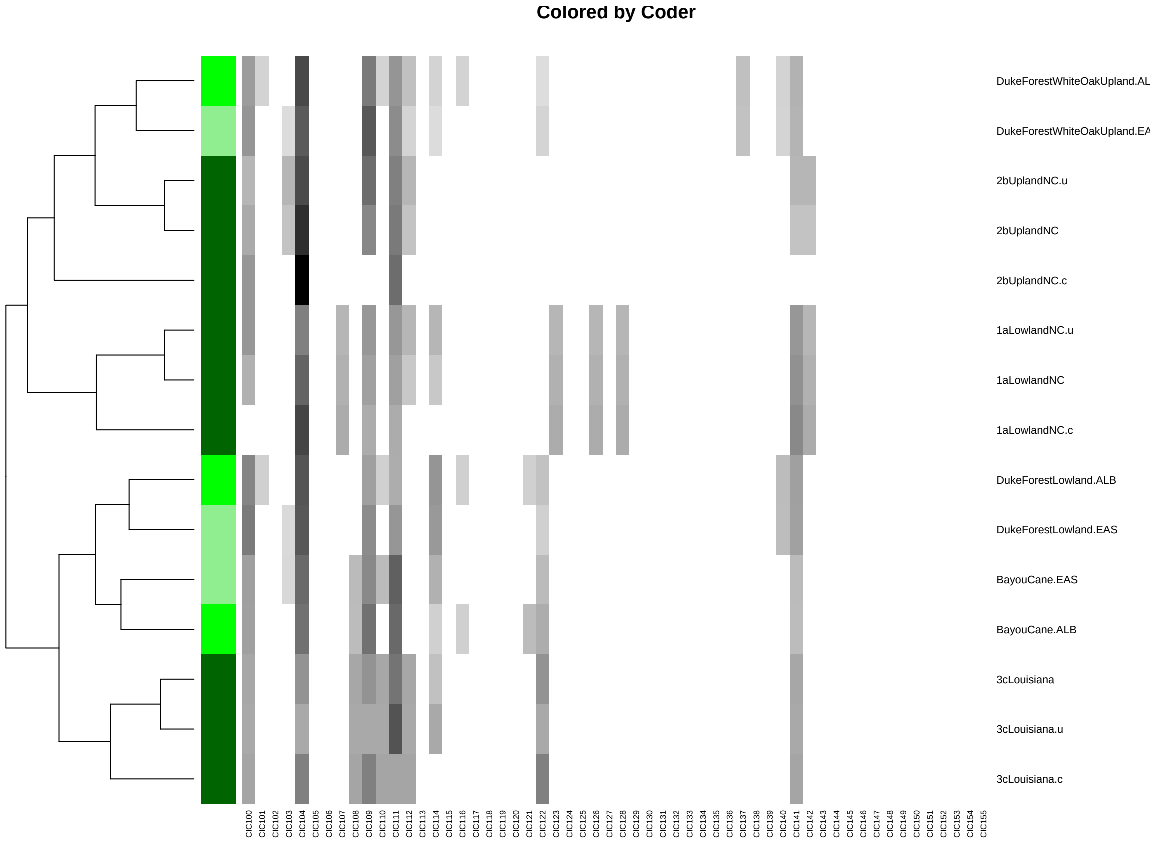


Figure 2.4: Same as figure 2.2, but with each flora coded by three different people. Note the inaccuracy induced: the upland floras remained distinct enough to cluster together regardless of coder, but there was confusion between the evergreen and deciduous lowland floras. Note that the floras labeled ‘BayouCane’ are the same as those labeled ‘Louisiana’ and the words ‘DukeForest’ and ‘WhiteOak’ can be ignored—they are artifacts of a change in naming convention.

Figure 2.5 shows the introduction of additional noise from coding source: here the colored bar to the left refers not to different coders but to different source publications, another variable that adds noise to the ecological signal. Dark blue indicates a flora from Braun (1950); light blue from Maycock (1994).

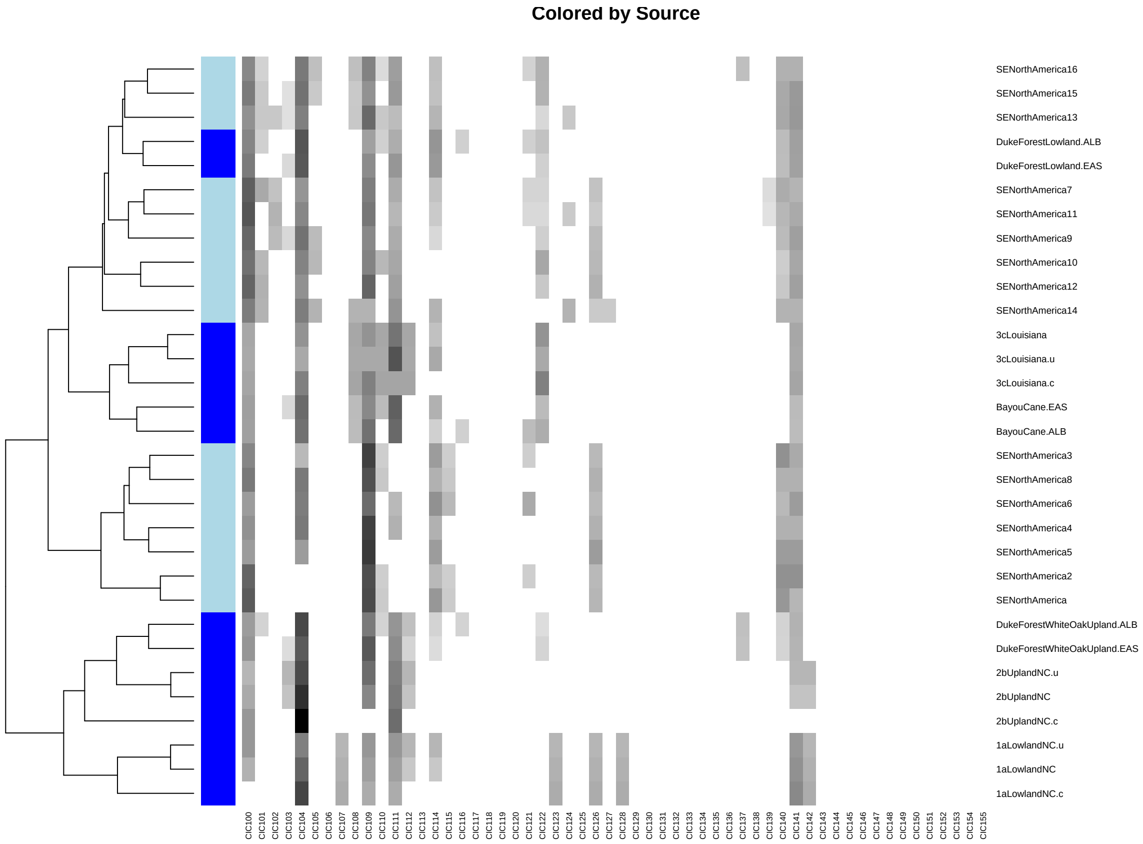


Figure 2.5: Addition of floras coded from floral lists in a different source: again note how there is source noise (different sources tend to cluster together), but that the source noise does not totally swamp the geographical or ecological signal.

Figure 2.6 shows all the North American floras in the data set: as with phylogenetic reconstruction, as more terminals are added, the precision with which their relationships are shown decreases. Nevertheless, despite the noise due to different coders and different source publications, strong signals like the clustering of northeastern mixed deciduous forests (the bottom five floras in the figure) remain evident. The red bar to the right indicates that all the floras are from the same region (North

America).

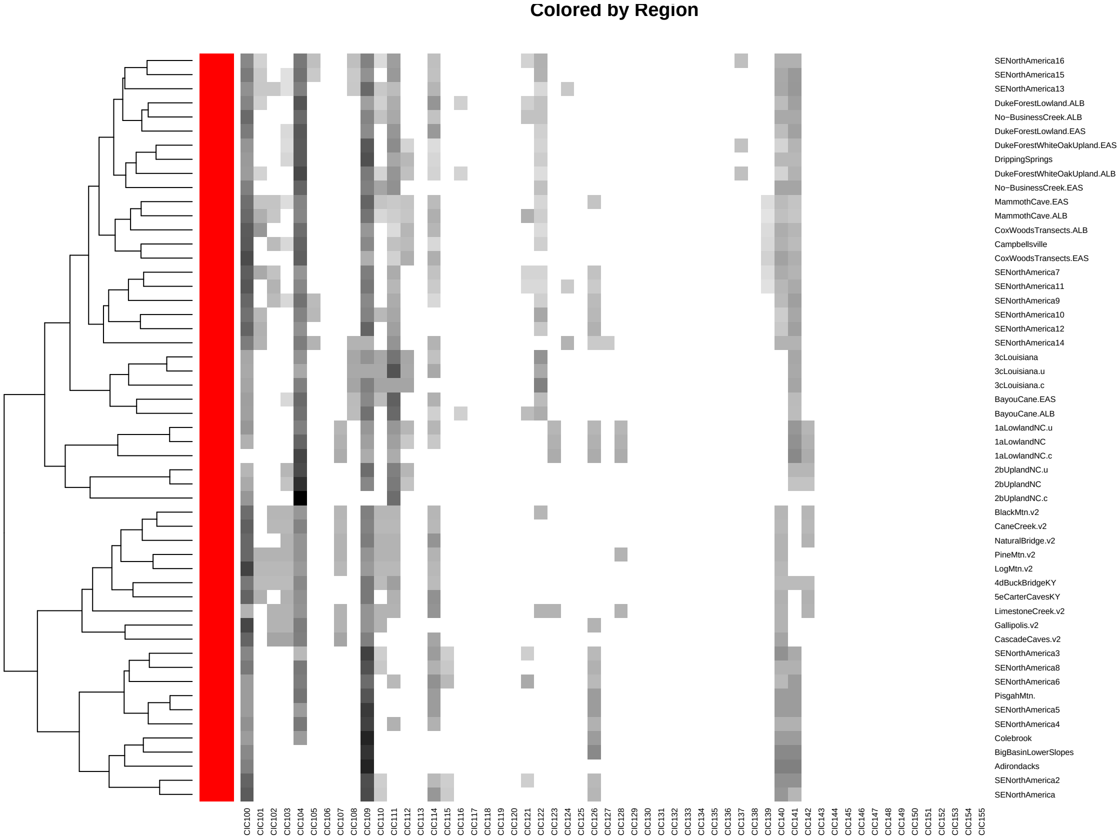


Figure 2.6: All floras from North America, showing how the precision of the classification decreases as more terminals are added, but that strong signals like the clustering of northeastern mixed deciduous forests (the five at the bottom of the figure) remain easily visible.

When floras from China are added as they are in figure 2.7, a very strong regional clustering is evident. Here the left colored bar shows the continent from which the flora was obtained. Red again indicates North America and pink China. Note that at this scale of analysis, the noise due to coding and publication bias is less influential than the larger distance between floras from different continents.

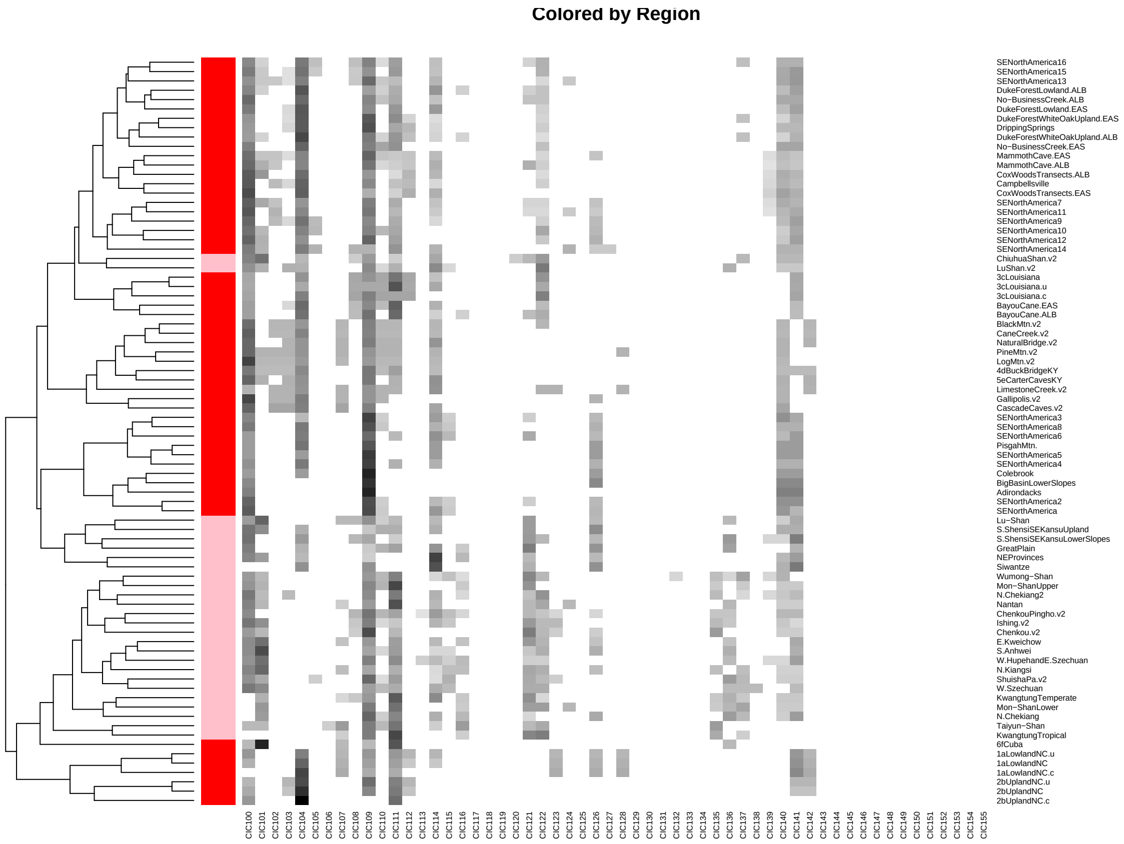


Figure 2.7: When floras from China are added; a very strong regional clustering is evident.

In order to illustrate the worst-case scenario, figure 2.8 shows as random sample of floras from the complete data matrix. A complex mixture of signals is apparent in which the sources of noise are mixed with real environmental signals. The colored bar at the left again indicates region: magenta for India; grey for Pakistan; green for Australia; blue for Latin America; orange for China; white for Africa; yellow for Pacific islands; black for North America, and cyan for Europe. Of particular interest is the Lebo fossil flora from the Paleocene Fort Union Formation of Montana Tamm (1989) which is nested within a number of modern floras from China. This supports the conventional wisdom that the best extant analogs to early Cenozoic floras from North America are found in southeast Asia because the current North American flora was decimated by the ice ages relatively recently.

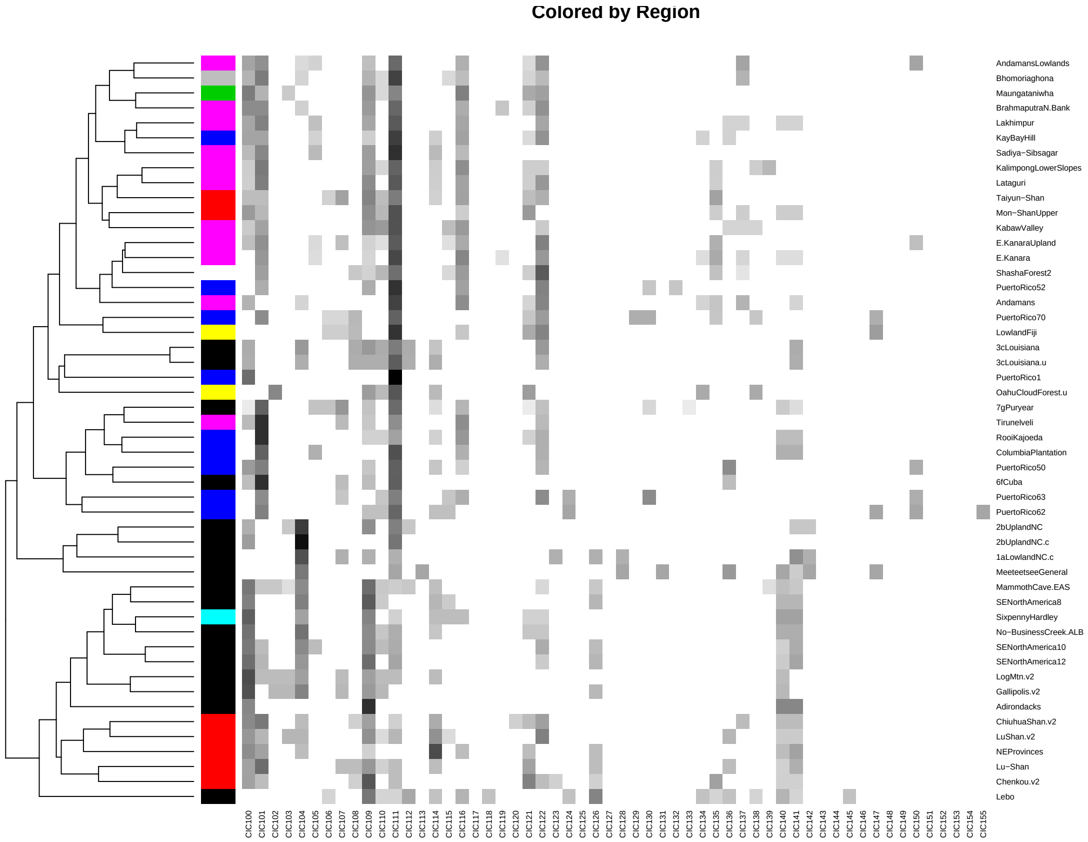


Figure 2.8: A random sample of floras from the complete data matrix. Note complex mixture of signals.

In contrast, to the primarily geographical clustering shown in the previous two figures, figure 2.9 shows a small group of floras from the Hawaiian islands which again were collected and coded homogenously from Mueller-Dombois and Fosberg (1998). Here instead of a geographical signal being evident, clustering is by canopy position: that is, for three floras each from four islands (one lowland, one cloud forest canopy and one cloud forest understory), the ecological types (canopy positions), not the islands, cluster together. This shows how even when an environmental signal is not obstructed by noise, it can signify different things: in figures 2.7 and 2.8, primarily geographic distance; in figure 2.9, primarily canopy position. This set of floras from Hawaii will be considered again in chapter 8.

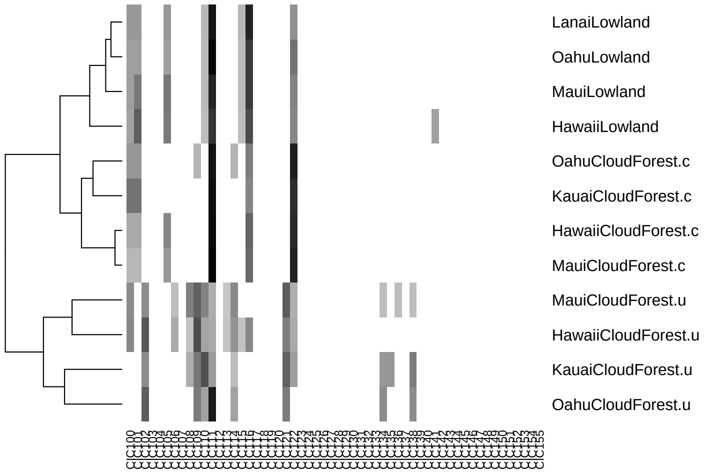


Figure 2.9: A small set of floras from the Hawaiian islands showing how leaf architecture sometimes carries information about canopy position rather than geography or climate.

2.4 Discussion

One thing that is not provided by the figures in the previous section is biologically interesting metadata. For instance, do floras cluster by temperature? The answer is that they do (even if these data do not tell us so). We know that geographic regions do cluster together both by temperature and by leaf architecture. Temperature is not a variable easily collected for these floras, nor would the figures in the previous section be the best way of identifying a gradient among floras with respect to temperature, even if temperature data had been collected. These examples were not intended to interpret the signal carried by leaf architecture, merely to show that the signal exists and is interpretable. Following chapters provide specific scenarios in which it is interpreted and applied to particular questions of biological interest.

2.5 Continued Work: The Standard Floral Paragraph (SFP)

A question remains: how to record and publish the sort of data examined in the previous section. Information about character states in systematic paleobotanical publications is published as character matrices and descriptions or as 'systematic paleontology'—standardized descriptive paragraphs that encode the relevant diagnostic information about each taxon described.

In this section I propose a form of standard paragraph for the publication and storage of the data collected and analyzed here. In contrast to systematic paleontology, it is vital that this information be easily machine-readable, but it is superficially more opaque than freer methods of describing floras. Like the descriptions of taxa in Airy Shaw (1973) or Mabberley (1997), however, the format is meant to be readable by people as well as computers.

There follows a paradigm and description of the fields in the format as well as a sample entry from Brown (1933). Although the data analyzed above have not yet been converted into this form, the form is presented here to indicate the direction future work will take. Scripts to automate the process of reading and writing to this format are provided in appendix B.

NAME_OF_FLORA. (2) Formation_binomial (3) Age (4) Organ(s)_preserved (P,W,L,R) (5) Preservation_type (O, I, M, C, P, R, A) (6) Number of OTUs (n/a/g/p/b/o) (7) Matrix (8) Facies (9) Current_location (Lat/Long/elev. or description). (10). Taxon_list (comma delimited). (11) References (comma delimited). (12) CICs (comma delimited). (13) Percentage_entire_of_angiosperm_leaves. (14) Average_size_of_angiosperm_leaves. (15) CLAMP (comma delimited). (16) Notes

(1) Name_of_flora
(2) Formation_binomial
(3) Age
(4) Organ(s)_preserved (P,W,L,R): Pollen/spore (microfloral disseminule), Wood, Leaf, Reproductive structure (seed, fruit, flower, cone, macrofloral disseminule)—list all that apply
(5) Preservation_type (O, I, M, C, P, R, A): Original material, Impression, Mold/cast, Compression/coalified original, Permineralization, chemical Replacement, Anatomical detail present—list all that apply
(6) Number of OTUs (n/a/g/p/b/o): All, angiosperm, gymnosperm, pteridophyte, bryophyte, other.
(7) Matrix
(8) Facies
(9) Current_location (Lat/Long/elev. or description).

(10) Taxon_list (comma delimited).
(11) References (comma delimited).
(12) CICs (comma delimited).
(13) Percentage_entire_of_angiosperm_leaves.
(14) Average_size_of_angiosperm_leaves.
(15) CLAMP scores (comma delimited).
(16) Notes

Where multiple references are given, the data they provide has been synthesized to produce the floral entry; * indicates that the entry does not come form the indicated reference or references and that the superseding authority is not given; a parenthesized entry after the asterisked field indicates the original (superseded) data. Thus, if Heer (1895) has described a flora from the Miocene, and Spicer (1996) has shown that the formation where Heer was working to be Jurassic, the entry for the age of the flora could be either:

'(3)Jurassic*...(9)Heer (1895)'
or
'(3) Jurassic*(Miocene)...(9)Heer (1895)'
or
'(3) Jurassic...(9)Heer (1895), Spicer (1996)'
or
'(3) Jurassic (Miocene)...(9)Heer (1895), Spicer (1996)'

EXAMPLE:

ASPEN. (2)Aspen Shale (3)UUAlb (4)L, R (5)I (6)26/23/3/0/0/0 (7)Volcaniclastic shales, siltstones, fine ss. (8)Fluvial (9)5 mi. NW of Kemmerer, WY, U.S.A. (10)Taxon_list (comma delimited). (11)Brown (1933) (12)CICS[0, 2, 0, 0, 6, 0, 4, 0, 0, 1, 0, 3, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0] (13)86 (14)Mesophyll (15)CLAMP[0.56, 0.32, 0.47, 0.50, 0.00 , 0.33, 0.78, 0.39, 0.00, 0.70, 0.30, 0.60, 0.50, 0.55, 0.55, 0.20, 0.57, 0.86, 0.87, 0.94, 0.27, 0.25, 0.20, 0.45, 0.56, 0.3, 0.80, 0.34, 0.23, 0.18, 0.22, 0.56] (16) Notes

2.6 Conclusions

Leaf ecophenetics is a new method for plant paleoecology. This chapter explains it and shows how despite several different sources of noise, an environmental or ecological signal is easily detectable in modern and fossil floras. The method produces a semi-quantitative description of each flora studied (based on the leaf forms found in the flora), classifies it with respect to other floras, and allows the raw data (the leaf forms that characterize a given flora) to be recovered from graphical representations of the floral spectra. Subsequent chapters provide examples of how this general method can be applied to biological question about relationships between and among floras distributed through time, across geographic space.

Methods of ecological reconstruction that rely on functional morphology are less subject to non-uniformitarian biases than those that rely on phylogenetic analysis. DiMichele et al. (2001) take an approach similar to the one suggested here: they describe a method for quantifying 'ecomorphospace' by coding particular taxa for a set of 22 'ecomorphic' characters and then clustering them mathematically using principle components analysis and an agglomerative clustering algorithm. They recover clusters that are roughly congruent with the traditional Linnaean classes, thereby supporting our expectation that in general the Linnaean classes correlate with broadly construed ecological niches. They then use this observation to support the macroevolutionary theory that early radiations 'fill' niche space in a comparatively short time, leaving subsequent evolution only certain niches to exploit under historical constraints, just as morphological evolution fixes the body-plan early in history and canalizes subsequent modifications. Although DiMichele et al. (2001) address the question of ecological patterns through evolutionary time, their paper is not intended as a general methodology, but rather as a way of answering a specific question about the ecological influences on plant evolution.

My approach is more pattern-oriented: instead of formulating a particular hypothesis and then finding a specific way to test it, I have developed a general way of displaying and publishing paleobotanical data so that fossil floras can be fit into the framework in which we analyze modern floras. It is known that particular communities are characterized by particular patterns of leaf architecture; here information on relationships among ecosystems is recovered from leaf architectural data in the way that the early plant synecologists like Warming and Raunkiaer used the 'physiognomic method of ecological characterization' to classify communities into formations (*i.e.* ecological characterization by growth form, not to be confused with leaf physiognomy. See Whittaker (1962)).

How does ecophenetics differ from other attempts to derive ecological information from fossil assemblages? First of all, leaf physiognomy is a method of climate reconstruction, not a step towards characterization of forest communities: it is essentially about estimating climatic parameters, not facilitating ecological pattern recognition. No doubt many of the patterns we hope to identify will be highly correlated with climatic parameters that are identified by leaf physiognomy, but the statistical tools used by CLAMP and related methods take complex and multivariate inputs and return a few variables. The approach used here is to reformat the complex relationships between ecosystems using leaf architecture as a proxy measurement so as to reveal hitherto unknown patterns in the history of plant evolution.

3 Leaf architectural patterns through time: An example from the Cretaceous/Tertiary boundary

The Cretaceous/Tertiary boundary extinction has long been considered one of the most important identifiable events in the course of Phanerozoic evolution. At times, the dramatic evidence for this has obscured the fact that any extinction event is selective and may not affect all groups of organisms in the same way. In this chapter, I examine a North American plant fossil database from the Mesozoic and Cenozoic eras in order to re-evaluate the evolutionary significance of the Cretaceous-Tertiary extinction on plants. A comparison of the leaf architectural profiles of fossil floras in each stage of the Cretaceous and epoch of the Cenozoic shows that the changes in leaf architecture at the Maastrictian/Paleocene boundary cannot be statistically distinguished from the population of changes at other boundaries. To the extent that patterns in leaf architecture reflect ecosystem structure, we can therefore conclude that despite the local species or morphotype extinctions that are known to have taken place at the boundary, the effect of the extinction on the structure of plant ecosystems was either minor or short-lived. Certainly, the extinction seems insignificant compared with the dramatic changes in leaf-architecture that accompanied the rise of angiosperms in the middle Cretaceous. This analysis also provides an example of the importance of time scales in the evaluation of macro-evolutionary pattern, and shows how the use of morphological categories instead of phylogenetic groups or simple diversity measures can produce rich and ecologically informative semi-quantitative proxy measurements of plant evolutionary patterns.[1]

3.1 Introduction

Since 1980, when Alvarez et al. (1980) proposed an extra-terrestrial impact as the extinction mechanism at the K/T boundary, questions about the significance of these extinctions for plant ecosystems have remained contentious, though there seems to be little question that the K/T extinctions dramatically affected the course of terrestrial vertebrate evolution. Long before the probable cause of the extinction was determined, the faunal discontinuity at the boundary was considered so important that it was used to identify major subdivisions of geological time. The floral record, however, does not show such a clear change at the K/T boundary. Did the K/T boundary event change the course of plant evolution, or did it merely cause some local species extinctions without affecting plant evolution in any lasting way? Were the K/T extinctions more or less influential than the rise

---

[1] Note on authorship: The material in this chapter has already appeared in print as Green and Hickey (2005). My co-author, Leo J. Hickey, collaborated with me on framing the research and provided editorial suggestions on the text; I performed all the data analyses and wrote the text.

of angiosperms in the middle Cretaceous? Opinions based on personal experience with the fossil record and on data sets of limited geographical and chronological extent have been numerous, but the difficulty of amassing a large set of comparable data on botanical macrofossil occurrences has restricted most data-based discussions to local dynamics and short-term effects.

In order to address these questions about the long-term effects of the K/T boundary extinction on plant evolution, I examined a database of fossil leaf occurrences through the Mesozoic and Cenozoic eras from an eco-morphological perspective. That is, I intentionally ignored the available taxonomic information (except the collection of specimens into operational taxonomic units, morphotypes, or species), instead examining only change in patterns of leaf architecture. This provides a more complex and ecologically informative view of change than most prior examinations of the floral record, which have focussed on species diversity as the sole response variable.

Consensus on the floral response to the K/T boundary event has varied widely during the past twenty years: one initial reaction (Hickey 1981:302) was that the North American macrobotanical record made it ‘unnecessary for the paleobotanist to invoke a catastrophic or unusual mechanism to explain the observed patterns of localized extinctions’. Subsequently, however, palynological data, better correlation, and more thorough sampling at good boundary sections modified this response to the point where Johnson and Hickey (1990), discussing a composite stratigraphic section in Marmarth, North Dakota, wrote that ‘results of this analysis of the terrestrial plant record are compatible with the hypothesis of a biotic crisis caused by extraterrestrial impact at the end of the Cretaceous’ (Johnson and Hickey 1990:433). The most recent discussions seem to reinforce this point of view; a new analysis of the Hell Creek/Fort Union boundary in North Dakota (Wilf and Johnson 2004), which is clearly the best-studied plant macrofossil boundary section, gives a local species/morphotype extinction rate of about 57%, and stresses the ‘sharp shift’ in floristic composition at the boundary.

But the question remains: have we in the past twenty years actually modified our ideas about the plant fossil record, or have we just agreed that other evidence for the impact is conclusive and then found a correlative signal in the macrobotanical record? Note that there is a substantive difference between on the one hand identifying such a signal or even suggesting that the extinction rate at the boundary is statistically different from background, and on the other hand arguing that such extinctions played a significant role in plant evolution. So it remains an open question: were plant ecosystems doing anything substantively different in the Paleocene from what they had been doing in the Cretaceous?

Such broad questions about plant evolutionary history have been asked for over a century (see, for example Ward 1883–1884), but our knowledge of the evolution of plant ecosystems is limited by the difficulty of collecting unbiased data on plant fossils distributed through time. Obviously any sample of past vegetation will be affected by taphonomy, by irregular sampling, by varying volumes of preserved sediment, by different intensity of research, and by differing taxonomic practices and the incomparability of modern and fossil taxa. Nevertheless, the few instances in which diachronic data on plant diversity have been assembled have given us some semi-quantitative support for assumptions about historical events like the rise of angiosperms and the K/T boundary event. In particular, the data set initially described by Knoll et al. (1979) and subsequently published as species diversity curves in Niklas et al. (1985) has been nearly as influential in paleobotany as the similar data on marine invertebrate diversity collected by Raup and Sepkoski (Raup 1972; Sepkoski et al. 1981; Sepkoski 2002)) have been in invertebrate paleontology. The picture of changing plant diversity through time provided by Niklas et al. (1985) has been independently substantiated in part by Lidgard and Crane (1988), but to the best of our knowledge, few subsequent attempts to elaborate on this picture have been particularly influential.

An eco-morphological approach, in which patterns of change in leaf architecture are treated as measures of ecological change, has the potential to tell a more detailed story. In this chapter, I examine the question of ecosystem continuity at the K/T boundary by looking at changes in leaf architectural profiles of fossil floras since the beginning of the Cretaceous as represented in the *Compendium Index of North American Mesozoic and Cenozoic Type Fossil Plants* (henceforth *Compendium Index*), a database that is more-or-less independent of and complementary to the diversity data of Niklas et al. (1985).

This is not meant either as a review of literature on the K/T boundary generally, or as a justification of the ecological interpretation of leaf architectural patterns. Instead, like Raup and Sepkoki's work on the marine invertebrate record, it is intended to identify patterns and correlations in the known fossil record, and interpret them based on the assumption that such patterns in the fossil record have evolutionary significance.

3.2 Data

The data from the *Compendium Index* that are considered here reside in a card index and computer files at the Peabody Museum, Yale University, with entries for published descriptions of fossil plant species from the Mesozoic and Cenozoic Eras. Each entry consists of the published illustration and description of the fossil species and a citation for the source from which the illustration and description were obtained. The intent of the *Compendium Index* has been to provide full descriptions of all North American fossil species, and although it is far from complete, it is probably at the current time the most comprehensive single reference in North America for identification of Mesozoic and

Cenozoic leaf fossils. Erling Dorf began compiling the *Compendium Index* at Princeton University in 1937 as an aid to taxonomic identification, and by 1940 he and his staff had amassed cards for some 4500 species from 126 references (Dorf 1940). In 1984, after Dorf's death, the *Compendium Index* was transferred to Yale University where it has since been maintained and augmented at the Peabody Museum.

Currently, there are about 9800 species or morphotypes represented from 233 references, and the list continues to grow at a rate of about 25 new references per year including both current publications and gaps in the coverage of the older literature. The intent has been to enter all newly published type material as well as important emendations and range extensions of species that are already represented in the catalog (that is, additional cards are added for species that have been revised or are reported from stages, localities, or formations from which they had not previously been known.) While the temporal resolution of each card depends on the publication from which it was obtained, the minimum resolution in the *Compendium Index* is to epoch in the Triassic, Jurassic, and Cenozoic, and to age in the Cretaceous. Spatial resolution also varies by source and for the most part has not yet been electronically recorded. Beginning in the 1980s, portions of the *Compendium Index* were entered into a computer database (originally dBase; the files are currently maintained in FileMaker Pro) allowing electronic manipulation of some of the data, and an electronic version of the *Compendium Index* was recently released as Hickey et al. (2006).

Partly because of the interests of those responsible for maintaining the *Compendium Index*, partly because identification of angiosperm leaves remains one of the most difficult and potentially rewarding pursuits for the systematic paleobotanist, there has been a distinct bias towards including leaves at the expense of other organs (there is only one category, for instance, for angiosperm wood and over fifty for angiosperm leaves). In addition, as in any taxonomic database, different researchers are predisposed to be 'splitters' or 'lumpers', thereby unnaturally proliferating or impoverishing the numbers of species recorded from a given locality. Like any record of fossil data, the *Compendium Index* is also affected by taphonomic sorting during fossilization: because the most favourable conditions for preservation (high sedimentation rates and anoxic sediments) occur in lowland floodplain facies, there is an enormous bias in the fossil record in favor of the plants that grow in those habitats. Moreover, mechanically strong and well-vascularized leaves, such as those produced by canopy trees, preserve much better than the soft, weakly veined leaves generally produced by herbs (Behrenmeyer and Hook 1992). Therefore the fossil record is largely a record of the woody component of lowland forests, though occasional catastrophically buried floras provide a more complete picture of the vegetation at particular localities Wing et al. (1993). Such extraordinarily well preserved floras are rare, however, so if we are interested in producing a relatively continuous picture of variation in time and space, we cannot limit our analysis to these fossil Lagerstätten.

Unlike the *Index of Generic Names of Fossil Plants* (Andrews 1970) and its antecedents (Lamotte 1952; Knowlton 1919) which were based on the *Compendium Index of Paleobotany* at the Smithsonian Institution, the Princeton/Yale *Compendium Index* provides an illustration and description for each fossil as well as its name and citation. Thus it can perhaps better be compared with regional monographs (Berry 1916), illustrated museum catalogs (Steward 1894; Stopes 1913; Reid and Chandler 1926, 1933; Chandler 1961)), or the *Traitè de Palæobotanique* (Boreau 1964–). Since the failure of the *Traitè* in the 1970s, when faced with the taxonomic difficulties of the plant fossil record in general and with fossil angiosperm organs in particular (Collinson et al. 1993), the *Compendium Index* remains the best approximation of a comprehensive paleobotanical reference work analogous to the *Treatise on Invertebrate Paleontology* (Moore et al. 1952–). New resources that could potentially supercede the *Compendium Index*, and with which we hope the *Compendium Index* will ultimately be integrated, include the Paleobiology Database (Paleobiology Database) and MorphoBank (O'Leary and Kaufman 2007).

Since we might like to follow Raup and Sepkoski's lead (Raup 1972; Sepkoski 2002) in using a reference work intended for identification (in their case the *Treatise on Invertebrate Paleontology*) as a record of macro-evolutionary change, a possible approach would have been to plot family or order diversities through time in the *Compendium Index* data. This, however, is impossible: if the *Compendium Index* were arranged taxonomically as the *Treatise* is, over half of all described species would fall into a single category—Class Angiospermae, Incertae Sedis—because they cannot be firmly attributed to any modern order. Therefore, from the outset, the *Compendium Index* was organized into a set of morphological categories to facilitate identification of fossils, based on aspects of gross morphology like overall shape and pattern of veination rather than on phylogenetic principles. The current index has 118 categories, which have been designed for further subdivision as groups are studied in more detail. These Compendium Index Categories, or CICs, which are numbered with three-digit numbers between 100 and 990, were originally illustrated and described in Ash et al. (1999), and are listed in a slightly revised form, as they appear in Green and Hickey (2005), in appendix C. They are based on characters like organ type (leaf, axis, reproductive structure), petiolar attachment (compound, central or marginal), primary vein course (pinnate or palmate), tooth presence, and tooth shape (serrate, dentate, or crenate). They are categories of convenience, designed for quick, simple, morphological classification using readily observed features and applicable to fossil plants in all degrees of preservation and to illustrations of variable quality. The theoretical issues surrounding ecological interpretation of data from such a classification system will be dealt with in detail in a future publication; here we are primarily concerned with identifying morphological patterns through time, not necessarily with relating particular morphological attributes to ecological variables. It should be noted, however, that some of the patterns in leaf architecture that are reflected in the CICs have been shown to correlate strongly with ecological and environmental variables (Bailey and Sinnott 1915; Wolfe 1993).

The data that are currently available electronically reside in a computer database file (.dbf) from which the actual data matrices used in the following graphs (see appendix A) were extracted. Our analyses are based on the recorded numbers of species (or morphotypes or operational taxonomic units) per time-division in each CIC. In some ways, this is roughly analogous to the initial, coarsely resolved data on marine invertebrate diversity collected by Raup (1972), but because the *Compendium Index* is organized by morphological category rather than by super-specific Linnaean taxa, it does not suffer from the incomparability of taxonomic ranks in different groups. That is to say, when we compare counts in two different CICs in the *Compendium Index,* we are not looking at a measure of diversity (which is sensitive to the degree of splitting or lumping that has been applied) but at the proportion of operational taxonomic units that have a particular morphology, which is comparatively insensitive to such biases. In other words, some taxonomists may be ‘splitters’ and others ‘lumpers’, but few systematically split ovate leaves and lump obovate leaves. In addition, the organization of the *Compendium Index* by morphological category instead of by higher taxon provides us with the ability to examine it for ecological rather than phylogenetic patterns.

Using shell commands and the open-source application *R* (R Development Core Team 2005), we extracted and tabulated the number of described morphospecies from each time period in each Compendium Index Category. At a coarse taxonomic level (for example Linnaean class), leaf shapes uniquely identify phylogenetic groups, so we can equate groups of leaf shapes with clades or Linnaean taxa (Dimichele and others 2001). At higher taxonomic resolutions, however, this assumption breaks down. Therefore to examine patterns through time in any more detail than by class (or at best subclass), we need to think in terms of architectural groups that may share functional and mophological attributes but are not necessarily genetically related.

Statistical manipulation at this stage has intentionally been restricted to an exploratory level because there are issues of sampling and systematic bias that must be fully discussed before applying specific confirmatory tests. While we are exploring the signal-to-noise ratio in the data and looking at first-order trends and patterns, it seems relatively unproductive to depend on techniques of variable reduction and parametric testing, though we do rely on relative correlations. It is easy to find statistically significant effects in the fossil record; harder to determine what they signify. As the data are of variable quality and unevenly distributed though time and space, we sometimes would like to draw conclusions based on very small samples, while at other times very large samples tell us nothing of biological interest. For exploratory analysis of this sort, graphical display of the data has proven to be the most effective tool (Tukey 1977; Tufte 1983). Despite the inevitable presence of noise in the data, the patterns we can now identify are robust enough that we are confident they will survive the cleaning up of the signal that is currently in progress.

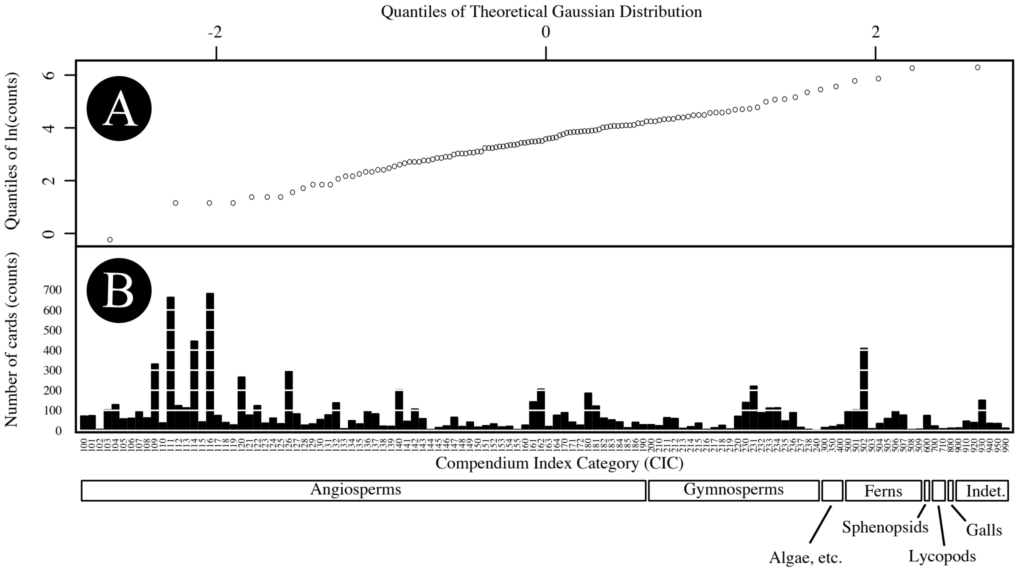


Figure 3.1: (A) Log-normal quantile-quantile plot of the counts in the *Compendium Index*. This is a scatter plot of the quantiles of the log-transformed counts plotted against the corresponding quantiles of the theoretical Gaussian distribution. Therefore a straight line indicates log-normal distribution of the count statistics. (B) Bar chart showing the actual counts in each Compendium Index Category (CIC). See appendix C for a description of the categories.

Figure 3.1 shows the overall frequency distribution of species in the *Compendium Index* by CIC, which is roughly log-normal, as is expected from a set of counts. When all the entries in the *Compendium Index* are sorted by morphological category, the vector of counts obtained ranges from almost 700 taxa with simple, entire, ovate, pinnate leaves with regular secondaries (category 116) to only a few taxa with peltate, lobed leaves (category 155). This vector is shown as a profile in figure 3.1B, with the actual number of species in each category plotted as a vertical black bar, and in figure 3.1A the dotted curve is a plot of the counts plotted on a log-log scale against a theoretical

Gaussian probability distribution function. In such a log-normal quantile-quantile plot, a variable with a perfectly log-normal distribution will be distributed along a straight diagonal line.

Compendium Index records include representatives of approximately 840 floras, almost exclusively from continental North America, but with a few additions from other parts of the world. By the nature of the geographic distribution of geological strata, any one time period is likely to be represented only by a small geographic area. For instance, the Turonian stage of the Cretaceous is represented almost exclusively by floras from Greenland. The *Compendium Index* does, however, include most the 'classic' American floras, including those described by Lequereux, Berrry, Knowlton, etc., so whatever biases are introduced by the patchy nature of the fossil record, the analysis of the *Compendium Index* should reflect the fossil record insofar as it has been cataloged by some two centuries of paleobotanical investigation.

Table 1: Approximate number of floras represented in the Compendium Index in each time period.

TIME PERIOD	APPROXIMATE NUMBER OF FLORAS	NUMBER OF FLORAS WITH >20 SPECIES
Pleistocene	1	0
Pliocene	34	3
Miocene	106	12
Oligocene	31	6
Eocene	193	10
Paleocene	106	10
Maastrchtian	65	7
Campanian	36	2
Santonian	33	1
Coniacian	7	0
Turonian	8	0
Cenomanian	26	6
Albian	50	2
Aptian	47	1
Barremian	36	0
Neocomian	7	0
Malm	6	0
Lias	7	0
Dogger	10	0
Triassic	30	0
TOTAL	839	60

3.3 Results

The 118 morphological categories into which the *Compendium Index* species records are sorted can also be lumped into aggregate groups which are synonymous with the higher Linnaean taxa used by Niklas et al. (1985). That is, we can add together all the angiosperm counts, all the conifer counts, etc., and plot the resulting sums as mountain charts through time, as I have done in figure

3.2. These aggregate groups are plotted alongside a redrawn version of the familiar plot from (Niklas et al. 1985:112). Note, however, that the data from Niklas et al. (figure 3.2A) are absolute numbers of described species in each time interval, while the *Compendium Index* data (figure 3.2B and 3.2C) are shown as the proportion of described species (or operational taxonomic units) in each morphological category in each time interval. The *Compendium Index* data are plotted twice: once averaged by epoch (figure 3.2B) and once at the finest chronological resolution now available (epoch in the Jurassic and Cenozoic, age in the Cretaceous, figure 3.2C). I am in the process of refining the chronological sampling by reference to the stratigraphic information in the *Compendium Index*. On the extreme right (figure 3.2D), are shown the absolute numbers of cards in the *Compendium Index* (which are roughly proportional to the numbers of described species). At the top (figure 3.2E) are estimates of modern species abundances from Raven et al. (1999). The chronology used is the Geological Society of America 1999 Geological Time Scale (Palmer and Geissman 1999) and the values are plotted at the midpoints of chronostratigraphic divisions.

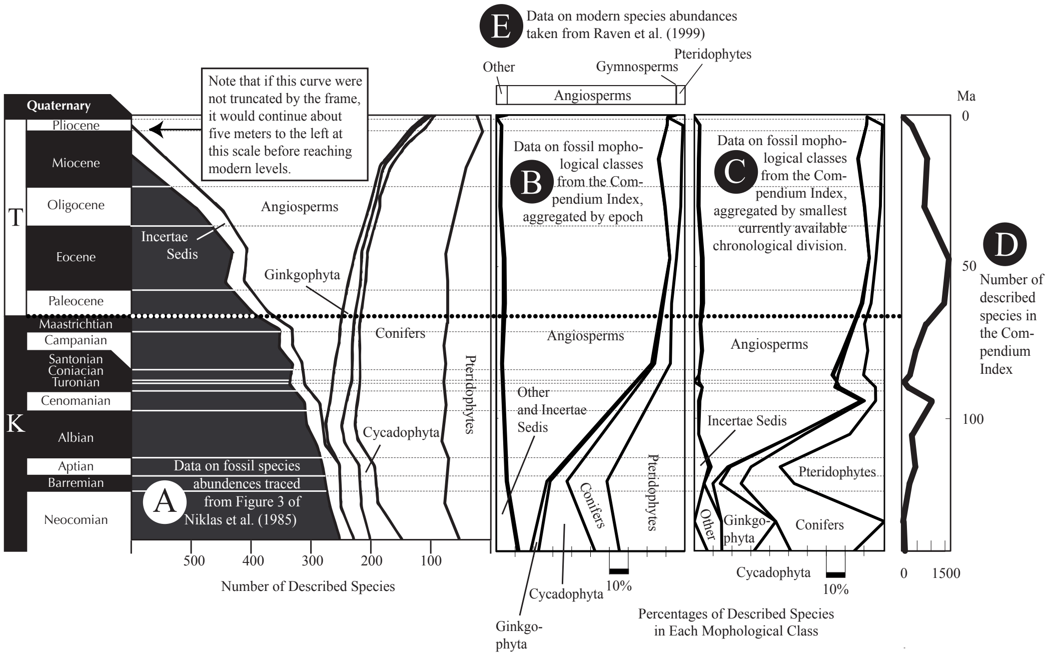


Figure 3.2: Juxtaposition of the curves from Niklas et al. (1985) giving absolute species diversities through time (A) with similar mountain plots showing the numbers of entries in the *Compendium Index* (B, C) in the morphological groups corresponding to each higher Linnaean taxon. Also shown are the absolute numbers of entries in the *Compendium Index* in each time division (D) and estimates of modern species diversities from Raven et al. (1999) (E). Note the smoothing artifacts like the implication in 2B that there are a large number of Neocomian angiosperms: this arises from lumping all Lower Cretaceous counts together in a single chronological bin.

Regardless of whether one examines the absolute species diversity data from Niklas et al. (1985) in figure 3.2A, or the proportional morphological data in 3.2B of the figure, it is apparent that there is no discontinuity at the K/T boundary. The picture obtained of floral change through time is uncannily similar to that provided by plate 56 in Ward (1883–1884), even to the extent of duplicating a substantial drop in angiosperm diversity in the Turonian. Because sample sizes are so low, we are not confident that this represents a true signal in our data, but the discovery of the same pattern in Ward's data raises the possibility that is is real.

The *Compendium Index* data do confirm a broad picture of relative floral stability since the Campanian (logistic growth curve) that has already been proposed (Lidgard and Crane 1988), but many of the criticisms of diversity-curve data are not applicable to our proportional morphology curves: as discussed above, incomparability of higher taxa and biases caused by differing intensities of study are automatically normalized by considering a proportional metric. Note also that in these data, in contrast to those of Niklas et al. (1985), the expansion of angiosperms seems precipitous (essentially taking place from the Albian to the Santonian stages of the Cretaceous), rather than happening slowly over the course of the Late Cretaceous and Tertiary. Because proportional counts of taxa in morphological classes are likely to reveal ecological dynamics while ignoring species turnover, we can conclude that the ecological expansion of angiosperms was indeed more precipitous than their rate of speciation: that is we see the broad-leaf angiosperms appear, restructuring ecosystems in an explosive radiation, but because we are looking at morphological categories we do not then notice the continued taxonomic turnover that we know occurred during the Tertiary. To use Gould's terminology (Gould 1991), angiosperms attained modern *disparity* levels by the end of the Santonian, but may have taken the rest of the Cretaceous and all of the Tertiary to attain modern *diversity.* Equivalently, it could be argued that there were as many species with angiosperm leaves (proportional to non-angiosperms) in, for instance, the Campanian as there were in the Miocene, but they were different species with similar leaves, and the replacement of one species by another with similar leaves is invisible in our record of eco-morphological change (figure 3.2B and 3.2C) while it is apparent in a diversity curve (figure 3.2A).

An intensification of this effect may come from the over-speciation of a number of early Late Cretaceous floras by zealous splitters, but because we are dealing with proportions, this effect could only be caused by inconsistent splitting of some morphologies more than others. (Note that this sort of inconsistant splitting is possible, as when a group of people working on a particular taxonomic group evolve a different implicit notion of the amount of morphological variation allowable at a given taxonomic rank. Reduction to proportions will minimize this error, but it will not necessarily eliminate it; therefore the resulting patterns will be most reliable when they are based on comparison of like with like as in the case of the angiosperm subset of the *Compendium Index* analyzed below. Associated with the rise of angiosperms was a proportional decrease in the numbers of cycadophytes, ginkgophytes, and pteridophytes, while the conifers remained relatively unaffected. This agrees with the results of Lidgard and Crane (1988) and supports the contention of Taylor and Hickey (1992, 1996) that the the earliest angiosperms occurred in disturbed habitats.

Lumping the morphological categories until they map onto Linnaean classes provides an overview of plant evolution through time and does not conflict with the story told by Niklas et al. (1985) and Lidgard and Crane (1988). It provides, however, little additional detail. To describe the K/T boundary dynamics in more detail, we can examine the categories (in this case just the angiosperm leaf categories) individually instead of lumping them together so that they correspond to higher taxa. Looking only at the angiosperm leaf categories (CICs 100–155) in each time interval (as is done in figures 3.3, 3.4, and 3.5), we can then set about quantifying the degree of discontinuity in eco-morphological dominance at the K/T boundary. First of all, however, we need a way of measuring how different the leaf architectural profiles of different geological time periods are from each other.

One way to address this question is by producing a bivariate plot with associated correlation statistics. Figure 3.3 provides an example of such a plot of the counts in each CIC of the species in the *Compendium Index* for the Maastrictian Age and the Paleocene Epoch, that is the stage transition that corresponds to the K/T boundary. The CIC numerical codes and cartoons of the leaf shapes they represent are plotted alongside the points in order to show which categories are represented. The points lying above the diagonal are more abundant in Paleocene assemblages while those lying below the line are more abundant in Maastrictian floras. The regression line, while based on a relatively small number of points, has a slope significantly different from zero, and this significance remains even if the two or three most influential points are omitted.

The correlation provides a rough measure of how similar the two sets of counts are. This measurement is difficult to evaluate except in the context of comparative measurements, so in order to determine whether the correlation between the Maastrictian and Paleocene is unusual, we must look at it in the context of other correlations between successive time periods.

Figure 3.4A shows this comparison: the solid line shows the ordinary parametric autocorrelation (Pearson's product-moment correlation, r, of each time period with its neighbors), plotted at the interval boundaries. The fine dotted line gives a 95% one sided confidence interval for these correlations. I also calculated two non-parametric correlation coefficients, Spearman's ρ, and Kendall's τ, shown by the two dashed lines. In general, these show the same patterns as the parametric correlation coefficient, their lower statistical power does not matter in this comparative context,

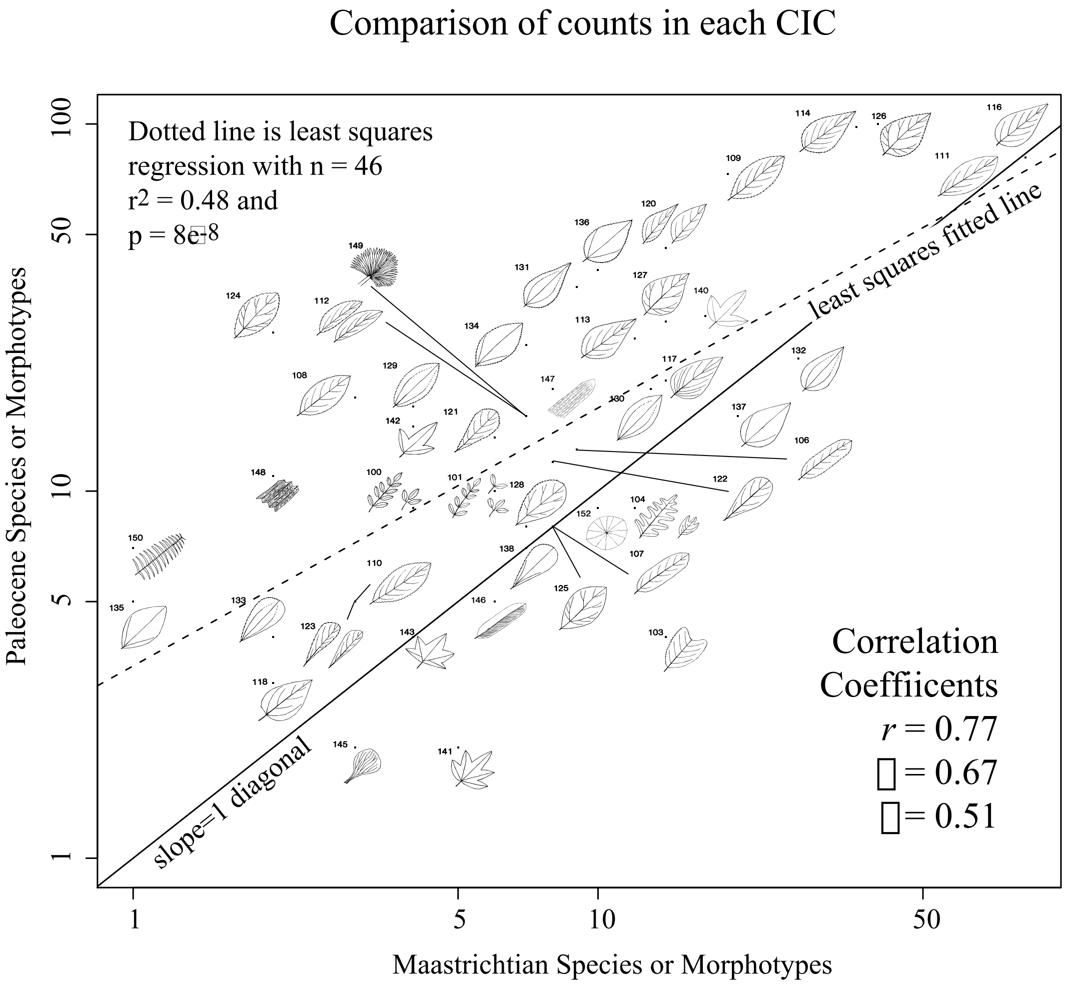


Figure 3.3: Comparison of Maastrichtian and Paleocene floras. In addition to the parametric correlation, r (formally, Pearson's product-moment correlation coefficient), I have calculated Spearman's rank-order correlation coefficient, ρ, and Kendall's rank-order correlation coefficient, τ. For further details of the algorithms used see the man page for `cor.test()` in *R*, and references therein. This plot shows the type of raw data from which the correlation statistics that are discussed below were obtained.

and they depend only on the rank-order similarity of their arguments, so they are more likely to be applicable to noisy and abnormally distributed data. Quantitative similarity measures like the

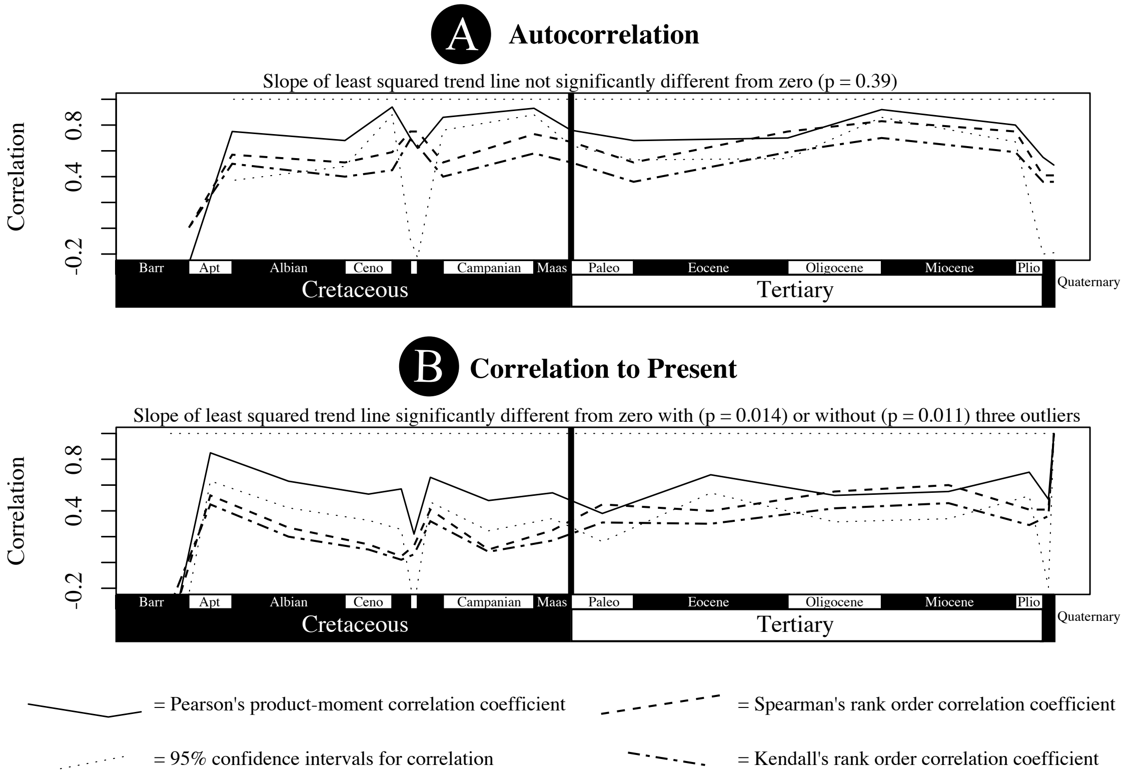


Figure 3.4: Autocorrelation (A) and correlation with the present (B) of each geological time period, using all three measures of correlation. The fine dotted line gives a 95% one sided confidence interval for Pearson's correlation coefficient. Note in both cases the small-sample effect giving an anomaly in the Coniacian.

Horn-Morisita index which defines distance d between vectors x and y each of length i as: $d(x,y) = \frac{2\sum_i x_i y_i}{\lambda(x)+\lambda(y)} \sum_i x_i \sum_i y_i$, where $\lambda(x) = \frac{\sum_i x_i^2}{(\sum_i x_i)^2}$ also show similar patterns but have not been plotted because it is not clear whether the ecological rational for using such measures applies in the case of percentage counts in morphological categories.

Low correlations imply times of leaf architectural change; high values suggest floral continuity. The trend line has no significant slope. From the figure, it is apparent that the Maastrictian/Paleocene boundary does not stand out as a low point, so we can conclude that it was not a time of dramatic change in leaf architecture.

In figure 3.4B the correlation of each time period with the present is plotted in the same way that the autocorrelations were plotted in figure 3.4A; again the difference between the Maastrictian and Paleocene does not stand out (if it did it would look like the vertical riser of a step), but in this case there is a somewhat significant trend, which somewhat surprisingly remains even in the absence of three outlying points. This provides an empirical test of the uniformitarian assumption that modern floras provide unbiased estimators of the distribution of architectural characters in past floras, which underlies the practice of estimating paleotemperatures from leaf morphology using leaf physiognomy (Wolfe 1993, 1995; Wilf 1997). The negative trend in modern-fossil correlations from left to right contradicts this assumption.

The argument that the Maastrictian/Paleocene (K/T) boundary does not stand out can also be made statistically explicit by plotting the empirical distributions of the correlation coefficients as histograms or smoothed densities as is done in figure 3.5 for all three correlation coefficients. In each case the position of the K/T boundary in the distribution is shown. For each correlation coefficient, the exact probability that the correlation at the Maastrictian/Paleocene transition is significantly greater than the mean age transition correlation is dependent on the exact shape of the distribution and since there are only eleven points the shapes of the distributions are poorly constrained. Nevertheless in all three cases it is clear that the Maastrictian/Paleocene boundary is not a low outlier.

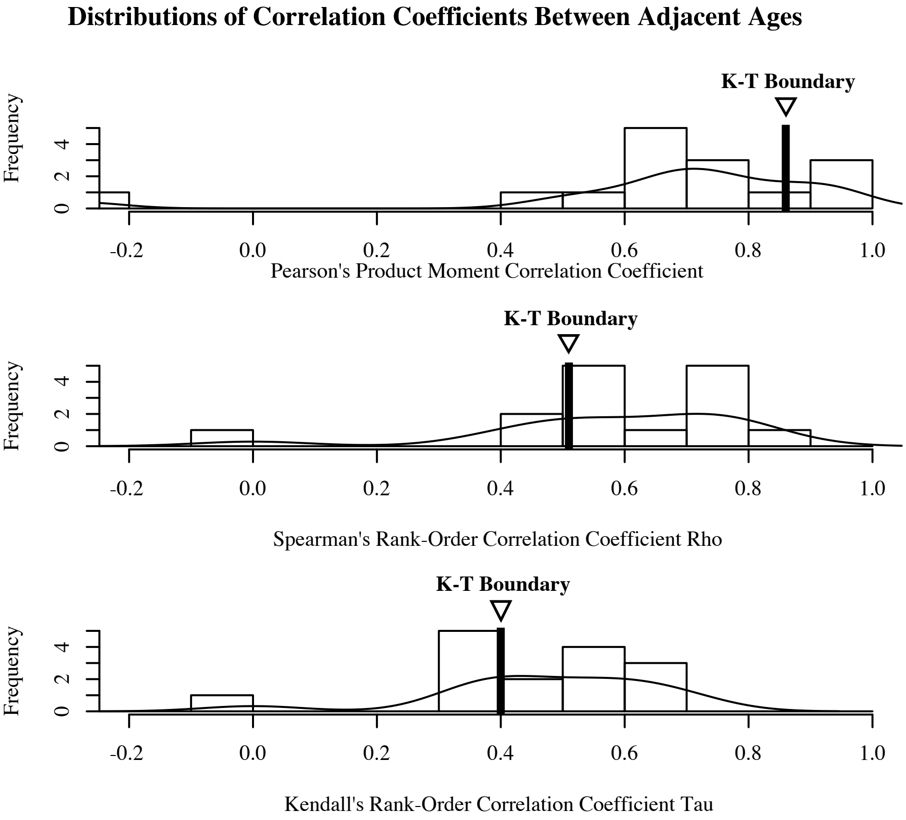


Figure 3.5: Plots of the empirical distributions of the three correlation coefficients (both as histograms and as smoothed densities) make explicit what can also be seen from the previous figures: that the Maastrictian/Paleocene (K/T) boundary is far from being a low outlier.

3.4 Discussion

Both the smooth curves across the Cretaceous-Tertiary boundary in figure 3.2 and the similarity of the Maastrichtian and Paleocene floral profiles suggest that the K/T boundary event had little lasting effect on the evolution of angiosperm leaf architecture. Qualitative appreciation of this point initially led a number of paleobotanists to doubt the extinction mechanism proposed by Alvarez et al. (1980) because of the apparent continuity of the plant fossil record examined at a resolution of geological stage (Clemens et al. 1981). Examined at an increased stratigraphic resolution, as it was by Wolfe and Upchurch (1986, 1987); Johnson and Hickey (1990), and Wilf and Johnson (2004), the K/T boundary does indeed show a concentration of last appearances, though it is still not clear that this concentration rises above a base extinction rate to a statistically significant extent. On the other hand, as we decrease the chronologic resolution, we effectively smooth the time-series with a low-pass filter. Comparing figure 3.2C with 3.2B, I accomplished an appropriate degree of smoothing by increasing the temporal bin size, but the same results would have been achieved by a moving average filter that replaced each value by the arithmetic mean of it and its neighbors. In this case, the higher frequency dynamics that we avoid by discussing figure 3.2B in preference to 3.2C seem to be due to small-sample effects. It is common for the quality of paleontological data to be too low and irregular for routine application of statistical time series tools like autoregression and spectral analysis.

It is worth examining in some detail, however, the apparent conflict between our data, which show no significant changes at the K/T boundary and the reported 57% extinction of morphotypes reported by Wilf and Johnson (2004). Two (non-exclusive) explanations for this difference present themselves: first of all, it could be entirely an effect of time (or space) averaging: that is because all our Paleocene data from North America are lumped together, we miss the dramatic changes that appear in the beginning of the Paleocene because we can only see the overall epochal average. Secondly, an examination of morphological attributes (which is in this context equivalent to an examination of proportional representation in architectural categories) may reveal ecological stability while a measure of species or morphotype diversity may show population level fluctuations. As Wilf and Johnson (2004) point out, even 57% of morphotypes is not a particularly large extinction compared with 95% of marine genera at the Permo-Triassic boundary or 100% of dinosaurs at the end of the Cretaceous.

In order to see whether these differences were a result of looking purely at continent and epoch scale averages, I collected the CIC attributions for 286 of the Hell Creek/Fort Union boundary section morphotypes from Johnson (1989) though I could not obtain the appropriate morphotype descriptions to recode the most recent data from Wilf and Johnson (2004). In figure 3.5 are plotted the profiles of four biozones defined by Johnson and Hickey (1990), three from the Hell Creek Formation (HC 1–3), immediately below the K/T boundary, and one from the Fort Union Formation (FU1), immediately above the boundary.

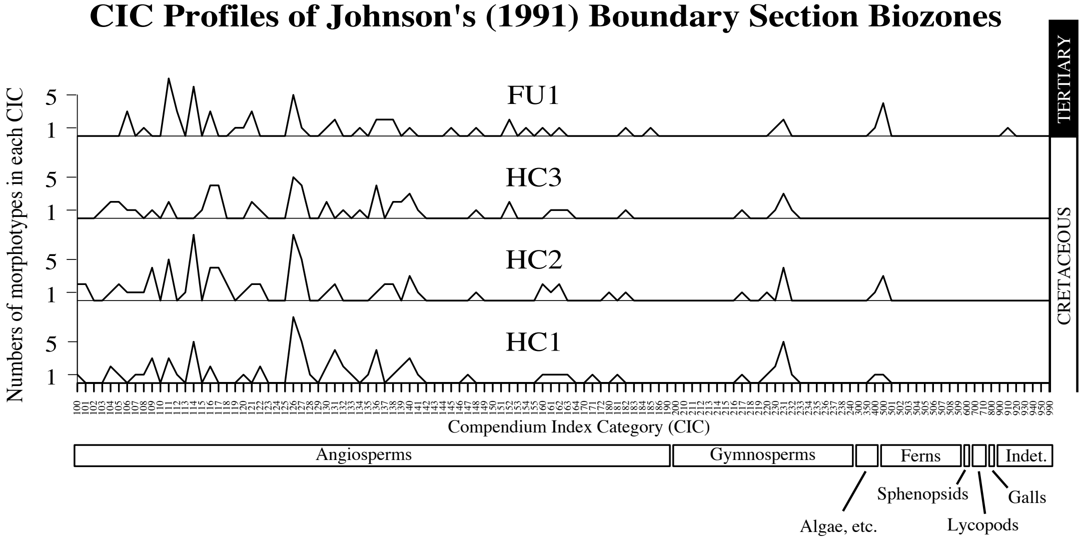


Figure 3.6: Leaf architectural profiles for each of the four biozones defined by Johnson and Hickey (1990): Hell Creek (HC) 1–3 and Fort Union (FU) 1. In this representation, there seems to be no dramatic change between HC3 and FU1, as one would expect if the K/T boundary extinction had had significant ecological effects. The FU1 profile does provide a slight echo of the fern spore spike following the K/T boundary, but otherwise there is no convincing indication of a change in leaf architecture across the boundary.

The profiles of the biozones do not seem to look radically different from each other and certainly the flora of the Fort Union biozone immediately after the K/T boundary does not seem to represent a dramatic break from the earlier Hell Creek floras. With only four points in our little time series, it is difficult to test this statistically, but again correlations among the zones give a certain ability to measure similarities of the counts in each architectural category across the boundary.

Figure 3.7 shows pairwise plots of the counts from each of the biozones against the counts from each other zone. The correlations amongst all of them are positive, but the weakest of the correlations is indeed between FU1 and HC3. When empty and singleton categories are eliminated,

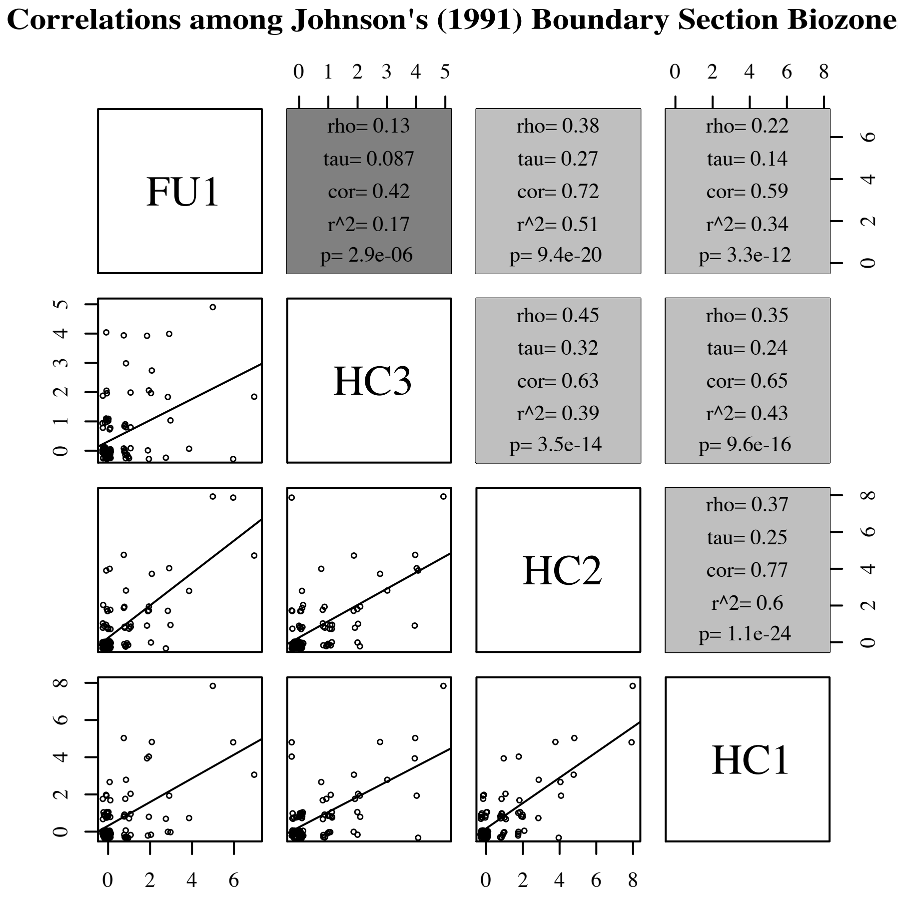
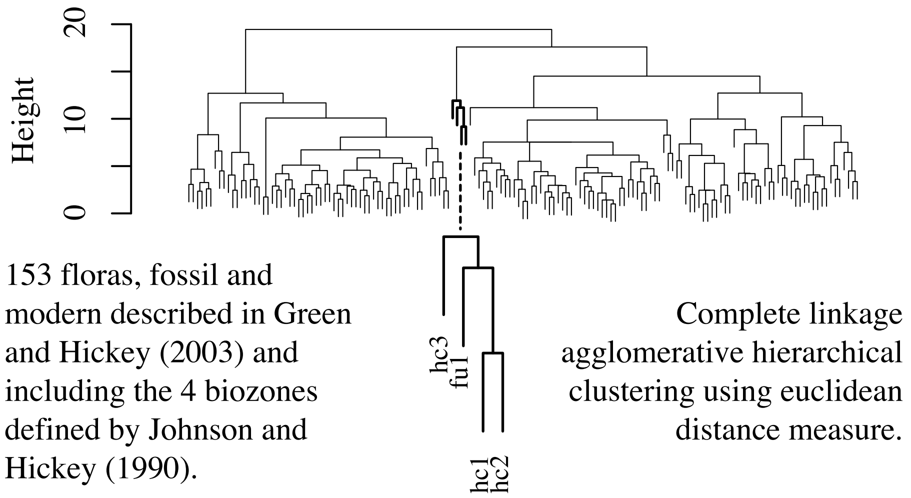


Figure 3.7: Pairwise plots of the same four biozones whose profiles are shown in figure 3.6. Note that according to the correlation statisitics, the relationship between the Fort Union flora and the floras of the three Hell Creek biozones is slightly more distant than the relationships among the Hell Creek biozones. This supports the contention of Wilf and Johnson (2004) that the K/T boundary was the most significant modification of the flora during the 2–3 million year period represented by this composite section. Note that for this plot the data have been jittered so as to reduce overplotting and zeros are left in the data (which will naturally increase the apparent significance of the regression).

the HC3:FU1 relationship loses its significance at the 5% level, while the other pairwise comparisons remain significant by at least some measure. Note that this is not a robust statistical result in which the HC3:FU1 comparison lacks real significance while the HC1:HC2 and HC2:HC3 retain it, but it is clear that of the correlations, the HC3:FU1 is the weakest.



153 floras, fossil and modern described in Green and Hickey (2003) and including the 4 biozones defined by Johnson and Hickey (1990).

Complete linkage agglomerative hierarchical clustering using euclidean distance measure.

Figure 3.8: Algorithmic hierarchical cluster map showing how the Hell Creek and Fort Union biozones, despite spanning the K/T boundary, are better correlated with each other than any one is with 149 other fossil and modern floras. The comparative data comes from related work (Green and Hickey 2003). Note that the exact arrangement of floras in such a dendrogram is sensitive to distance measure, data standardization, and clustering method. The example presented here is for illustrative purposes only; it is not possible to determine from it, what we have concluded from broader experimentation, that the clustering of the Hell Creek and Fort Union floras is relatively robust.

A final illustration of the similarity of Johnson and Hickey's biozones to each other is provided by a hierarchical cluster analysis. Though the particular arrangement of the floras is far from robust (experimentation has shown that it is dependent on data pre-treatment, distance metric and clustering method), the clustering of the biozones does continue to appear despite variation in clustering technique. Thus we can conclude that the leaf architectural difference between the Hell Creek and Fort Union floras is much less dramatic than most other differences between floras, fossil and modern.

So by the leaf architectural or eco-morphlogical measure we are using, the K/T boundary shows only a weak preeminence in a section representing about three million years. This suggests that time and space averaging alone are not enough to explain our data: even when the plant record is examined at the smallest possible temporal and spatial scale, the ecological effects of the K/T boundary were not dramatic. The K/T boundary event does seem to have been marginally the most important effect on plant systems in three million years, but the evolutionary question is whether or not it had any effect that lasted on a longer time scale, and every examination of this question seems to indicate that forests responded homeostatically to the K/T boundary event.

This is comparable to results obtained by Cantrill and Poole (2005), in the southern hemisphere (Patagonia and Antarctica) where the K/T boundary is still visible in the fern spike, but minor, transient, or totally absent in the macrofloral record. It is worth also comparing this result with the analysis of the same boundary section by Labandeira et al. (2002), which showed a drop in the diversity and variety of insect feeding traces on leaves at the beginning of the Paleocene. My demonstration that leaf architecture did not change substantially at the same time, supports one of Labandeira's scenarios in which the disappearance of feeding traces was caused by insect extinctions: a crash due to failure of demand rather than supply.

So in answer to the question: 'How did the K/T boundary event affect the evolution of forest ecosystems?' the response suggested by these data is 'Not fundamentally.' Whatever species extinctions occurred must have been replaced within a few million years by migration of closely related species or re-radiation into empty niches. At the scale at which we can examine it, the plant macrofossil record shows no dramatic change in ecosystem structure, though of course the possibility remains likely that even shorter term fluctuations were more dramatic. If all the forests in North America were burned over in a single summer, that would clearly count as a dramatic ecological effect, but it seems intuitively likely that such an effect would have few or no effects that lasted longer than the time taken for the forests to regrow. The effects of migration plus seed banks would make any lasting change in the physiognomy of forests unlikely, and these theoretical predictions are (in general terms) substantiated by natural experiments like the catastrophic destruction of Krakatoa in 1883 (Thornton 1996). For plants the boundary event should be considered a ‘mass death’ not a ‘mass extinction’, and this raises the evolutionary question of whether plant ecosystems have ever suffered a setback comparable to animal mass extinctions.

An important caveat is that the data shown here capture information on the architecture of woody dicot leaves. Therefore what I am arguing is that the leaf-litter on the forest floor in the early Cenozoic was not significantly different from the leaf-litter on the forest floor in the late Cretaceous. The spread of grasslands in the Miocene, of course, restructured terrestrial ecosystems in a way that is probably almost as dramatic as the rise of angiosperms, but notice that it is entirely invisible in our data. One reason for this invisibility is that all angiosperm leaf forms were lumped together in figure 3.3, so in that particular representation differences in leaf shape among angiosperms are not apparent. More important, however, is the sampling bias in the fossil record: because the fossilization potential of herbaceous plants is vanishingly small in the climatic and edaphic conditions in which grasslands occur, we expect grasslands to go essentially unrepresented in the plant macrofossil record. So we should interpret our data as a history of lowland forest ecosystems, not as a history of terrestrial environments. The rise of angiosperms and the extinction at the K/T boundary were two of the most dramatic evolutionary events that restructured terrestrial ecosystems during the post-Paleozoic. Traditional descriptions of plant evolution and an examination of figure 3.3 substantiate the importance of the rise of angiosperms in the middle of the Cretaceous, but the K/T boundary event seems to have had no lasting effect on leaf architecture, and therefore by analogy on the ecological structure of forests.

A final theme for discussion is the potential for generalizing the use of morphological bins rather than diversity in the production and analysis of paleontological time series. Diversity, which is peculiarly susceptible to biases due to differing taxonomic practices, has been relied upon very heavily not only by paleontologists, but also by modern ecologists who have taken it to be one of the most fundamental attributes and best scalar descriptions of an ecosystem. Nevertheless, as criticism of Raup and Sepkoski’s data base, for example by Peters and Foote (2001), and the ecological literature on the connection between diversity and stability in ecosystems (Tilman 1999) have shown, evaluating the meaning of diversity can be highly problematic. Because of the nature of macrobotanical remains and the way they can best be organized, our data set suggests a useful alternative or supplement to the diversity curve for summarizing evolutionary change: choose a set of morphological bins (or, alternatively, a set of morphological variables), split the data set into the smallest units possible (operational taxonomic units), and record their proportional representation in each morphological bin or variable. Like any proportional metric, this minimizes the effect of biases in time series data by automatically normalizing for sample size, intensity of study, interval length, et cetera. Moreover, it provides a multivariate data set rather than the univariate time series provided by a measurement of diversity. Since environmental variables frequently influence morphological attributes in simple, predictable ways, functional explanation is also easier than in the case of diversity, which can interact with the environment in complex ways. Most importantly, diversity is a single variable known to be controlled by a number of factors and it is manifestly impossible to deconvolute complex interactions of many variables from a single composite measurement. Proportions of operational taxonomic units in morphological classes provide multivariate data distributed through time, so if the difficulties of dealing with correlated, non-normal, multivariate data can be overcome, then we have a reasonable hope of being able to investigate macro-ecological change in the remote past in more detail.

3.5 Conclusions

In the *Compendium Index* data, we find no indication of a change in the leaf architectural profiles between the Maastrictian and the Paleocene comparable to the changes that are seen earlier in the Cretaceous, associated with the rise of angiosperms. This lack of change at the K/T boundary highlights the influence of evolutionary innovation compared with catastrophic extinction on evolutionary history. Also, if we accept that there is a connection between ecosystem properties and leaf morphology, we must conclude that plant extinctions at the end of the Cretaceous, while they may have selectively eliminated certain species, do not seem to have restructured plant ecosystems in a way that was significant on an evolutionary time scale. Finally, in comparison with datasets with better stratigraphic resolution, the patterns identified show the importance of temporal scale in any discussion of macroevolutionary dynamics.

4 The problem of irregular sampling: A meta-analysis of the Climate Leaf Analysis Multivariate Program

The Climate Leaf Analysis Multivariate Program (CLAMP) is an established methodology for physiognomic analysis of dicot leaf floras. This chapter uses a meta-analysis of four studies that provide CLAMP data on 245 floras from Asia, Africa, and North and South America to demonstrate the application of a new analytical methodology for the exploration of the relationship between leaf morphology and environment. This methodology involves the application of a generalized 'pairs' plot or scatter plot matrix (SPLOM), a form of graphical analysis for multivariate data. It is compared with the results from regression, hierarchical cluster analysis, principle components analysis, and canonical correspondence analysis. Analysis of the available data using pairs plots reveals extensive multiple covariation among explanatory leaf physiognomic variables and identifies sources of systematic error that eigenvector ordination methods tend to conceal. Pairs plots provide a supplementary method for analyzing complex multivariate data on leaf physiognomy and contribute to biological understanding of leaf–environment interactions. Because pairs plots allow more flexible investigation of multivariate data than existing eigenvector and regression-based approaches, they may prove useful no only for analyzing CLAMP data, but also in exploring multiple covariation in other complex paleontological data sets.[2]

4.1 Introduction

As paleontological data become more quantitative, multivariate, complex, and voluminous, the choice of tools for data analysis acquires a greater influence over the biological and geological conclusions that are drawn from a given body of data. Either the data must be processed, summarized, its dimensionality reduced, and its details obscured; or we need new tools to handle the presentation and publication of larger, more complex data sets. In this paper, one such new tool—the 'pairs' plot—is suggested as a way of improving the standard procedures for examining the relationship between leaf architecture and environment.

The Climate Leaf Analysis Multivariate Program (CLAMP) is a method of analyzing fossil leaf assemblages or floras (specifically those deriving from the woody dicotyledonous component of forest ecosystems) by quantifying a set of significant morphological or architectural leaf variables and relating these variables (averaged over the flora) to climate parameters. Such a procedure also allows estimation of ancient climate parameters by uniformitarian extrapolation of patterns found in the distribution of leaf attributes in modern vegetation (Wolfe 1993, 1995; Wolfe and Spicer 1999).

---

[2] The material in this chapter has already appeared in print as Green (2006).

This general notion, which is sometimes referred to as 'leaf physiognomy', has been accepted since the early twentieth century when Bailey and Sinnott (1915) pointed out the strong correlation between the temperature in which modern forests grow and the proportion of the species that compose them that have 'entire' (i.e. untoothed) leaves. From this observed correlation in the modern world, a determination of the percentage of species with entire margins in a fossil flora allows the estimation of the temperature in which it grew. With the introduction of computers that could handle algorithmic classification and ordination of multivariate data, Wolfe (1993) proposed a multivariate method of coding leaves (originally based on 29, but later updated to 31 variables) that was intended both to improve the precision of temperature estimates over the univariate linear regressions that had preceded it, and to allow the estimation of other climatic variables.

In addition to temperature, other variables that have been more-or-less successfully estimated using CLAMP data are precipitation (e.g. Wilf et al. 1998), and moist enthalpy, which can be used to calculate paleo-elevation (e.g. Forest et al. 1999). The linear relationships between leaf size and precipitation and between leaf physiognomy and moist entropy are less clear than the relationship between leaf margins and temperature. As (Kennedy 1998:32) points out: 'It appears that CLAMP provides a relatively accurate estimation of temperature, but only a general estimation of precipitation.' Other variables (including many related to timing of changes in temperature and precipitation, like growing season precipitation, or warm month mean temperature) have been studied less extensively.

Despite the relatively widespread application of CLAMP methods, its procedures have been criticized as overly complex and no more informative than simple regression models (Wilf 1997; Wilf et al. 1998). Nevertheless it provides the only well-known procedure for collecting multivariate data on leaf morphology and in certain contexts has become a standard way of determining ancient terrestrial climatic parameters. Therefore the focus is almost exclusively on the CLAMP method, though some of the issues identified may also apply to other recent leaf physiognomic studies like the 'digital' approach of Huff et al. (2003), and Royer et al. (2005). Much of the debate about the advantages of the CLAMP method over various regression models centers around statistical details: the goal has been to maximize the 'explanatory power' of the method and minimize the standard errors of the temperature estimates that it provides. This may not, however, be the best way to choose an analytical methodology, because we have no satisfactory mechanistic explanation of the relationships between most leaf morphological characters and climatic variables. Thus we are by definition engaged in data analysis: that is, we are trying to determine what measured quantities signify and to design empirical models to predict them, not trying to test models based on theory against real data. Minimizing analytic error and maximizing explained variance produce a model that best explains a given set of data. Whether this model will ever explain any other data, be of practical predicative utility, or suggest fruitful lines of future inquiry, is a very different question.

In the following consideration of the available data, several issues with the CLAMP method that ought to be addressed become apparent. The focus throughout is on the analytical choices made, not on the collection of raw data: for the purposes of this discussion, it is assumed that the matrix of CLAMP scores is a relatively good reflection of the woody dicot leaf forms present in a living flora. No coding scheme is perfect, but the CLAMP method is the only such coding method that has been widely applied. In contrast, the statistical methods for analyzing CLAMP data represent only a small fraction of the available procedures for multivariate data analysis. Therefore, it seems necessary to explore to what extent the results of a CLAMP analysis are sensitive to the analytical methods chosen and to inherent biases in the data. Are the eigenvector and regression techniques that are generally applied to these data appropriate? What other techniques should be tried? In short: how can we improve the methods used to analyze CLAMP data? The alternative or supplementary analytical program proposed by this paper is based on graphical data analysis using pairs plots and seems to show substantial advantages over eigenvector approaches for exploring the relationships between dicot leaves and the environments in which they grow. Though the focus of this examination is specifically aimed at one type of paleobotanical data, the general issue of whether exploratory graphical analysis is more appropriate than data-reduction is applicable to many other paleontological data sets.

4.2 Materials and Methods

The data available for this analysis come from four studies. The first and largest, WOLFE173, is an updated version of the data published by Wolfe (1993), sometimes referred to as 'CLAMP 3B' (e.g. in Jacobs 2002). It is available on the web as two Excel spread sheets (.xls files) containing respectively the morphological and environmental data for 173 floras, points representing which are colored black in the figures in this paper. For the 103 floras that were published in 1993, the number of species in each flora and its latitude, longitude, and elevation were typed in from Wolfe (1993). The geographical distribution of this data set is mainly restricted to the continental United States and Japan, though there are a few floras from Alaska and continental east Asia. As of April, 2005 these data could be downloaded from the web at:
<http://www.open.ac.uk/earth-research/spicer/CLAMP/Physg3ar.xls> (the climatic variables) and
<http://www.open.ac.uk/earth-research/spicer/CLAMP/MET3AR.xls> (the morphological leaf scores).

The second data set, JACOBS, is from Jacobs (1999, 2002) which give CLAMP scores and associated environmental data for 30 floras in tropical Africa. This study used the original 29-variable coding scheme, so there are two variables with all values missing. Points from JACOBS are colored red in the figures in this paper. The third data set, GREGORY, is from Gregory-Wodzicki (2000), provides CLAMP scores and environmental data for 12 floras in Bolivia, and is colored green here. Finally, Kowalski (2002) provides CLAMP scores and environmental data for 30 floras in tropical South America, which are represented by blue points here.

The data in JACOBS, GREGORY, and KOWALSKI were scanned in from tables in the cited publications, processed for automatic text-recognition and then hand-edited and spot checked for accuracy. The data were read into the open-source program *R* (R Development Core Team 2005) from tab-delimited text files, which are provided in appendix A, and preprocessed so that all studies were in comparable form. The code used is given in the script file in appendix B. The data matrices are not printed because all the data have appeared in print before.

The completed data set consists of 245 floras and is stored as a series of data frames in R with the suffix -ALL for the raw frames containing both morphological and environmental data, and a suffix -CLAMP for the cleaned CLAMP scores. The complete data set is ALL, and the supplementary material typed in from Wolfe (1993) is a separate data frame called WOLFE1993. Stranks (1996) provides additional data from New Zealand that have not yet been processed.

Table 2: Variables in 31-dimensional Matrices

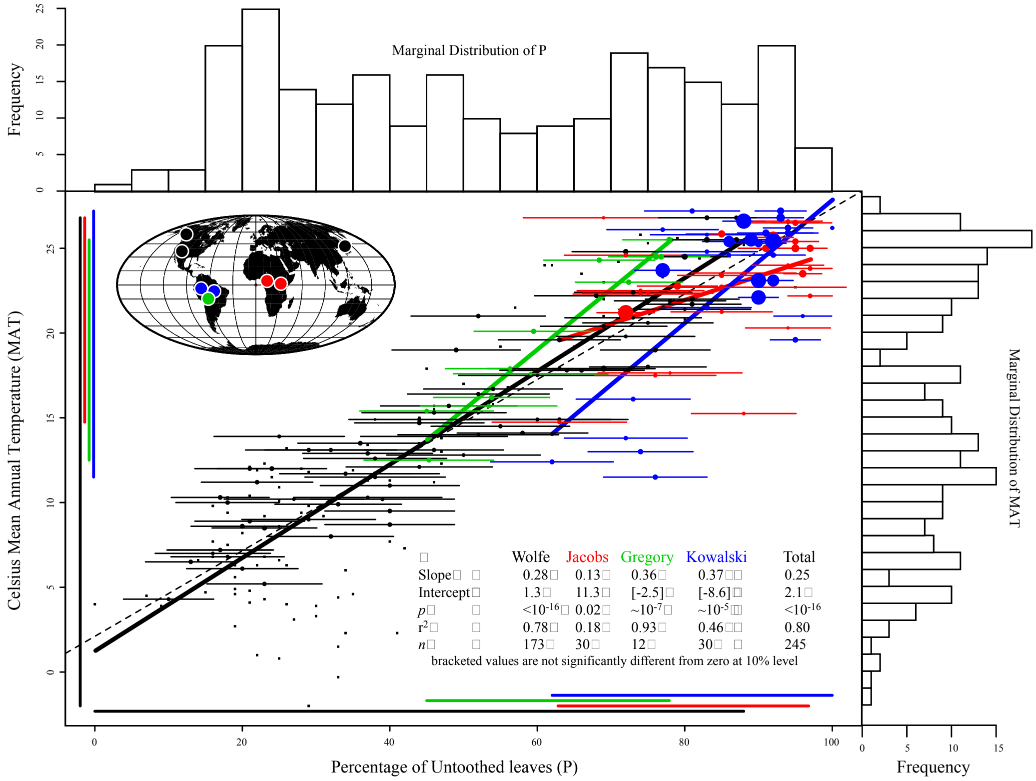
	VARIABLE NAME FROM Wolfe (1993)	ABBREV. USED HERE	COMMENTS
1	Lobed	Lobd	
2	No.Teeth	Entr	
3	Regular.teeth	TReg	
4	Close.teeth	TCls	
5	Round.teeth	TRnd	
6	Acute.teeth	TAcu	
7	Compound.teeth	TCmp	
8	Nanophyll	ZNan	Missing in JACOBS
9	Leptophyll.1	ZLe1	
10	Leptophyll.2	ZLe2	
11	Microphyll.1	ZMi1	
12	Microphyll.2	ZMi2	
13	Microphyll.3	ZMi3	
14	Mesophyll.1	ZMe1	
15	Mesophyll.2	ZMe2	
16	Mesophyll.3	ZMe3	Missing in JACOBS
17	Emarginate.apex	AEmg	
18	Round.apex	ARnd	
19	Acute.apex	AAcu	
20	Attenuate.apex	AAtn	
21	Cordate.base	BCor	
22	Round.base	BRnd	
23	Acute.base	BAcu	
24	L.W.< 1:1	Rlt1	
25	L.W.1-2:1	Rb12	
26	L.W.2-3:1	Rb23	
27	L.W.3-4:1	Rb34	
28	L.W.> 4:1	Rgt4	
29	Obovate	SObo	
30	Elliptic	SElp	
31	Ovate	SOvt	

The 31 physiognomic variables described in Wolfe (1993, 1995) are listed in Table 1. Unfortunately, unless growing season precipitation is taken to be the same as annual precipitation, the only environmental variable that appears in all four data sets is MAT, so our comparison of different studies is restricted to a single response variable. This is unfortunate because the main point of applying a multivariate framework is to elicit information about multiple response variables. Because little of the true uncertainty in a temperature estimate comes from analytic error in the explanatory variables (this contention is defended below), it is highly unlikely that a multivariate framework will really improve temperature prediction much, however much it can be made to reduce residual error of the regression.

4.3 Results

Exploring the relationship between a single pair of variables is simple and intuitive. For instance, as is well-known, a plot of the percentage of untoothed leaves (P) against mean annual temperature (MAT) shows a strong linear relationship. Figure 4.1 shows this relationship by plotting all available CLAMP data along with associated least-squares regression lines. The regression line for all the data is shown as a dotted line and limiting the regression to the floras for which information is available on the total number of species coded does not change the line perceptibly. The points are colored to show which study they came from and the thicker, colored lines show the results that are obtained when separate regressions are performed for each study. These regression lines are cropped to the extreme ranges of each data set, which are shown by colored bars near the edges of the plot.

Above and to the right of the main bivariate plot are histograms showing the marginal distributions of each variable. Note that both of these distributions are polymodal, probably because of irregular geographical sampling: there are relatively few floras representing the intermediate temperatures because the latitudes that would supply them (the horse latitudes) are kept dry by Hadley circulation and therefore have not provided as fertile a source for appropriate floras to study.



	Wolfe	Jacobs	Gregory	Kowalski	Total
Slope	0.28	0.13	0.36	0.37	0.25
Intercept	1.3	11.3	[-2.5]	[-8.6]	2.1
p	$<10^{-16}$	0.02	$\sim10^{-7}$	$\sim10^{-5}$	$<10^{-16}$
r^2	0.78	0.18	0.93	0.46	0.80
n	173	30	12	30	245

bracketed values are not significantly different from zero at 10% level

Figure 4.1: Comparative bivariate plot of temperature (MAT) against percentage untoothed leaves (P) for 245 floras and four studies, color-coded by study. Least-squares regression lines, marginal distributions, regression statistics, and the approximate geographical locations of the study areas are also shown. Points for which data are available on how many species are represented have error bars showing plus and minus the binomial sampling error. The size of the point plotted is also proportional to the sample size, except in the case where data are missing. The same colors representing the four studies are used in all subsequent figures.

It is evident from this univariate exploration that the 'study effect' (the effect on the regression line of which study the data is drawn from) is important, though it cannot be determined from this representation whether it is due to poor repeatability of the coding or whether it is caused by spatial autocorrelation. In this regard, note how the slope of the regression line in the two South American studies (GREGORY and KOWALSKI) is very similar, though the intercepts differ, while Jacobs's African data have a slope that is quite different from that found in the other three (predominantly New World) studies. This phenomenon of slope having a greater spatial autocorrelation than intercept has also recently been pointed out by Mosbrugger et al. (2005). This incomparability of models based on training sets from different regions has also been frequently remarked upon (Stranks 1996; Jacobs 2002; Spicer et al. 2004; Greenwood et al. 2004), but with equal frequency has been ignored when citing binomial sampling errors or standard deviations as if they were true uncertainties.

To expand our consideration from one explanatory and one response variable to 31 explanatory and one response is not trivial. Perhaps the simplest solution is the reduction of all 31 explanatory variables to a single distance metric. Clustering the available 245 floras hierarchically shows imperfect clustering by study (the 'blocks' of color) in figure 4.2, which shows dendrograms produced by an agglomerative hierarchical algorithm using the Euclidean distance metric under two clustering procedures (single-linkage and complete linkage), with different properties. (Single linkage clustering finds spherical clusters of objects in *n*-space; complete linkage finds strings of closely-connected objects.)

With a few exceptions (e.g. Traiser 2004) such clustering procedures have not been used extensively in leaf physiognomy, perhaps because they produce no explicit models or quantitative estimates of independent variables, but merely give a visual display of similarities among floras. From such a display, we can nevertheless qualitatively conclude that the studies do cluster together, but not without noise.

Much more prevalent—perhaps even ubiquitous among explicit considerations of CLAMP data—are eigenvector techniques for rotating multivariate vector spaces and re-projecting data along a few major axes of variation. The simplest and most general of these is principle components analysis (PCA). Originally Wolfe (1993) relied on correspondence analysis, and then (in Wolfe 1995) switched to canonical correspondence analysis, or CCA (Ter Braak 1986). Both methods were specifically designed for comparison of environmental data with species distributions, and have become fashionable in community ecology. Ter Braak (1986) makes it very clear, however, that:

> 'The rationale of the technique [CCA] is derived from a species packing model wherein species are assumed to have Gaussian (bell-shaped) response surfaces with respect to compound environmental gradients.' (Ter Braak 1986:1168)


Complete Linkage Dendrogram

Single Linkage Dendrogram

Figure 4.2: Complete linkage (top) and single linkage (bottom) agglomerative dendrograms colored by study showing (imperfect) clustering by study and area for 245 floras clustered according to the CLAMP variables.

and that

> ‘The vital assumption is that the response surfaces of the species are unimodal, the Gaussian (bell-shaped) response model being the example for which the methods performance is particularly good. For the simpler case where species-environment relationships are monotone, the results can still be expected to be adequate in a qualitative sense.…The method would not work if a large number of species were distributed in a more complex way, e.g., bimodally.’ (Ter Braak 1986:1177)

There are *a priori* reasons to expect species to have unimodal or linear distributions along envirnomental gradients, but this logic does not necessarily hold for morphological variables like the proportion of leaves with attenuate apexes. As can be seen in the 9th and 30th rows in the 32nd column of figure 4.5, discussed below, some of these relationships between morphological variables and temperature are arched or parabolic. Therefore the theoretical applicability of CCA to morphological variables averaged over floras on a continental or global scale is highly questionable. Like many other statistical methods, CCA is also vulnerable to non-linearity and multiple colinearity:

> 'When the data are collected over a sufficient habitat range for species to show nonlinear, nonmonotonic relationships with environmental variables, it is inappropriate to summarize these relationships by correlation coefficients or to analysis the data by techniques that are based on correlation coefficients, such as canonical correlation analysis.' (Ter Braak 1986:1167)

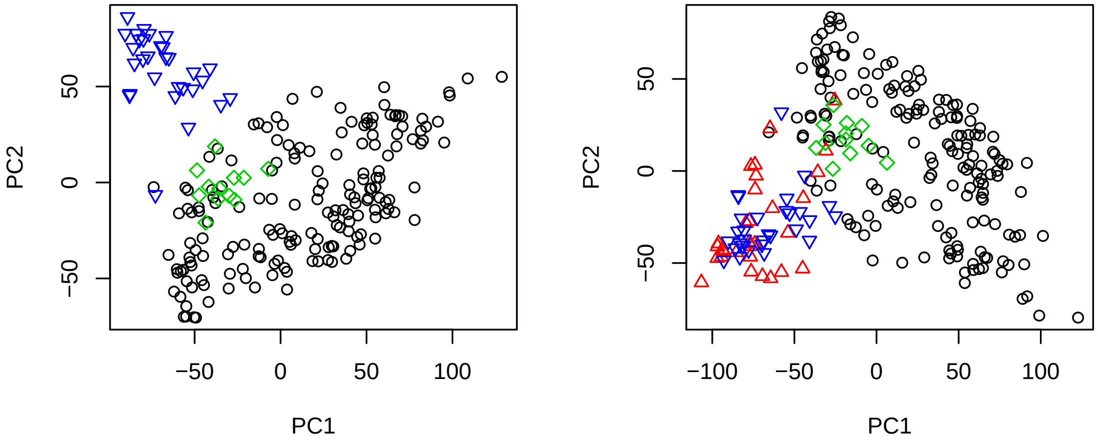
and

> 'When the environmental variables are strongly correlated with each other—for example, simply because the number of environmental variables approaches the number of sites—the effects of different environmental variables on community composition cannot be separated out and, consequently, the canonical coefficients are unstable. This is the multicollinearity problem.' (Ter Braak 1986:1170f.)

Many statistical procedures—including simple linear regression—work in practice even when their assumptions are unrealistic, so this alone would not invalidate the application of CCA to CLAMP data. The argument made here is not that CCA produces incorrect results, but merely that the ubiquitous application of it to CLAMP data may be evidence of excessive analytical rigidity.

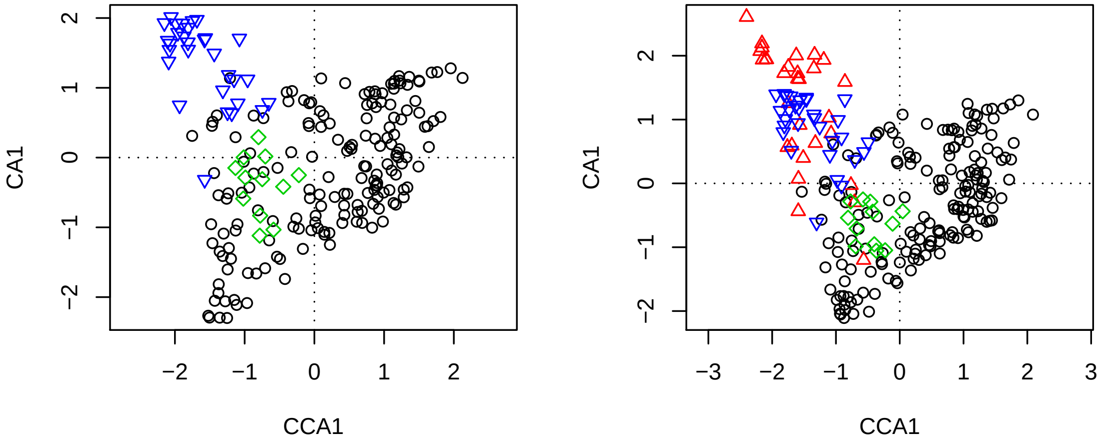
Most publications explicitly using CLAMP have followed Wolfe's lead even to the point of using Excel spreadsheet macros in the files that can be downloaded from the CLAMP website and a commercial program called CANOCO (Lepš and Šmilauer 2003) specifically designed to perform CCA. In fact CCA is now available in many general-purpose statistical packages, including three different implementations for *R,* and therefore continued reliance on compiled, proprietary software seems additional evidence of methodological canalization. As figure 4.3 illustrates, moreover, the analysis of CLAMP data is not a case in which PCA and CCA produce significantly different results: the top two bivariate plots are of principal components; the bottom two are canonical correspondence axes.

Note how in the upper left quadrant of figure 4.3, the bivariate plot of the first two principle components clearly discriminates KOWALSKI from the other two studies. When JACOBS is added, however, (the upper right quadrant of figure 4.3) the gap seems much less distinct. This illustrates



31−variable data set, Jacobs data omitted

29−variable data set, all studies



31−variable data set, Jacobs data omitted

29−variable data set, all studies

Figure 4.3: Principle components analysis biplots (top) and canonical correspondence analysis biplots (bottom) of the data. The left pair of biplots uses all 31 variables present in three of the studies; the right pair reduces the number of variables to 29 and shows all four studies. Note the basic similarity between the PCA and CCA plots: the scattters of points are viewed from different angles, but the relationships between studies are similar.

how sensitive this form of analysis is to sampling. The more data that are added, the harder it is to discriminate clusters that looked distinct when there were fewer points. Although only PCA and CCA have been tested, it is difficult to imagine a realistic situation in which other related eigenvector methods would lead to radically different interpretations of CLAMP data, although they may—like PCA and CCA—differ in the exact values they produce.

Both in the case of hierarchical cluster analysis and eigenvector analysis, it is apparent that the study effect contributes some structure to the data but by no means determines them. Formally, we could also use the multivariate t test (Hotelling's T^2) to check the pairwise null hypotheses of multivariate equivalence of means. In all $\binom{4}{2} = 6$ cases we are forced to reject the null with p-values less than 10^{-6}. In simple terms, the studies could not possibly all be equivalent. Note that the Hotelling code (see Green 2006) can not deal with missing values, so the number of variables had to be reduced to 29 in the three cases out of the six pairwise comparisons where variables were missing. Whatever the statistical logic, the data are clearly affected by the source from which they were obtained, though it cannot be determined with the available information whether this is due to the studies being in different geographic regions or whether people actually code leaves differently.

A dendrogram reduces 31 variables to a single distance metric, eigenvector methods reduce 31 variables to a few principle components, of which two are shown in figure 4.3. What about the remainder of the variables? One response is: the first two principle components account for a large proportion of the variance, so the other variables do not matter much. This seems to be a limited way of looking at the process of data analysis: if only the axes of maximum variance are of interest, then why collect multivariate data? Multivariate data are often collected to answer more than one question and a variable that answers a particular question (like the presence of teeth answering questions about temperature) may say nothing about another question (regarding, for instance, plant growth form). To choose variables exclusively from mathematical criteria like variance maximization seems to abdicate the responsibility for interpreting results biologically.

One way of looking at more variables is called a 'generalized draftsman's display' by Chambers et al. (1983), a scatter plot matrix (SPLOM) by Cleveland (1985) and Basford and Tukey (1999) or, more simply, a pairs plot. Figure 4.4 shows all the pairwise relationships between the original two variables plotted in figure 4.1 (P and MAT), the first two principle components (PC1 and PC2), the first canonical component (CA1, the primary axis corresponding to the matrix of sites) and the first constrained canonical component (CCA1, the primary axis corresponding to the environmental matrix).

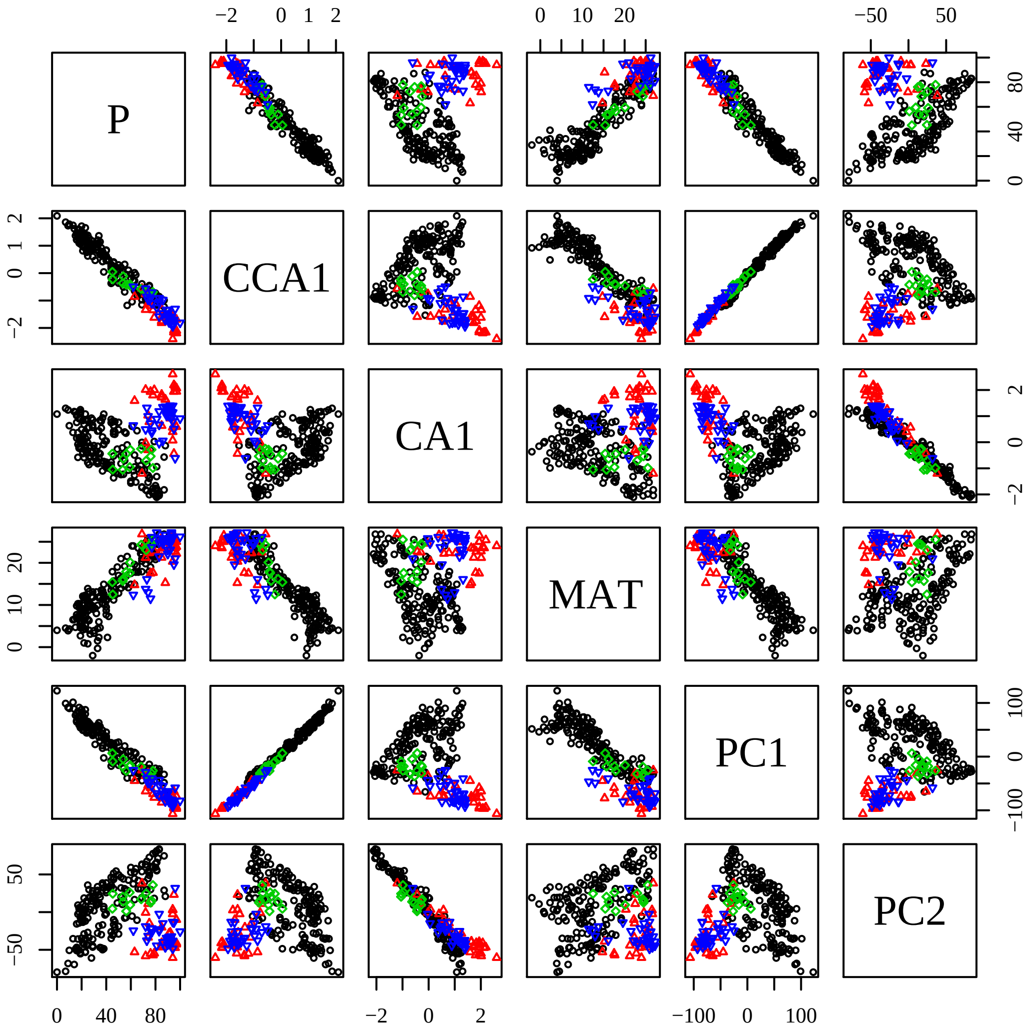


Figure 4.4: Traditional pairs plot as used in Basford and Tukey (1999). This is simply a matrix of scatterplots showing the relationships between each pair of a set of variables. All the data here are restricted to the original 29 CLAMP variables. In particular, note the similarity between PC1 and CCA1, demonstrating how little it matters which eigenvector method is chosen. Plots above the diagonal are inverses of the plots in the lower diagonal. Note that the plots in the third row, second column and sixth row, fifth column are exact duplicates of the right-hand two plots in figure 4.3.

As can be seen, a pairs plot allows the plotting of a very large number of multivariate data in a compact form. The question then arises: what is the value added by eigenvector methods of data reduction if it is possible plot and examine the raw data themselves? In figure 4.5, all the 31 explanatory variables and MAT are presented in this pairwise fashion, with additional details as described in the figure caption.

This is a very concentrated way of presenting data; it plots 32 × 245 two-digit numbers, the equivalent in characters of about twelve and a half manuscript pages of text. Each of the small plots above the matrix diagonal is a similar bivariate plot showing the relationship between two of the 32 variables. Thus the scatter plot in the 32nd column and 2nd row of the pairs plot is a reduced version of figure 4.1; it is simply the bivariate plot of P against MAT. The second and 32nd of the diagonal cells also correspond to the marginal histograms in figure 4.1. The shadings below the diagonal are obtained by performing four two-sided hypothesis tests for each cell:

H_0: slope of the least squares regression line = 0
H_0: Pearson's product-moment correlation coefficient = 0
H_0: Spearman's rank order correlation coefficient = 0
H_0: Kendall's rank order correlation coefficient = 0

The cell is colored white if the mean of the three correlation coefficients is positive and if all tests reject H_0; black if the mean of the three correlation coefficients is negative and all tests reject H_0; and medium grey if all tests fail to reject H_0. If some but not all of the tests reject H_0, the cell is colored light grey or dark grey depending on the sign of the mean of the correlations of the tests that are significant. In all cases the color of the text is black if the mean of the three correlation coefficients is positive and white otherwise. All tests are made at the level $\alpha = 0.05$ / number of comparisons, where the number of comparisons is (number of variables) × (number of variables - 1) / 2, i.e. the 5% level with Bonferroni correction for multiple comparisons.

This representation of the data allows us to examine complex multiple-covariation among the explanatory variables in detail. For instance, compare the second column with the third-through-seventh column block. They are inverses of each other, as they must be, because the second column represents the percentage of species lacking teeth and the third-through-seventh columns give the percentage of species with particular types of teeth. Another interesting block of covarying values is provided by the leaf sizes: columns and rows 8 through 16. Here, the smallest four leaf size categories are all strongly positively correlated with each other as are the largest four leaf size categories, while there is a strong negative correlation between the small and large blocks. Only the middle three size categories are not strongly collinear. Graphical display of this sort of data makes the strong covariation among the variables apparent and indicates that any statistics calculated from them that

See Plate 1

Figure 4.5: Elaborated pairs plot of 31 explanatory and one response variable. The variables are represented along the top of the matrix by the figures from Wolfe (1993) that were originally used to illustrate them, and along the left side by their abbreviated names. In two cases—‘teeth regular’ and ‘teeth close’—the line drawings colored pink also show the opposite of the character being coded, ‘teeth irregular’ and ‘teeth distant’. Groups of related variables are shown by the colored bars along the bottom and left and those groups of related variables that are constrained to sum to 1 are so marked along the bottom of the matrix. Above the diagonal are scatter plots like the example shown in figure 4.1; below the diagonal, the white, black, and grey squares show whether the relationships between the variables are statistically significant. The white squares show a strong positive relationship between the variables, the black a strong inverse relationship, and the intermediate shades of grey show weaker or absent relationships. Along the diagonal are plotted histograms of each of the 32 variables (the marginal distributions for the bivariate plots).

assume independence should be treated with caution.

At such a small scale, it can be difficult to see details of the scatter plots, so figure 4.6 shows another pairs plot of a subset of the variables.

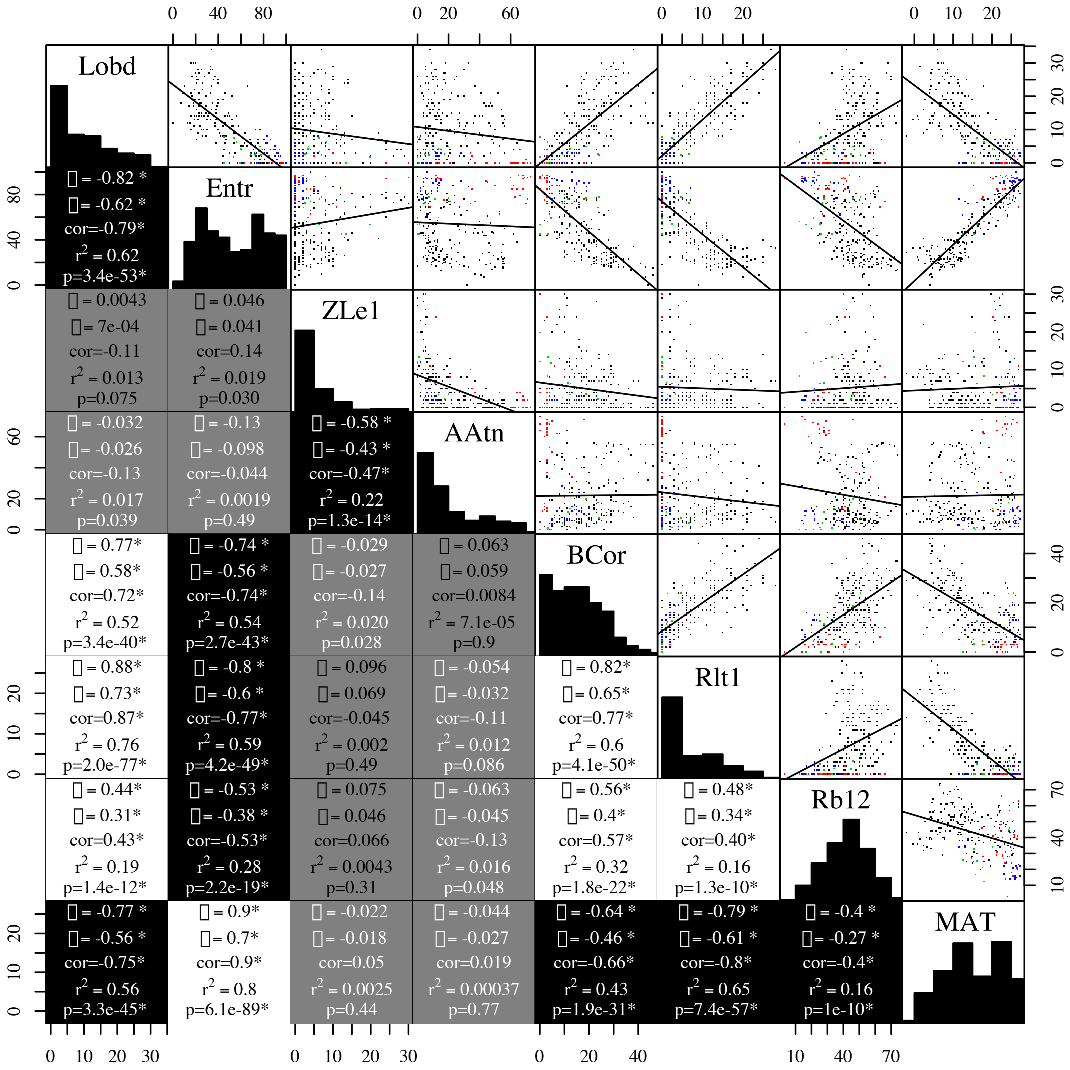


Figure 4.6: Pairs plot of selected variables. ρ is Spearman's rank order correlation coefficient; τ is Kendall's rank order correlation coefficient; cor is the ordinary (Pearson's) product-moment correlation coefficient; and the r^2 and p-values relate to the least-squares fitted line. Where the associated correlation tests (and in the case of the p-value, the regression) are significant at the 5% level, the statistics are followed by asterisks. The blocks are white with black text if the mean of the three correlations is positive and all four of the tests are significant, and they are light grey with black text if some, but not all of the tests are significant. The blocks are colored black and dark grey with white text in the reverse scenarios. If none of the tests are significant, the blocks are a medium grey and the sign of the mean of the three correlations is shown only by the color of the text printed on them: black for a positive and white for a negative correlation.

In this plot, in addition to the scatterplot matrix above the diagonal, the numbers in the blocks below the diagonal give all of the pairwise correlation coefficients (the two common non-parametric correlation coefficients, Spearman's ρ and Kendall's τ as well as the ordinary product-moment correlation coefficient) and the r^2 and p-values for the least-squares fitted line.

This reduced set of variables could have been selected by a stepwise multiple regression procedure with formal rules for adding or subtracting variables from a model based on information criteria. This has been done and it is easy to produce models in which all terms are significant with r^2 as high as 0.8 and $p < 10^{-16}$. Since there are several such stepwise techniques, all of which produce about the same quality of model from different variables, there is no reason to choose one model over another, and the variables for this smaller plot were selected based on their perceived interest instead of on a formal stepwise procedure.

This particular choice of explanatory variables happens to produce a multiple linear least-squares model for MAT with r^2 of 0.86 and $p < 10^{-16}$; adding a factor to the model showing which study each observation came from as an additional explanatory variable increases the r^2 significantly to 0.88.

These models could certainly still be improved further by continued massaging (for instance, interactions, non-linear terms, and variable transformations were not even tried). It does not seem valuable, however, to spend time massaging a model until the issues that the univariate case brought up (like incomparability of studies) are taken care of.

4.4 Discussion

The CLAMP method can be criticized at two levels: data collection and data analysis. This paper is concerned primarily with improving the methods of analysis, but there are a few problems with the analyses that are based in the data collection.

First is the fact that the raw (by species) scores have, except in doctoral dissertations (Stranks 1996; Kennedy 1998) seldom or never been published. This means that some of the most important and interesting questions about phylogenetic distribution of leaf morphological variables and the differences among plants of different growth form (habit) can not be asked. Some recent studies like that by Kennedy et al. (2002) have not even published the CLAMP scores averaged by flora, but have only printed biplots of eigenvector loadings. This form of presentation is so highly processed and incorporates so many assumptions that it makes interpretation of the results difficult and reanalysis of the data impossible.

Second, the selection of characters in Wolfe (1993) was explicitly based on preliminary eigenvector analyses, for instance:

> During one stage of the study, I expanded the character set to include about 20 character states additional to [the original 29]. Judging from eigenvalues and percent of total variance accounted for, these characters either added nothing or even lowered both eigenvalues and percent variance. (Wolfe 1993:20)

To reject potentially interesting variables from a coding scheme on the basis of low eigenvalues and percent of variance accounted for is to allow the statistical horse to bolt: you proceed quickly, but have little control over the direction you are traveling. Among the character states rejected, were those relating to compound leaves, spinose teeth, and inrolled or thickened margins, all of which have clear mechanical adaptive significance. If we hope to obtain ecological or environmental data from leaf physiognomy as well as information about climate, such character states should be retained.

The same criticism goes for lumping together characters like ‘teeth round’ and ‘teeth appressed’ merely because ‘combining the states produced both higher eigenvalues and percent variance’ (Wolfe 1993:24).

Thirdly, the description of some of the character states seems ambiguous. Though it is not possible without a comparative study to say for certain that interpretations of the character descriptions would vary, it is not clear whether ‘0.25 if the teeth are both regular and irregular and some leaves have teeth’ (Wolfe 1993:24) should be interpreted as ‘0.25 if the teeth are regular and/or irregular and some leaves have teeth’ or as ‘0.25 if the teeth are all regular or all irregular and some leaves have teeth’.

In the case of deeply lobed leaves, the leaf size is supposed to be scored from a single lobe, but the aspect ratio and shape still refer to the overall leaf, whereas in the case of a compound leaf the leaflet is what is scored for size, aspect ratio, and shape. This is particularly problematic in genera like *Rosa* in which a plant can have compound, deeply lobed, and simple leaves on the same branch.

Fourthly, the scored variables are divided into sections relating to common topics. Leaf size, for instance is coded as proportions of leaves falling into nine size classes. The scores in some of these sections, like leaf size, aspect ratio, or shape must sum to one while the scores in the section describing teeth and lobation do not have a constant sum. This means that the presence of teeth is implicitly weighted more heavily in the overall description of the flora than, for instance, the leaf size, and it is not clear that any normalization procedure can correct this bias. The restriction of groups of variables to a constant sum introduces dependence and implicit weighting that are hidden by eigenvector analyses but made apparent by a graphical approach (see figure 4.5).

Despite these issues with the process of coding, no morphological coding scheme could be ideal and these criticisms of CLAMP are offered in a spirit of improving what is the only such system currently available in the published literature. In particular, Wolfe's original article (Wolfe 1993) was much more broadly focussed than some subsequent publications: a discussion of axes of variation other than those corresponding to temperature and precipitation made it not only a contribution to paleoclimatology but also ecologically and botanically interesting.

More important than these problems in the coding are the true uncertainties associated with the estimation of paleoclimatic variables. It is generally accepted that the leaf physiognomy of a flora indicates the general climate regime of the area in which it grew: 'tropical', 'sub-tropical', or 'temperate', and 'wet' or 'dry'. These are categories that not only any botanist, but many laymen would recognize from simple leaf silhouettes. Beyond this there remains doubt as to the degree of precision and reliability that leaf physiognomy can provide, but there has been relatively little general discussion of what causes the real uncertainty in the procedure of estimating ancient environmental parameters from leaf morphology.

It is noticeable that two doctoral dissertations that have examined CLAMP data in detail are less sanguine about the errors associated with the methodology than most published articles. Stranks (1996) cautions that 'the method is still in in a developmental stage with many questions remaining unanswered' (Stranks 1996:122) and 'that a relationship exists between physiognomy and climate is clear. Whether it can successfully be applied to fossil floras in order to extract climate and altitude, however remains to be resolved' (Stranks 1996:124). Though she does not use the term 'spatial autocorrelation', she correctly observes that 'the response of southern hemisphere sites in general cannot be compared to those of northern hemisphere sites' (ibid) and Greenwood et al. (2004) support this contention. Kennedy (1998) lists several sources of potential error and admits that 'qualitative sources of error, such as subjectivity in morphotyping and taphonomic bias, could potentially introduce large amounts of uncertainty into palaeoclimatic interpretations' (Kennedy 1998:20). In contrast to this conservatism, many publications suggest that: 'CLAMP...is a powerful paleoclimate proxy with the ability to yield quantitative data on past temperatures, precipitation, growing season length, and humidity, as well as enthalpy' (Spicer et al. 2005:429).

Some of the sources of error that must be dealt with are, in rough increasing order of relative importance or difficulty of quantification:

1. Binomial sampling error. This is the simple and well-understood error associated with the random selection with replacement of n leaves out of a population of which a proportion P have untoothed margins. If this selection is repeated many times, the standard error of P should approach $\sqrt{\frac{P(1-P)}{n}}$. This imposes a minimum error on the order of a few degrees with floras of about 30 species.

In floras that have many more species (e.g. >100), the binomial error becomes insignificant (Wilf 1997).

2. Repeatablity of coding. At this stage, it is not clear what errors may be produced by different people coding the same floras, so this potential source of error is not readily distinguishable from spatial autocorrelation or the study effect discussed above. Future work will investigate this source of error using blind experiments.

3. Spatial autocorrelation and irregular sampling. The current sampling distribution is very poor, but can be improved by collecting more samples where they are lacking, by gridding the available locality data on a raster and applying statistical tools spatially, and by creating spatially distributed artificial floras from species range data as has been done by Traiser et al. (2005). Unfortunately, climate station data are seldom or never available from exactly the same places as floras are collected. Up to a point, this can be addressed by appropriate methods of interpolation, but errors introduced by microclimatic variation and patchy species distributions may continue to remain problematic.

4. Inherent time-averaging. This is not an issue if MAT is the only dependent variable, but MAT is a grossly time-averaged quantity that will be perceived differently if data on, e.g., mean monthly temperatures are compared across studies. It is easy to illustrate how dramatically plants have evolved to respond to the *timing* of temperature change: CAM plants open their stomata during the night when it is cool and transpiration is reduced. As soon as one calculates average daily temperatures—much less monthly or yearly means—from an hourly record, one loses the ability to explain an entire evolutionary strategy that allows thousands of species of plants to exist. This is an extreme example, but the more general point that different temporal scales will affect the significance of variables like temperature must be taken into consideration.

5. Other sources of noise (elevation, microclimate, disturbance, soil type, systematics, taphonomy, etc.) All of these variables are known to be of importance at particular spatial and temporal scales, and must be considered. Is the sample skewed by collection of more low-altitude floras than high-altitude? Do secondary-growth (recently cleared) forests respond differently than primary forest? In the absence of clear answers to these questions about systematic biases, calculation of a stochastic binomial sampling error becomes nearly irrelevant.

6. Uniformity through time. How far back in time can spatial patterns observed in the modern day be extrapolated? This is a broad question facing all methods of reconstructing ancient climates; a simple criterion that is often implicitly invoked is that a method must work increasingly well as it approaches the present; hence error must increase as we go back in time.

The error figures usually associated with estimates of mean annual temperature (MAT) from leaf morphology are usually one- or two-standard deviation analytic errors calculated by assuming only binomial sampling error or normally distributed stochastic variation in the explanatory variables and then propagating this error through a regression line. When the number of species increases much beyond a typical 30, these analytical errors are dramatically reduced, which has led to the appearance in the literature of, for instance, temperature estimates of plus or minus a few degrees (Burnham et al. 2001; Kowalski 2002; Kennedy et al. 2002). Even errors of under a degree have appeared, which as Miller et al. (2006) point out is incompatible with a rigorous error analysis of the relationship between P and MAT.

Errors 4–6 may be ultimately unquantifiable and uncorrectable, but there is abundant evidence that the issue of spatial autocorrelation can be handled. Work by Thompson et al. (1999) provides graphical tools for plotting floras in ecological space and Traiser (2004) and Traiser et al. (2005) give spatially distributed leaf physiognomic data from synthetic floras for the whole continent of Europe. In concert with the sort of exploratory data analysis that is presented here and in chapter 6, these techniques may make it possible, not only to improve estimates of terrestrial paleoclimates, but also to extract additional types of data about how environments and plant ecosystems have changed through time.

4.5 Conclusions

A graphical exploratory investigation of CLAMP data reveals further serious and unaddressed statistical issues with the standard procedures used to analyze the data. *Exploration* and *estimation* are different goals. If the only utility envisioned for fossil leaf floras is the production of ever-more-precise but possibly inaccurate climate estimates, then the current methods of publication and analysis of CLAMP data are satisfactory. In order, however, to understand the ways in which plant leaves respond to environmental stimuli in the context of real communities, we need application of looser, more flexible tools for data analysis, an appraisal of uncertainty that accounts for systematic bias and unquantifiable noise as well as trivial stochastic errors, and the publication of raw data in a form that can be compared between studies. Graphical techniques like pairs plots are effective methods of exploratory analysis of multivariate data, but theories of biological interest like mechanistic models of leaf response to environmental variables cannot be tested against such data unless the standard forms in which the data are currently published are extended to include the raw (by species) scores.

Leaf morphology remains a valuable and under-exploited source of multivariate data. The CLAMP method is not ideal, but it it is the best source of data currently available. It can give satisfactory results if the data it produces are published and analyzed appropriately. From a biological as opposed to a strictly paleoclimatological perspective, appropriate analysis consists of taking the ‘climate’ out of CLAMP and allowing multivariate data on leaf architecture to illuminate broader ecological questions. The pairs plot as a tool for graphical exploratory analysis provides information on complex covariation among leaf-physiognomic variables, and allow evaluation of systematic errors in CLAMP data, neither of which can be done with eigenvector methods of data reduction or with hierarchical clustering. This has the potential to make multivariate leaf-physiognomic data interesting not only to paleoclimatologists, but also to plant morphologists and functional ecologists. Moreover, the exploratory graphical approach advocated here may prove valuable in other paleontological data sets where current analyses obscure interesting detail in complex, multivariate data.

5 Leaf architectural patterns in space: Constructing synthetic floras for North America

For almost a century, leaf margin analysis has been an important method of estimating Cenozoic palaeotemperatures. However, the relation between the proportion of toothed versus entire leaves and temperature has so far been calibrated and characterized using relatively small numbers of opportunistically and irregularly collected floras. Here I present the results of a spatially-distributed analysis of the relation between leaf margins and temperature for North America. I used species range maps of native dicot trees to derive synthetic local floras for each 50 km grid cell in North America and compared the percentage of entire-margined species with the mean annual temperature in each cell. Leaf margin type for each species was assigned using illustrations from appropriate regional North America floras. This analysis confirms a strong inverse relationship between toothed margin percentage and mean annual temperature between 0 and 25°C in eastern North America. The estimated temperatures are roughly similar to those obtained from floras in east Asia and previous analyses in North America, but the detailed curve is more complex, showing a distinct change in slope above a mean annual temperature of 20°C. [3]

Wetter climate areas of the coastal western North Amercia do not show any significant correlation between leaf margin percentage and temperature. This may be due to the confounding effect of diverse topography, or it may be caused by the low diversity angiosperm tree flora of this area. Like previous studies, this paper demonstrates that there indeed is a strong relationship between leaf margin percentage and temperature on a regional scale in eastern North America. However the relationship does not have the same linear behavior as results obtained from previous analyses. I also discuss the implications of the results obtained here for palaeoclimatic reconstruction in the context of other leaf margin analyses based on North American sites.

---

[3] Note on authorship: an article written collaboratively with Jonathan M. Adams and Yangjian Zhang of Rutgers University containing the material in this chapter is also in review at the journal *Global and Planetary Change.* Division of work between the author (WAG) and collaborators JMA and YZ was as follows: coding of taxa was shared equally between WAG and JMA/YZ. GIS analysis used in this version was performed by YZ; an earlier analysis based on a smaller data set was performed by WAG. Figures were produced by WAG. Text was produced by WAG and JMA; this version has been edited to remove all wording identifiably from JMA. The coding done by JMA/YZ relied on web searches by Latin name to reveal images and or descriptions of leaves, several of which were inspected before assigning a categorization. The coding done by WAG relied only on published sources cited here.

5.1 Introduction

It has long been known that trees of cooler climates in the northern hemisphere tend to have toothed leaves, while trees with entire-margined leaves are more common in subtropical and tropical climates.

The adaptive significance of this relationship between leaf form and temperature remains unknown. The presence of teeth has been related to herbivory (Brown and Lawton 1991), to efficiency in the design of distributary networks (Givnish 1978), and to the deciduous habit (Baker-Brosh and Peet 1997). Another hypothesis is that the thinner shorter-lived leaves of cooler climates require teeth or lobes to allow them to flex in the wind without tearing (Baker-Brosh and Peet 1997). Most recently, Canny (1990), Baker-Brosh and Peet (1997), and Royer and Wilf (2006) have developed an explanation for the adaptive significance of teeth based on their ability to increase transpirative water flow by decreasing the apparent thickness of the aerodynamic boundary layer around leaves, thus supplying extra nutrients to toothed leaves early in the growing season when low evaporation rates due to low temperatures limit photosynthesis.

Despite the continuing doubt as to the adaptive significance of teeth, linear models relating P (the proportion of entire leaves in a flora) and MAT (mean annual temperature) have been used to estimate past temperatures from fossil leaf floras since the early 20th century (Bailey and Sinnott 1915).

In the past 30 years, this relationship has been recalibrated many times using more data from floras in North America by Dolph (1976, 1979), Wolfe (1979, 1993), Wilf (1997), and Kowalski and Dilcher (2003). Leaf margin analysis has also been calibrated from modern floras elsewhere in the world (Herman and Spicer 1996; Stranks 1996; Kennedy 1998; Jacobs 1999, 2002; Gregory-Wodzicki 2000; Kowalski 2002; Kennedy et al. 2002; Greenwood et al. 2004; Traiser et al. 2005), and used to estimate palaeotemperatures in the Tertiary (Wing and Greenwood 1993; Wolfe 1995), around the K/T boundary (Wolfe and Upchurch 1986, 1987), and as far back in time as the Albian (Miller et al. 2006).

In addition to examining leaf margins, Wolfe (1993), Wing and Greenwood (1993), Huff et al. (2003), and Traiser et al. (2005) have also tried to improve temperature estimation by using multivariate leaf data, but the advantage of a multivariate approach has not been conclusively established, so I restrict my current consideration to leaf margin analysis.

Most recently, Traiser (2004) and Traiser et al. (2005) have suggested the use of synthetic floras for grappling with the spatial variation in the relationship between leaf physiognomic and climatic variables in western Europe. In addition to considering a number of other variables, including leaf size, they identified a general latitudinal trend in leaf margin percentages, but the narrow temperature range in the moist climates of Northern Europe (southern Europe is dominated by more arid Mediterranean climates with a precipitation-limited winter growing season) limits the applicability of their results in North America, which has a far greater range of temperatures in the moist eastern forest zone, ranging from boreal in Canada to subtropical in southern Florida.

In this chapter I apply an approach based on that of Traiser (2004) to the North American flora in order to test the applicability of linear models for temperature estimation via leaf margin analysis and provide a recalibration of the relationship between temperature and leaf margin percentage based on spatial analysis at a continental scale.

5.2 Methods

As a source of data on tree floras, I used the online version of the *Atlas of North American Trees* (Critchfield and Little 1966; Little Jr. 1971, 1976, 1977, 1978), which has been digitized as shapefiles suitable for analysis using geographical information systems (GIS) software by Thompson et al. (1999) and is available on the world wide web at (<http://pubs.usgs.gov/pp/p1650-a/>). In theory, all native Canadian and US trees and large shrubs (defined as >3m tall in any part of their range) are included in the *Atlas.* The maps are generally thought to be accurate to at least the county level.

The range maps were compared with a 50 km grid covering North America. This grid size was chosen to provide samples at a relatively fine geographic resolution, while minimizing the effects of microclimatic variability, where a range of environments within a gridcell might confound the analysis. A 50km cell exceeds the maximum distance over which leaves might be expected to be transported before being deposited and preserved in the fossil record, and is roughly equivalent to the implicit county-level accuracy of the original range maps. A larger grid-size would lose data present in the original range maps; smaller would imply greater precision that is really available from the originals.

Each of the 512 broadleaved dicots in the online *Atlas* as of 2006 was included. Monocots, conifers, and leafless species (e.g. cacti) were excluded. Leaf margins were classified using Gleason and Cronquist (1963), Sargent (1905), Little Jr. et al. (1964, 1974), of North America Editorial Committee (1993–), Britton and Brown (1913), and Hickman (1993), as well as web-based resources.

Species were classified as toothed (rather than lobed) if indentations in their leaves were present and extended less than 0.25 of the distance from the outermost margin towards the mid-vein of the leaf. Following the procedure of Wolfe (1993), species with only entire leaves were assigned a value of 1, those with only toothed leaves were assigned a 0. Species in which presence or absence of teeth is a variable trait were assigned a value of 0.5. A list of the categories assigned to each species and whether its leaves are deciduous or evergreen can be found in appendix A.

The score for each grid square was obtained by averaging the values obtained for all the species whose ranges overlapped into that grid square. Note that this methods section refers to the final collaborative analysis the results of which are presented here. The original GIS analysis done by the author used 1/2 degree or about 60km grid cells and only about 250 species. See footnote at the beginning of this chapter for full details of collaborators' input.

5.3 Results

Summary maps and scatter plots (figures 5.1, 5.2) show that the percentage of entire margins reaches its highest point in the south-eastern North America, and is generally high in the east, compared to other parts of the continent. There is clear general latitudinal trend, complicated by differences in high relief regions of the west. Grid cells with fewer than 20 codable species are not colored.

Figure 5.1 shows the points coded by mean annual temperature. Because of the very good correlation between latitude and mean annual temperature, this is roughly equivalent to latitudinal coloring, but note, for instance, the cold anomaly where the Appalachian mountain chain in eastern North America is shown by green and light-blue 'tongues' pointing southwest, parallel to the east coast.

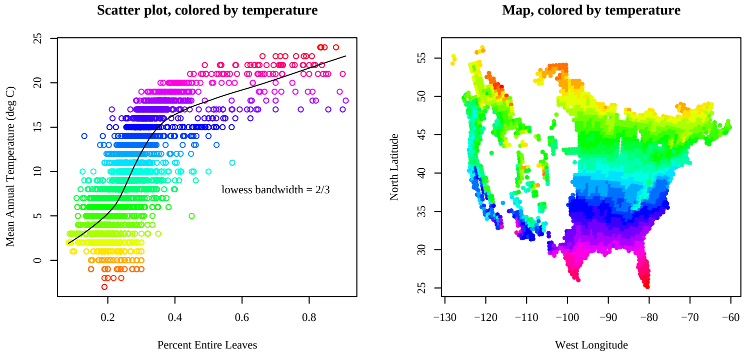


Figure 5.1: Summary map and scatterplot showing proportion of entire-margined species in the North American native tree flora color coded by mean annual temperature. Black line is lowess non-linear regression line.

Figure 5.2 shows the relationship between temperature and leaf margin percentages colored by the proportion of entire leaves. With this coloring scheme, the effect of continentality is also visible: the floras with the smallest proportions of entire leaves (colored orange) are not those with the very coldest mean annual temperatures, but in the continental interior (southern Quebec and Ontario). This suggests that at a given mean annual temperature, more equable continental areas (dryer, with greater mean annual range of temperature) will have fewer entire leaves than coastal regions.

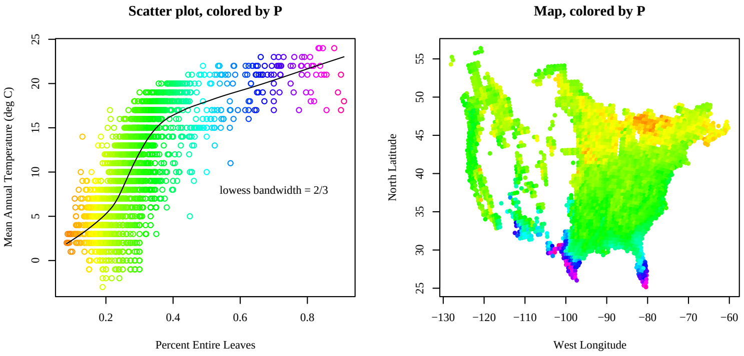


Figure 5.2: Summary map and scatterplot showing proportion of entire-margined species in the North American native tree flora color coded by the proportion of entire leaves. Black line is lowess non-linear regression line.

Note that in both of these figures, there is a clear relationship between leaf margin percentage and mean annual temperature for the forest region of the eastern half of the continent. If the eastern coast is examined independently, the relationship between mean annual temperature and leaf margin percentage becomes even ‘tighter’, though it remains clearly non-linear (figure 5.3).

Margin percentage changes relatively slowly with respect to temperature between 0 and 20°C with a wiggle to the right at around 10°C, showing reduced continentality. The rate of change with temperature then accelerates at around 20°C, and keeps a more-or-less linear relationship up to about 25°C, at the geographical limit of southern Florida. These data show more scatter than is evident in previous work on this subject in China (Wolfe 1979) and North America (Wilf 1997), perhaps because more samples were included across a diverse range of soil types, elevations, and microclimates.

Previously published studies for parts of North America and eastern Asia have suggested only a linear relationship between temperature and leaf margin percentage (figure 5.4). The relationship found here intersects with the linear relationships suggested by other studies at lower and higher temperature values, but curves away from that line at intermediate values. Note that the data points used by Wilf (1997) happen to fall only at those extreme values, and the line drawn between them is

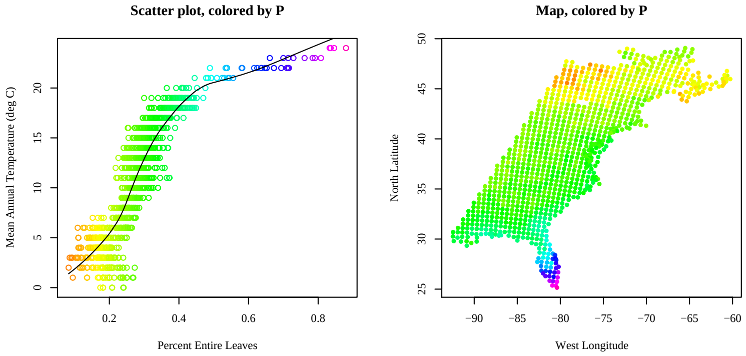


Figure 5.3: Proportion of entire leaf margins against temperature in the eastern North American forest region demarcated above. Points color coded as in figure 5.2 by the proportion of entire leaves. Black line is lowest non-linear regression line.

therefore an interpolation. The data analyzed here suggest that if intermediate temperature zones were included, the true relationship would be arched. In retrospect, a close examination of Wolfe and Spicer's complete (173 flora) data set, including high-elevation floras, suggests the possible presence of this non-linearity, which can now be documented in detail. In figure 5.4, I show a comparison of the east-coast subset of the new synthetic data with six previous studies. In broad terms the new data substantiate previous work, but the strongly non-linear relationship has not been observed before.

On the western side of North America, areas between northern California and Alaska receiving more than a meter of rain per year show only a weak correlation with temperature over a range of some 25°C (figure 5.5). The same is true if the coastal region in general is included, irrespective of rainfall (figure 5.6). (Note that although the slope of the least squares regression lines are statistically different from 0 in both cases, the spatial autocorrelation of the gridded data means that the p-value of the regression cannot be trusted.) The poor correlation apparent from the plots in figures 5.5 and 5.6 is surprising because important early studies of the correlation between leaf margins and temperature by Wolfe (1993) were heavily based on data from western North America.

Examinination of the relationship between percentage of broadleaf evergreen leaves and temperature, geography, and percentage entire leaves (figure 5.7) reveals expected patterns: there are more broadleaf evergreen species in low latitudes, high temperatures and where most of the leaves are entire. Percentage of evergreen species does not, however, do as good a job at predicting temperature

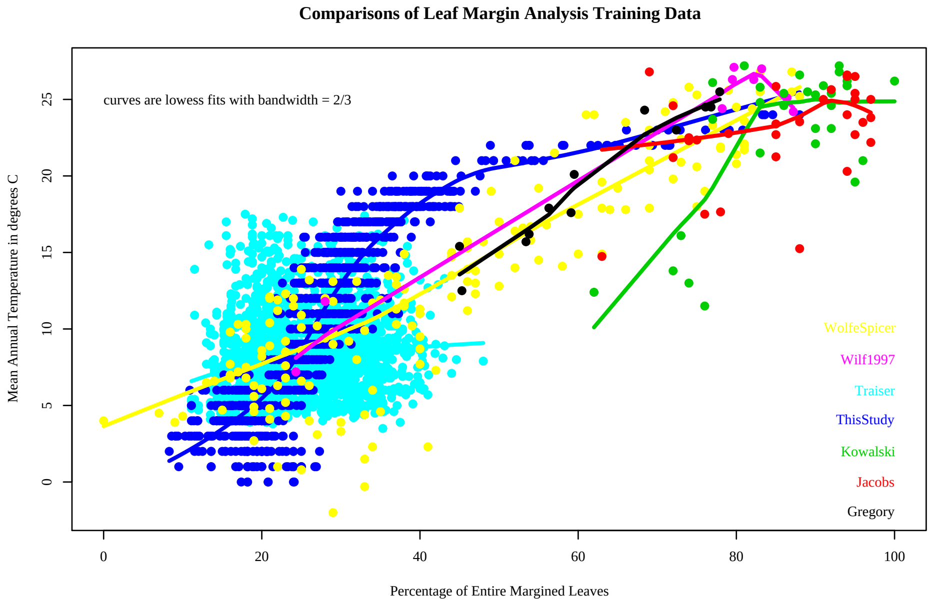


Figure 5.4: The results of the present study for eastern North America compared to other previously published studies

as percentage entire margin species. Since it is usually impossible to determine whether a fossil leaf was deciduous or not, from the point of view of paleoclimatic reconstruction the observation that deciduousness is not a better temperature predictor than teeth is irrelevant. It is more interesting to consider using percentage entire margins to predict deciduousness in the fossil record. This seems a reasonable procedure, though subject to the relatively broad scatter in the plot of percent broadleaf deciduous and percent entire in the lower left corner of figure 5.7.

Comparing the maps in figures 5.2 and 5.7, we can also see that while shading percentage entire produces relatively horizontal colored bands, shading percentage broadleaf evergreen shows crescent-shaped bands running north along both coasts of North America. This implies that coastal regions are identifiable by a particularly high ratio of evergreen species to entire species, in other words, near the coasts, there are more evergreen species than would be expected based on the number of entire species. Since deciduousness may be related to cold winter temperatures and physiological drought, this effect is not hard to explain by reference to the milder (more equable) and wetter climates near large bodies of water.

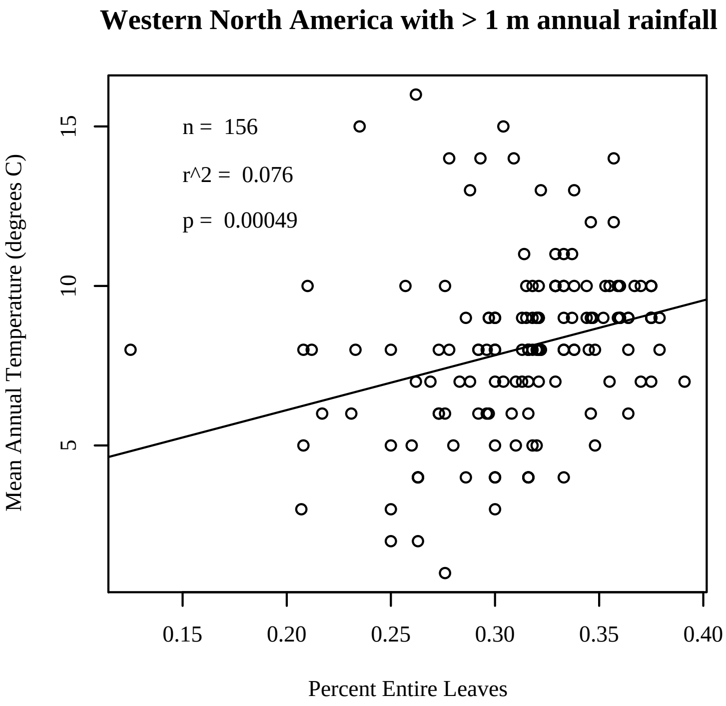


Figure 5.5: Scatterplot of margin proportion for areas in western North America with >1 m annual rainfall.

5.4 Discussion

The results obtained in this study in some respects confirm previous findings on leaf margin analysis. In other respects, however, they call for a reappraisal of earlier published conclusions on the nature of the temperature relationship.

On one hand, the positive relationship between entire leaf margin percentage and temperature is confirmed as being robust in the extensive forest area of eastern North America, between around 0 and 25°C. Given the scatter of the data points, it would appear possible to designate temperature to within about ±3 or 4°C between about 0 and 20°C. This is slightly less than previous studies had suggested (accuracy to within ±2.5°C), but still appears likely to give useful conclusions about paleotemperatures, especially in combination with other sources of evidence. The scatter decreases markedly at temperatures above 20°C, at least in eastern North America. In the range 20–25°C precision seems to be within ±1°C, though this apparent precision may partly be due to more

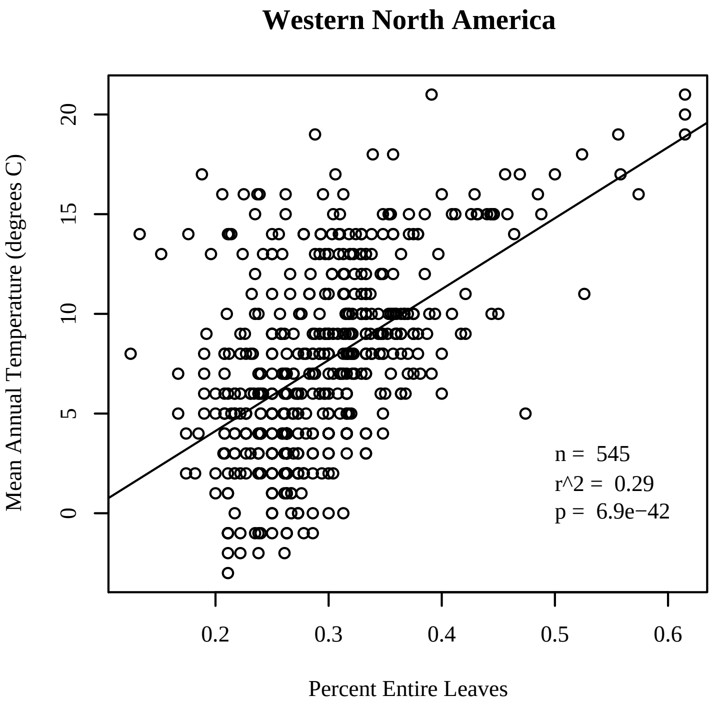


Figure 5.6: Scatterplot for areas in western North America regardless of annual rainfall.

homogenous edaphic environments in the warmer parts of the continent.

This spatial analysis reveals limits in the accuracy of temperature estimates due to scatter and non-linearity, which are not evident in smaller data sets. The arch away from the straight line at intermediate temperatures may call for reassessment of some earlier paleotemperature estimates from the fossil record that were based on leaf margin analysis. It is possible that the flattening of the curve that also brings about the arch relates in some way to the swampy soil conditions common in Florida and along the Gulf Coast. Burnham et al. (2001) and Kowalski and Dilcher (2003) have noted that swampy environments tend to have different proportions of entire-margined species than non-swampy environments. However, although swamps are widespread across southern Florida, this is by no means the case throughout northern Florida, or in Texas where the mean annual temperature is greater than 20°C, above the inflection point of the curve. Hence, it seems rather unlikely that swampy soils alone cause this trend.

For the western coast of North America, the lack of any strong correlation with temperature was

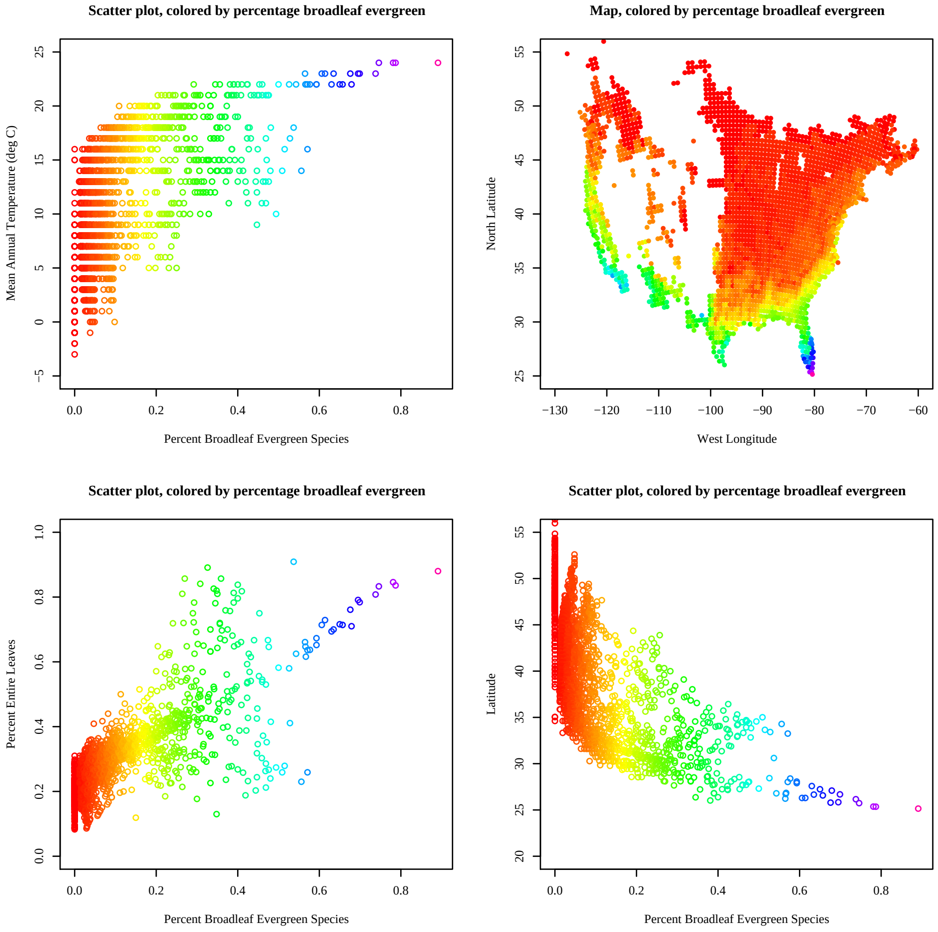


Figure 5.7: Map of percentage of broadleaf evergreen species (where there are more than 20 species to a grid cell) and scatter plots showing the bivariate relationships among percentage broadleaf evergreen, percentage entire, mean annual temperature, and latitude.

unexpected. Although the topography of this area is variable, is is surprising that the lowland flora failed to produce a detectable latitudinal trend. The first classic studies in leaf margin analysis by Wolfe (1979, 1993) were also based on data from western North America, and indicated a strong correlation between percentage entire leaf margins and temperature.

It is possible that the nature of the sampling in western North America by Wolfe (1993), on local deciduous stands of particular physiognomy, identified a linear relationship which is not evident when sampling is more thorough and regular. Wolfe pointed out that the lowland floodplain settings supporting forests in modern arid areas are just the sort of facies that preserve plant fossils. On the other hand, the vagaries of taphonomy and preservation of fossil leaf floras surely require that error analysis of the relationship between temperature and leaf margins use a wide range of forest types. Thus, the results presented here may provide a more realistic interpretion of the variation found in the fossil record. The lack of a broad scale effect in western North America may also be partly due to filtering of the flora. Species extinctions during glacial phases may have left behind fewer forms with entire margins (Adams and Faure 1997). In this sense, interpretation of Tertiary fossil leaf floras from western North America might be more accurate if based on training data from the eastern half of the continent, which is not so depleted.

Synthetic floras such as are produced here are different from species lists obtained from individual sites as previous authors have done. Therefore diversities will be inflated by counting range-through taxa, but there seems to be no reason why this should bias the proportion of leaves with entire margins. If anything, the broad scale sampling used here should be expected to yield a more accurate representation of true relationship between leaf form and temperature than sampling local sites, which are subject to the vagaries of microclimatic variation. The smaller sample numbers and rather more selective or *ad hoc* sampling used in previous studies may have disguised the true non-linear relationship apparent in eastern North America. It is unclear whether the correlation in eastern Asia would be the same if more data points from intermediate temperatures were added in.

If the non-linear functional relationship obtained here reflects the true relationship between leaf margin percentage and temperature, what are the implications? Many of the estimations of temperature from leaf margin percentage in fossil leaf floras may have to be altered, leading to different paleoclimatic conclusions. It appears from the scatterplot of eastern North American data in figure 5.3 that for mean annual temperatures between about 12 and 19°C, the earlier calibrations tended to underestimate temperatures by between 3 and 5°C. However, for warmer climates above about 20°C, the earlier linear models seem to have overestimated temperature by 1 or 2°C.

This study appears to confirm that leaf margin analysis can be useful in paleotemperature analysis in some areas, such as eastern North America. The relationship, however, may be unreliable in regions that are floristically depleted or have complex topography, such as the western half of the continent. In regions where the relationship is strong, it should be used with caution, because of non-linearity in the relationship and the evident scatter within the data. These findings emphasize that interpretation of palaeotemperatures from fossil leaf floras is best done in combination with other indicators (such as general floristics and a combination of other leaf characteristics).

6 Leaf architectural patterns across lineages: An ecophenetic examination of specimens in the the National Cleared Leaf Collection

Previous chapters have been concerned with how leaf architectural characteristics are distributed through time (chapter 2) and space (chapters 4 and 5). This chapter deals with their distribution across taxonomic and phylogenetic space. I explore the hypotheses that leaf architectural characters, insofar as they can be represented by Compendium Index Categories (CICs) as described in chapters 2 and 3, can be used to: (1) assist identification of unknown leaves to taxonomic groups and (2) reveal phylogenetic relationships among taxa.

6.1 Introduction

In this chapter, I will distinguish between *taxonomic* groups (defined here as named sets to which leaf specimens are assigned) and *phylogenetic* groups (defined here as monophyletic clades representing explicit hypotheses about the historical relationships between the plants that produced the leaf specimens). Note that it is perfectly possible (depending on how well or poorly taxonomic groups reflect phylogenetic relationships) for leaf characters to allow perfect identification of an unknown leaf to a taxonomic group while containing no phylogenetic signal and being useless for the purpose of revealing phylogenetic relationships. Hence, it is necessary to examine two related hypotheses reflecting taxonomic and phylogenetic attribution. Leaf identification to taxonomic groups does not depend on phylogenetic reconstruction of plant relationships so it is possible for hypothesis (1) to be accepted while (2) is rejected (the numbers 1 and 2 referring to the hypotheses given in the preceding paragraph). It is not logically possible, however to accept hypothesis (2) without also accepting (1). In other words, classification is necessary (but not sufficient) for phylogenetic reconstruction, while phylogenetic reconstruction is not necessary for classification. If (2) can be demonstrated, however, it is sufficient also to show (1).

The data to be examined comes from the Yale Peabody Museum holding of the *National Cleared Leaf Collection* (NCLC), which consists of about 6767 accessioned leaf specimens cleared, stained, and mounted between glass slides for easy examination by transmission light microscopy. Cleared leaves housed in the National Museum of Natural History at the Smithsonian Institution were not examined.

6.2 Data

Using the *Manual of Leaf Architecture* (Ash et al. 1999) and Green and Hickey (2005), each specimen in the NCLC was assigned from one to four CICs. All data were entered into the museum collections database, and this chapter is based on an examination of a version of the database exported on August 21, 2006, which can be found in appendix A.

The museum database was read into *R* (R Development Core Team 2005) and laundered using a script file, which can be found in appendix B. The data are recorded as specimens, and genus, species, and family are given for each entry in the museum database. As given, these are taxonomic categories that imperfectly reflect phylogenetic relationships. The families are from Cronquist (1981) with some additions and emendations from Takhtajan (1997). There are 306 families with at least one specimen that can be attributed to a single CIC.

Specimens that were too imperfectly preserved or ambiguous to be referred to a single CIC were removed from consideration. As further examination may show, it is possible that this cropping of the data may influence the conclusions, because variation is not distributed homogeneously across specimens. Therefore in order to compare families with complete statistical rigor, a different, hierarchically nested, stratified sampling program would be needed. Given the constraints of this data set, however, it is clearly not possible to examine variation at all hierarchical levels, so I have chosen to look at three levels in this initial analysis: specimen, family, and order. At the ordinal level, two alternative phylogenetic schemes (ways of assigning families to orders) are available, one from Cronquist (1981), the other from the Angiosperm Phylogeny Group (Stevens 2001–; Angiosperm Phylogeny Group 2003). The 6767 specimens in the database represent 3732 species and 1733 genera so a reevaluation of each attribution would be too time-consuming for the present study. Instead, I rely on the attributions to family as given and assume that species and genus sampling within family have not skewed results. Only superfamilial relationships are examined critically.

Family level is a particularly interesting level at which to look at variation because it has traditionally been where a relatively large proportion of angiosperm variation has been expressed. In other words, family-level groups have been more stable, less controversial, and more natural-seeming than groups at other levels of the hierarchy.

In order to discuss the distribution of leaf architectural characters across families, I will define the following terms: The family *mode* is the most commonly found CIC among specimens attributed to the family. Where there is a tie, the family is considered to have more than one mode. The family *spread* is the total number of CICs represented by all specimens attributed to the family. The family *max* is the number of specimens in the modal CIC.

If a good correspondence is found between CICs and families, *i.e.* if all the specimens in a family have leaves that fall into the same CIC, and if this CIC is different for each family, then it is easy to attribute an unknown leaf to a family. On the other hand if all the families have a wide spread of leaves across CICs or if the modal CICs in different families are the same, then it becomes difficult to identify unknown leaves to family.

Figure 6.1 is a bar plot showing the distribution of the spread for all 306 families represented in the data.

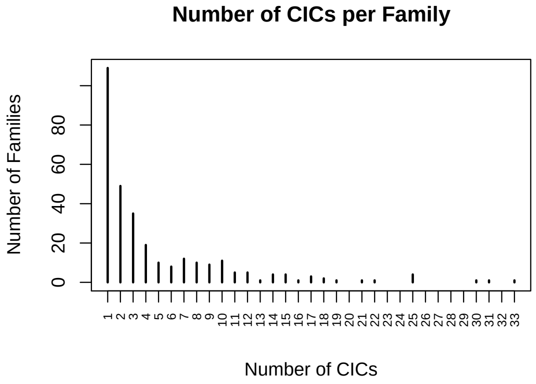


Figure 6.1: Bar plot showing how many CICs are represented by the leaves in each family in the NCLC. The heavy right skew of the distribution is at least partially an artifact of taxon sampling; *i.e.* it is not corrected (or rarified) by examining the same number of specimens from each family.

Figure 6.2 shows how the shape of the distribution in figure 6.1 is skewed by sampling. While figure 6.1 is heavily right-skewed, the rarified distributions in figure 6.2 (to sample sizes respectively of 10, 20, 30 and 40 specimens) show how most families have a spread across about 10 CICs. Assuming that the specimens are randomly chosen from within each family, at least 20 specimens are needed in order to estimate a family's spread.

A table of the modes, spreads, maxes and total number of specimens for each family can be found in the appendix A, as can other tabular summaries of the data used below.

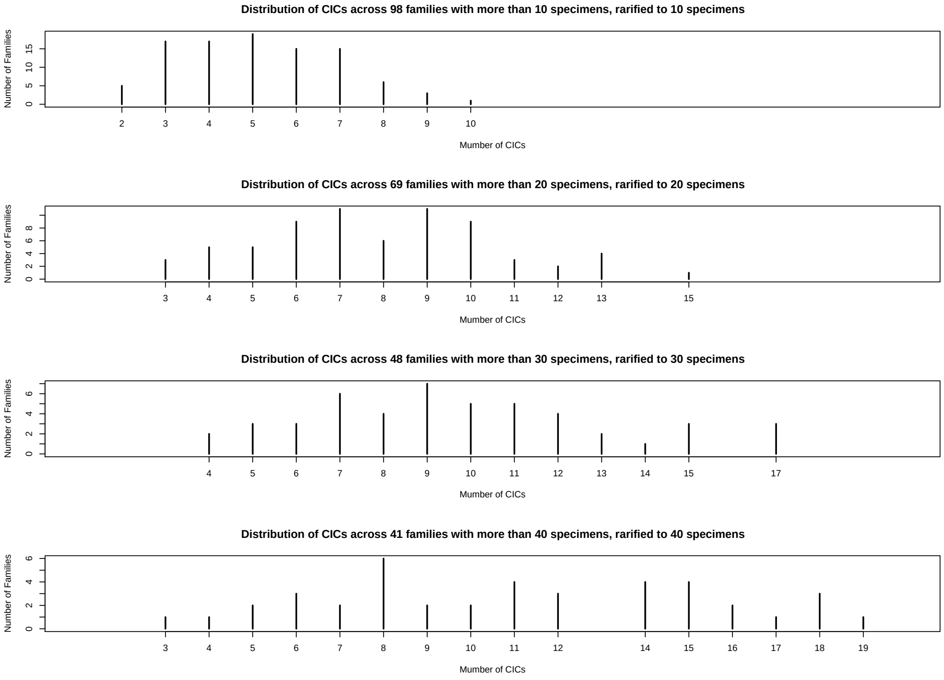


Figure 6.2: Bar plots showing how rarefaction affects apparent family spread

6.3 Results

To illustrate these data in more detail, I have chosen to focus on two families: Rosaceae and Roridulaceae. Rosaceae is the family with the largest number of specimens in the data examined, while Roridulaceae (next to Rosaceae in alphabetic order) is represented by only a single specimen. Figure 6.3 shows the distribution of specimens across CICs in these two families as ‘spectra’ or ‘profiles’ (cf. figure 4.6). The bar plots are not analogous to the summaries in the previous section, but show instead the actual frequency of specimens in each CIC for the two families.

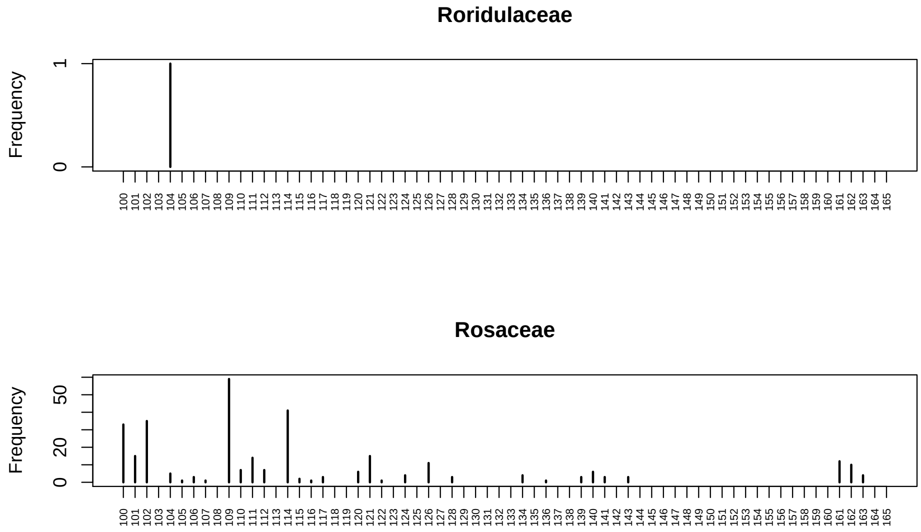


Figure 6.3: Profiles or spectra of CICs for two families.

Roridulaceae has a spread of 1, a mode at CIC104, a max of 1 and a total of 1 specimen. In contrast, Rosaceae has a spread of 30, a mode at CIC109, a max of 59 and a total of 313 specimens. Therefore both Rosaceae and Roridulaceae appear in figure 6.1 (Roridulaceae at the left end of the distribution as one of the 109 families with a spread of 1; Rosaceae near the right end as the single family with a spread of 30, the third-broadest spread). Roridulaceae does not appear in figure 6.2, however, because it is only represented by a single specimen so it is impossible to say from the available data very much about what its family spread would be if more specimens were examined. Rosaceae must appear in all four distributions in figure 6.2 because it has 313 specimens, but it is not clear where exactly it is in each distribution since a random selection of a subset of its 313 specimens was made to determine its rarified spread.

The data summaries in the previous section and the figure 6.3, while giving an idea of the likely accuracy of identifications to family of an unknown leaf do not provide enough information actually to assist such an identification. For instance, given an unknown leaf in CIC104, one would need to examine all 306 family profiles in order to see where the leaf might probably fall. This is a difficult graphical challenge to meet, but there is a better way of representing the same data, which somewhat simplifies the task.

For instance, in figure 6.4 the same data are shown as scatter plots in which the CICs on the x-axis are unchanged, but each point on the y-axis relates to a particular specimen, rather than (as in figure 6.3) showing the frequency of counts in each CIC.

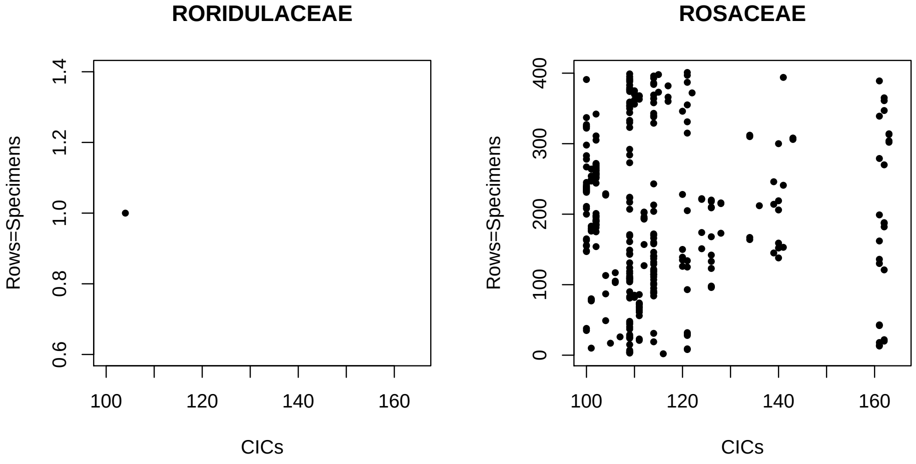


Figure 6.4: Scatter plots for two families showing the same data as figure 6.3 in an alternative representation.

In principle this strategy provides a way of representing all 6767 specimens on a single plot as shown in figure 6.5. Unfortunately, at reasonable resolutions it is difficult to identify a particular CIC in any of these little scatter plots, so the general picture is all that can be obtained. This general picture substantiates some of the claims made in the previous section based on data summaries: that most of the families have only a few specimens and that families with enough specimens to get an idea of the spread include leaves spread out over about 10 CICs.

See Plate 2

Figure 6.5: Scatter plots as in figure 6.4 for all 306 families in the NCLC.

How can this information help us to identify an unknown leaf? One route is to lump data together into a smaller number of groups than the 306 families; e.g. into Linnean orders. The defect of this is that interfamilial relationships (the composition of orders and subclasses) has historically been much less stable than the constitution of families. Many families, like the Graminae and Compositae, have been accepted taxonomic and phylogenetic groups since the 18th century, while superfamilial groups like the Hamamelidae and Dilleniidae, which were considered phylogenetically meaningful as recently as the 1980s (Cronquist 1981) have now been broken up and their parts redistributed into other groups. In order to try to allow for this instability in the constitution of superfamilial taxa, I have used both a recent system based largely on genetic data (Stevens 2001–; Angiosperm Phylogeny Group 2003) and an older system from Cronquist (1981).

In each case, I lumped together data from all the specimens in each family in the order and produced order profiles like the family profiles shown in figures 6.4 and 6.5. Instead of plotting them as bar plots (figure 6.4) or scatter plots (figure 6.5), however, I employed grayscales intermediate between black and white to indicate whether a particular CIC was densely occupied in a particular order. This matrix of orders by CIC weights can be seen as a grey-scaled image in figure 6.6 with (in addition) the orders clustered using a complete linkage hierarchical clustering algorithm with euclidean distance metric (for details see Everitt 1974; Hartigan 1975; Kaufman and Rousseeuw 1990; Gordon 1999) and the manual page for `hclust()` in R Development Core Team (2005). Before clustering, the data were also scaled using the Hellinger distances (between rows) (Oksanen 1983; Legendre and Gallagher 2001).

Note that the dendrogram on the left of the shaded spectra, which suggests 'relationships' between the orders does not reflect the phylogenetic hypothesis in Stevens (2001–). This is not unexpected: the data being examined are occupation of leaf architectural classes—an imperfect measure even of leaf architecture, much less of phylogenetic or historical relationships between plant lineages. The clustering structure here shows which orders have similar leaves (based on the available data), not how the orders are related. We expect to find cases where orders that are related do have similar leaves, but this systematic component to the leaf architectural signal is obscured by ecological influences or environmental plasticity, which may be much more influential.

Note that the lack of corresponedence between leaf architectural clusters and phylogenetic or systematic relationships is not limited to a particular system. Figure 6.6 shows equally poor correspondence to the super ordinal groupings postulated by Cronquist (1981).

At the scale of family, the same pattern can be observed, both in the case of the families attributed to order Rosales by APG (see figure 6.8) and in the case of Cronquist's system (figure 6.9).

In neither case are the relationships between families within the order Rosales well estimated

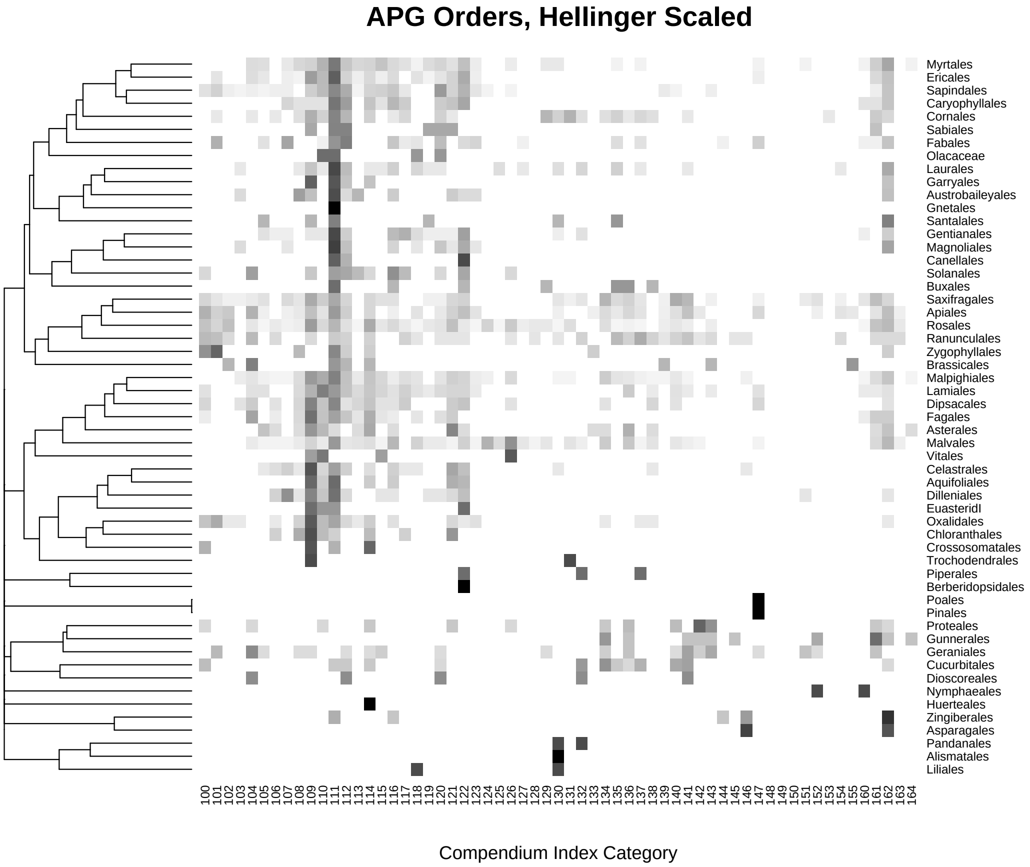


Figure 6.6: A shaded image showing CIC weights for the orders recognized by Stevens (2001–). Orders have been hierarchically clustered and reordered by a complete-linkage clustering algorithm using the euclidean distance measure for Hellinger-scaled CIC data.

by the leaf architectural groupings. This result effectively allows us to reject hypothesis (2) from the introduction. The data from leaf architecture that we have considered in this study does not by itself allow effective phylogenetic reconstruction of dicot relationships. This result should not be taken to suggest that leaf characters should be rejected from phylogenetic analyses. As can be seen from certain subsets of the data like the correct sister-relationship of Roridulaceae and Neuradaceae in figure 6.9, there clearly is a phylogenetic signal present. It is, however, confounded by other (probably environmental) signals such that (at the scale of orders within families or families within order Rosales) it is not strong enough to produce an accurate estimate of phylogenetic relationships.

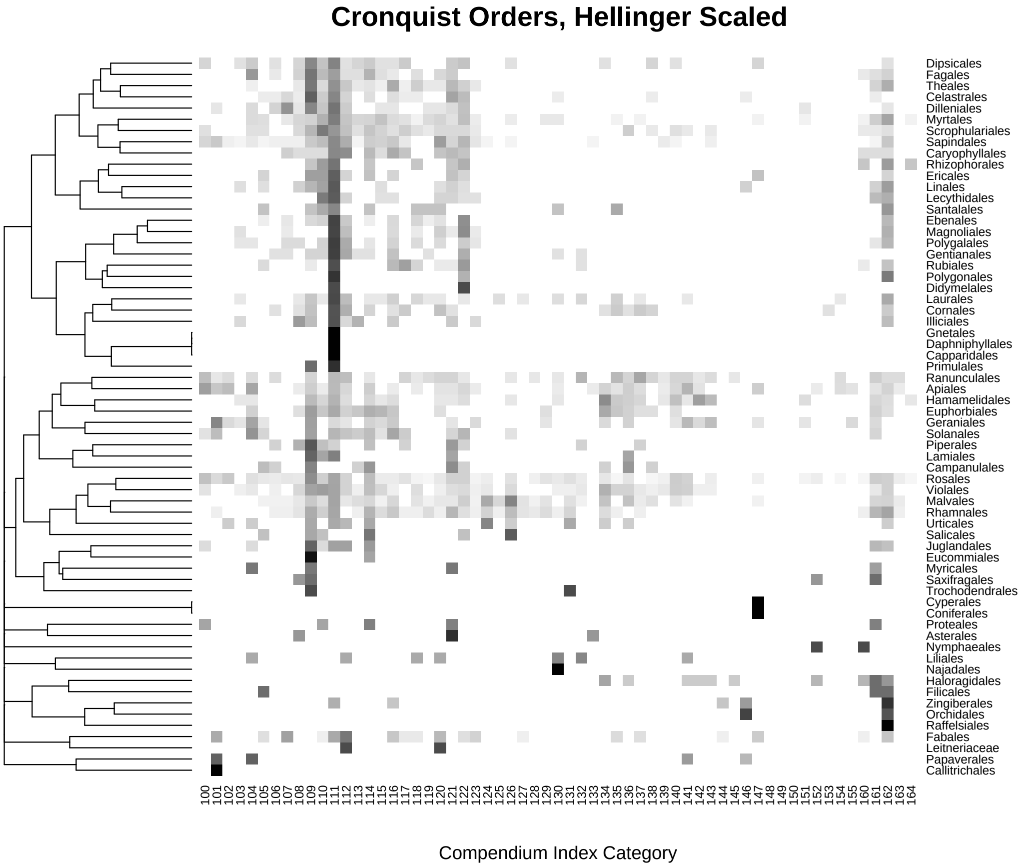


Figure 6.7: A shaded image showing CIC weights for the orders recognized by Cronquist (1981). Other than the choice of different ordinal assignations for the families in the NCLC, this shows the same data as are displayed in figure 6.6.

6.4 Discussion

So if ecophenetic data based on CICs cannot recover phylogenetic groups, what are they useful for? Given the task discussed in the previous section of identifying to taxon a leaf of unknown attribution in, for instance, CIC104, figures 6.6–6.9 take on a different meaning: the column of grey blocks above CIC104 becomes a representation of the probability of finding the leaf in each of the taxonomic groups. Where there are dark blocks, there is a high probability of finding similar leaves; where the area is light, there is a low probability of finding similar leaves. For instance, imagine a hypothetical scenario in which all leaves can be red or green (the color carrying no functional advantage or disadvantage). Purely by chance, some families and orders will have more red leaves

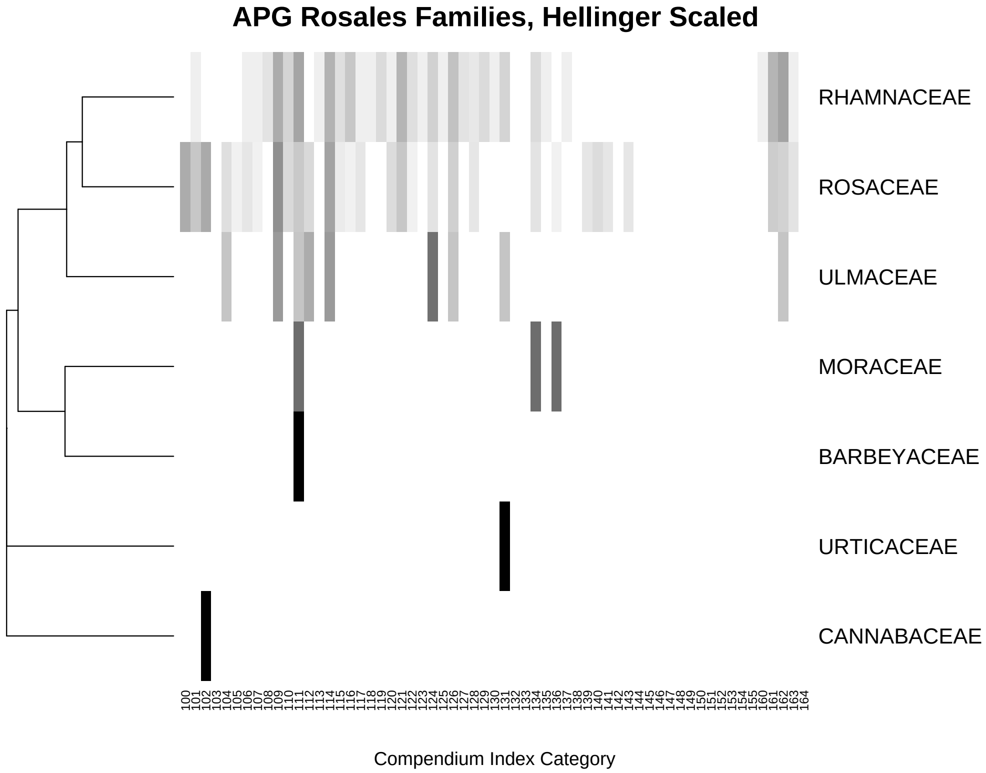


Figure 6.8: An examination of the families included in Rosales according to Stevens (2001–)

and some more green leaves, so if you want to identify a red leaf, it makes sense to look first in groups (orders and families) that have mostly red leaves. This is true unless red and green leaves are distributed uniformly across groups, which as can be seen from the figures is not the case.

For instance, to identify the leaf to a Cronquist order, reading along the column of shaded blocks above CIC104 reveals 21 possibilities, which are listed here beginning with the most likely: Papaverales, Myricales, Solanales, Fagales, Geraniales, Apiales, Liliales, Dipsicales, Urticales, Euphorbiales, Myrtales, Dilleniales, Juglandales, Ranunculales, Rosales, Scrophulariales, Polygalales, Hamamelidales, Sapindales, Violales, Theales.

Papaverales, having only 4 families (only one of which, Papaveraceae, has any leaves in CIC104), is boring to examine in more detail. The breakdown into families of Rosales, however, which is also reasobably high on the list of possibilities, is shown in figure 6.9. There the unknown leaf in CIC104 (provided it does belong in Rosales) is attributed to Roridulaceae or Neuradulacea or possibly to Pittosporaceae, but not to any of the other families.

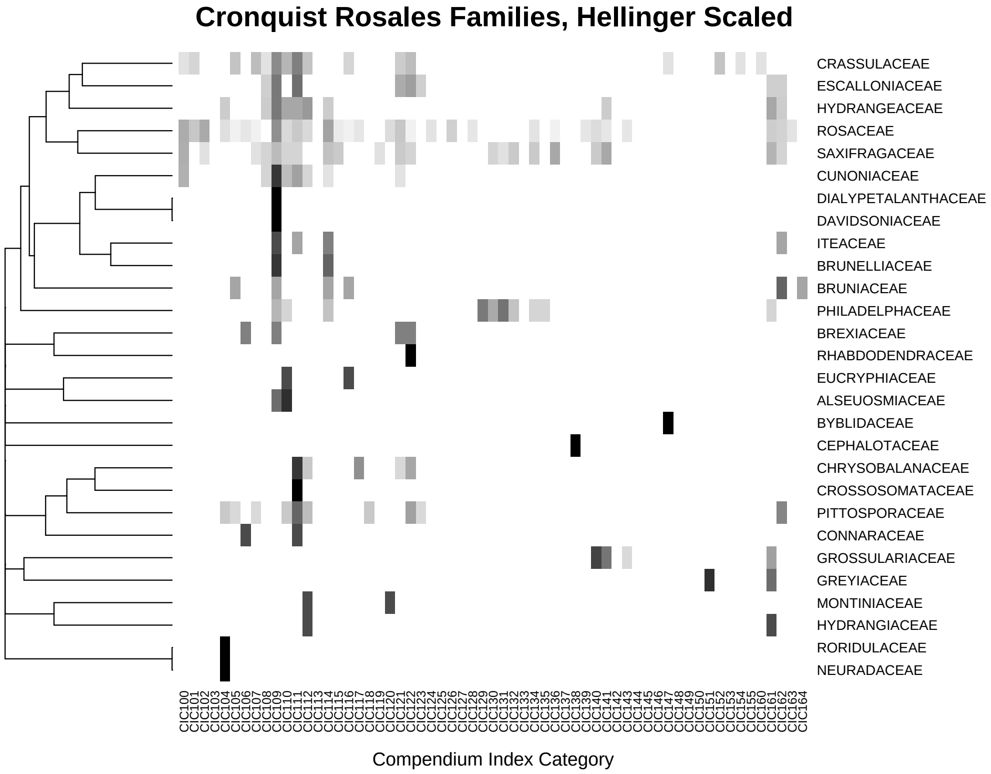


Figure 6.9: An examination of the families included in Rosales according to Cronquist (1981)

The same procedure can be used to identify an unknown leaf to APG taxonomic groups. The point is that taxonomic identification can be improved even if the contribution of phylogenetic history makes little or no contribution to the distribution of characters across groups. If the taxonomic groups used for identification match phylogenetic relationships perfectly, then taxonomic identification will also serve as phylogenetic analysis, but in order to demonstrate hypothesis (1) above, it is only necessary to show how leaf architectural data can be used to help assign leaves to well-defined morphological groups, not to demonstrate that these groups are also phylogenetically significant.

Another feature of the plots is the inclusion of orders (*e.g.* Gnetales in figure 6.7) that are not angiosperms and therefore cannot properly be classified using the CICs 100–164. As discussed in chapter 4, there are CICs, which are not examined here, intended to describe non-angiosperms. Also it should be pointed out that CICs 160–164 designate poorly-preserved specimens and therefore should not be relied upon for the purposes of classification or identification.

A final point to be noted is that the images shown in figures 6.6–6.9 can be considered probabilistic tabular polyclave keys. That is, they are similar to diagnostic keys laid out in a tabular format in which entry can be made from any character (as opposed to dichotomous keys, in which the order in which characters are examined is fixed), and in addition to showing the possibility of identification, they also provide a measure of how probable each possibility is.

6.5 Conclusions

From this data examination, we can conclude that leaf architectural data does provide a helpful tool for taxonomic identification. Although an unknown leaf will seldom be unambiguously attributable to a single family or order based on CICs alone, data summaries like figures 6.6–6.9 allow large numbers of families or orders to be ruled out in the process of identification. Moreover, subject to the constraints of sampling, these figures allow higher taxonomic groups to be ranked in order of which are most likely to contain a given leaf form (i.e. in which group are the greatest proportion of leaves of the given form found). It is not immediately clear, however, that this taxonomic identification contains a large amount of useful phylogenetic information. That is we can accept the first hypothesis but have only weak evidence supporting the second.

7 Classification based on structured factors

A problem that often appears in the social and natural sciences is the problem of clustering a set of objects based on a number of observed variables. This issue has been divided into so-called q-mode analysis (clustering the objects) and r-mode analysis (clustering the variables) and these modes of analysis can also be combined into 'two-way' clustering in which both variables and objects are clustered either independently or each by reference to the structure obtained for the other direction (Legendre and Legendre 1983). What characterizes all three of these methods is no assumption of *a priori* structure or weighting of objects or variables. The intent is to determine impartially if and where 'natural' or 'real' breaks appear in the data. There are many issues surrounding the definition of what a 'real' break may be, how clustering is affected by missing values and data pre-treatment, and what distance metric or clustering method should be used. We cannot discuss these issues here. Instead, this paper addresses a particular scenario in which we know more about the variables than about the objects. That is, we either have no prior classification of the objects or do not want our prior classification to affect our results, but we do know, and would like our clustering to be informed by, the way the variables are structured (which is frequently hierarchically). Note that this is more often the case than it might seem: there are very few cases in which we have absolutely no preconceived notions of interactions or covariation among a set of measured variables. The mere choice of certain variables implies some degree of classification, so to treat the variables chosen as if they are unstructured is usually to ignore important prior constraints on the data.

In many statistical packages and in particular in my analytic tool, the open-source programming language *R* (R Development Core Team 2005), it is conventional to represent objects as rows and variables as columns, so I will use 'row' synonymously with 'object' and 'column' with 'variable'. Specifically, therefore, we would like to to split up a matrix into strata or blocks, perhaps hierarchically structured, in a way that reveals potentially unexpected breaks in the rows, while adhering insofar as is reasonable to the known structure of the columns. Of course the distinction between 'objects' and 'variables' or 'rows' and 'columns' is entirely for the sake of convenience and clarity, because we have two sets of random variables that we would like to treat differently. Formally they are both random variables taking measured values from a sample space and there is no necessary correspondence between things that we think of as objects or variables in the physical world and the rows or columns to which these things are assigned for analysis. In practice, because of the asymmetry of the procedure outlined below, the random variable about which more prior knowledge is available should be assigned to the columns. In the case considered here, the values are all counts or count-equivalents (and the simulated data are therefore drawn from Poisson distributions), though with some variation, the same general procedure should also be applicable to continuous variables.

In this chapter I will use the following notation: a matrix or contingency table, M, takes values x_{ij} for i in n rows and j in m columns; the row marginal sums are $x_{i.}$; the column marginal sums are $x_{j.}$, and the sum of all counts is $x_{..}$. I am designing a procedure to determine where and when to partition this matrix; some desirable properties of this procedure are:

(1) The overall number of observations of each object or variable (the marginals, $x_{i.}$, and $x_{.j}$) should not unduly affect the clustering. This suggests using a metric in which the proportional rather than the absolute counts are used.

(2) The variables should not be treated as equivalent. It is the actual proportional counts in each variable, not the distribution of counts that is important, i.e. two rows with identical distributions of counts may be very dissimilar if the high counts and the low counts are in different columns.

(3) There should be a stopping rule that prevents the continued partitioning of blocks when there is no evidence that they have meaningful internal structure. We need, however, to be able to tune this stopping rule in order to allow the partitioning to proceed to a greater or lesser extent.

(4) There should be a graphical method of concisely displaying some combination of the raw data, its *a priori*-known structure, and the clusters that are derived from the procedure.

(5) Since we wish to ignore any *a priori* information we may have about the classification of the rows, we should reshuffle them as necessary in order to make better partitions in the matrix. In a case where the column structure is known, on the other hand, the column order must remain fixed.

In order to determine what partitions are 'better' a test statistic is needed to compare the actual counts x_{ij} with an alternative of multinomial independence. Because there are likely to be a number of empty cells, a chi-squared test may give unreliable results and a permutation test is computationally more expensive than a maximum likelihood approach. Therefore we use a likelihood ratio test to compare:

H_0 : X is distributed in proportion to the independent row and column marginals: each row has the same conditional distribution $p_{ij} = p_{i.}p_{.j}$ with likelihood:

$$L(x_{ij}|\hat{\theta}_0) = \frac{x_{..}!(\frac{x_{i.}x_{.j}}{x_{..}^2})^{x_{ij}}}{\prod_{ij} x_{ij}} \tag{1}$$

with:

H_A : $X \sim$ Multinomial$(x_{..}, p_{ij})$, i.e. $x_{..}$ samples independently drawn with replacement from population X with positive probabilities p_{ij} summing to 1, and likelihood:

$$L(x_{ij}|\hat{\theta}) = \frac{x_{..}!}{x_1!....x_k!}{p_1}^{x_1} \cdots {p_k}^{x_k} = \frac{x!}{\prod x_{ij}!}\prod_{ij} {p_{ij}}^{x_{ij}} \tag{2}$$

This gives a likelihood ratio test statistic,

$$\lambda = 2log\frac{L(\hat{\theta})}{L(\hat{\theta}_0)} = 2log\frac{\prod_{ij} {p_{ij}}^{x_{ij}}}{\prod_{ij}(p_{i.}p_{.j})^{x_{ij}}} \tag{3}$$

Taking logs and simplifying, we get a proportional log likelihood,

$$pll = \sum_{ij} x_{ij}(log\frac{x_{ij}}{x_{..}} - log\frac{x_{i.}}{x_{..}} - log\frac{x_{.j}}{x_{..}}) \tag{4}$$

Like the chi-squared test statistic, this will be be 0 if all the rows (or columns) have proportionally identical counts and its distribution will be asymptotically chi-squared on $(n-1)(m-1)$ degrees of freedom as $x_{..} \rightarrow \infty$, provided $x_{i.} > 0$, $x_{.j} > 0$

7.1 Algorithms

Echoing approaches found in Hartigan (1975), code has been written to do two separate things with the contingency table, M.

First of all, the function `part()` partitions or clusters the rows of M in such a way as to minimize the test statistic, *pll*. The algorithm picks a random starting assignment of the rows to two clusters and then tries each row in turn in each cluster, moving rows about whenever a smaller value of the test statistic is discovered. When the test statistic stops changing (by more than a small, settable tolerance ϵ) or when a maximum number of cycles through all the rows and clusters (defaulting to 5 complete cycles), the algorithm adds in an additional cluster and continues to optimize. The process ends when all of the clusters could have occurred by chance at the assigned significance level or when there are $n-1$ clusters. Though by default there is only a single initial randomization, a greater number of random starts can be chosen, so as to avoid being caught in a local minimum. Options are provided for several levels of verbose output, and for the production of plots in which the rows are reordered and colored appropriately to reflect the clustering that has been produced. Although the details differ, the intent of this function is similar to the 'one-way direct splitting' algorithm described by Hartigan (1975).

This optimization uses all the variables (which are implicitly weighted equally) and only partitions the input matrix horizontally, so the function `part.recursive()` provides a wrapper that first partitions its input matrix horizontally, but then, when all the remaining blocks could have occurred by chance, looks at subgroups of variables given by any known structure in the columns to see if there are sensible subsets of columns that will partition blocks of rows that could not be partitioned using all of the columns. The graphic output of `part.recursive()` is a series of nested sub-matrices together supplying the complete, hierarchical partition of the rows of the original matrix, M.

The second main function is provided by `mondrian()`, which takes a given row and column structure and compares the available vertical and horizontal partitions of M (based on the structures of the rows and columns), in each case making the split that is more likely. It uses the same stopping rule (remaining blocks could have occurred by chance at a given significance level), but behaves symmetrically with respect to rows and columns. That is, it splits either horizontally or vertically, starting with the deepest nodes in the given structures of the rows and columns, and continues to split until all the nodes have been tried or the stopping rule is met. This is analogous to 'two-way direct splitting' Hartigan (1975). Note also that for convenience (because the results of `part()` have been found to agree well with hierarchical clusters), `mondrian()` assigns a tree structure to either of the variables (rows or columns) that are not assigned a prior structure in the function call.

The graphical output of `mondrian()` consists of a plot of the matrix partitioned by lines of varying thickness, which show the hierarchical levels at which the partitions were made (figure 7.3) producing an overall effect reminiscent of the later modernist paintings of Piet Mondrian (1872–1944).

The functionality supplied by `part.recursive()` and `mondrian()` are meant to be applied in sequence: `part.recursive()` supplies a posterior structure for the rows of an imput matrix, which can then be compared with the prior column structure via `mondrian()`. Note, however, that `mondrian()` is algorithmically symmetric so it can also be used to compare rows and columns both of whose structures are known *a priori.* Note that this sequencing is not yet available; the output of `part.recursive()` is not in the correct form to be passed to `mondrian()`, a defect that is in the process of being corrected.

Subsidiary helper functions used by all of these are `pll.calc()`, which calculates the test statistic for a matrix, `null.check()`, which compares this test statistic with a chi-squared probability, and two functions (`sort.rows()` and `merge2matrix()`) that are needed to translate between parenthetical tree format (e.g. *((A,B),C)*), binary tree format (e.g.

$$\begin{array}{r|ccc} & A & B & C \\ root & 1 & 1 & 1 \\ node1 & 1 & 1 & 0 \\ terminal1 & 1 & 0 & 0 \\ terminal2 & 0 & 1 & 0 \\ terminal3 & 0 & 0 & 1 \end{array}$$

), and `merge` objects (e.g. for $j_A = 1, j_B = 2, j_C = 3$),

$$\begin{array}{r|cc} i = 1 & -1 & -2 \\ i = 2 & 1 & -3 \end{array}$$

the form in which hierarchical trees are stored in *R*. (For a more detailed descriptions of this form, see the man page for `hclust()` in (R Development Core Team 2005). In general, a binary tree format representation of n terminals will have a maximum of $2n - 1$ rows and will have exactly $2n - 1$ rows if and only if it is strictly dichotomously branching.

These functions are provided in appendix B and can be read directly into *R* using `source()`. They should run directly in *R* 2.0.1 under Mac 0S 10.2; the only know alternation that is necessary to run them on other platforms is replacing the graphics driver `quartz()` with an appropriate non-Macintosh-specific driver. Unfortunately at this stage documentation consists only of comments in the source code.

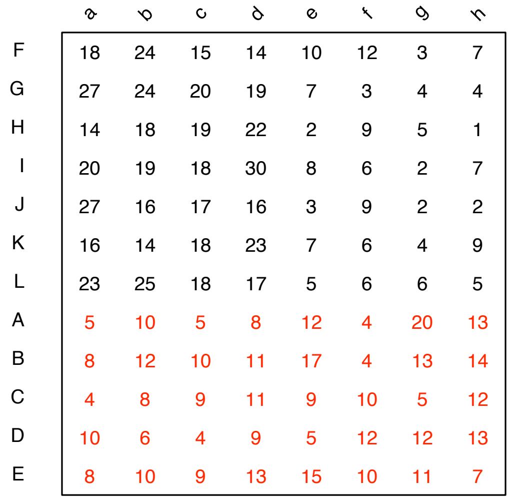
7.2 Simulation

Before these functions are applied to actual data, their operation is best explained by applying them to a small simulated matrix. Consider a simple 8×12 contingency table in which the counts in the first 5 rows are randomly drawn from a Poisson(10), the left half of the bottom 7 rows are drawn from a Poisson(20) and the right half of the bottom 7 rows are drawn from a Poisson(5).

	a	b	c	d	e	f	g	h
A	5	10	5	8	12	4	20	13
B	8	12	10	11	17	4	13	14
C	4	8	9	11	9	10	5	12
D	10	6	4	9	5	12	12	13
E	8	10	9	13	15	10	11	7
F	18	24	15	14	10	12	3	7
G	27	24	20	19	7	3	4	4
H	14	18	19	22	2	9	5	1
I	20	19	18	30	8	6	2	7
J	27	16	17	16	3	9	2	2
K	16	14	18	23	7	6	4	9
L	23	25	18	17	5	6	6	5

The function `part()` reliably divides this matrix into two strata as shown by the colors in figure 7.1.

Note that the clusters produced by `part()` are not dissimilar to those produced by a hierarchical algorithm using complete linkage and a euclidean distance measure, as shown in figure 7.2. Though in the case of the `part()` output, the default significance level of 0.05, allows retention of unresolved



	a	b	c	d	e	f	g	h
F	18	24	15	14	10	12	3	7
G	27	24	20	19	7	3	4	4
H	14	18	19	22	2	9	5	1
I	20	19	18	30	8	6	2	7
J	27	16	17	16	3	9	2	2
K	16	14	18	23	7	6	4	9
L	23	25	18	17	5	6	6	5
A	5	10	5	8	12	4	20	13
B	8	12	10	11	17	4	13	14
C	4	8	9	11	9	10	5	12
D	10	6	4	9	5	12	12	13
E	8	10	9	13	15	10	11	7

Figure 7.1: A partition of M into strata or clusters of rows based on the `part()` function.

polychotomies when there is insufficient evidence to partition clusters further, while the hierarchical algorithm provides clusters even when they are not statistically supportable.

Application of `mondrian()` to the same matrix (also with a default significance level of 0.05) produces the partition shown in figure 7.3. In this case, the tree structures associated with both rows and columns are produced automatically (using a complete linkage algorithm). The resulting row tree is shown to the left of the main matrix and the column tree is depicted above the main matrix as batteries of (respectively) vertical and horizontal lines. The matrix of counts is divided by lines whose relative thickness corresponds to how deep the split appears in its respective marginal tree structure. Thus, the first split that is made, between the 5th and 6th rows of the original matrix, is tried first because it is closest to the root of the *a priori* dendritic structure of the rows, for which reason it also receives the heaviest line weight.

To see how the algorithm can be tuned, we can experiment with both lower and higher (figure 7.4) significance levels. As is apparent from the figures, the degree of resolution (how far the partioning proceeds) can be set by the adjusting the level of probability at which a block is considered to have no meaningful internal structure.

If the significance level is set to 1, of course, partitioning can proceed until no block has more than a single row or column, as is shown in figure 7.5

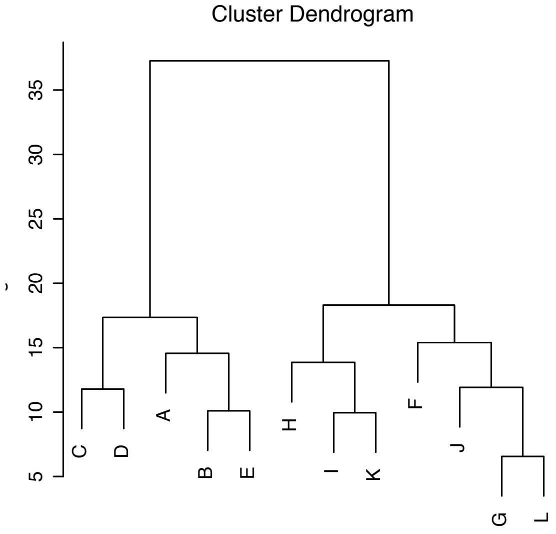
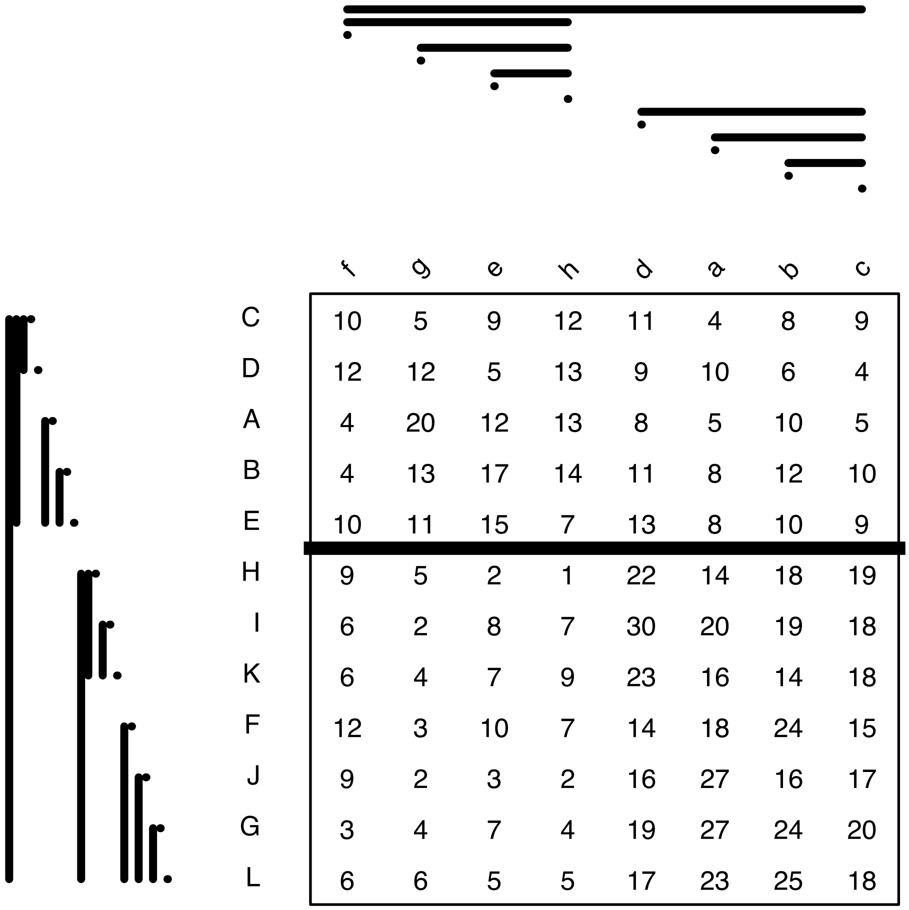


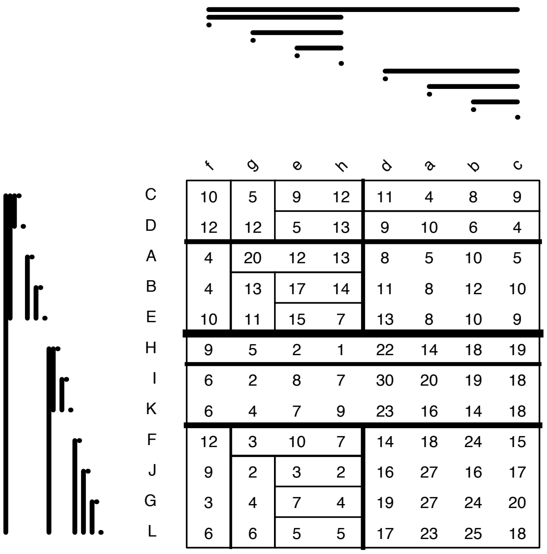
Figure 7.2: Complete linkage euclidean hierarchical clustering of the rows of M.

This output on an artificial matrix gives us some confidence that the algorithm is producing sensible results; in the following section we see how it responds to real data.



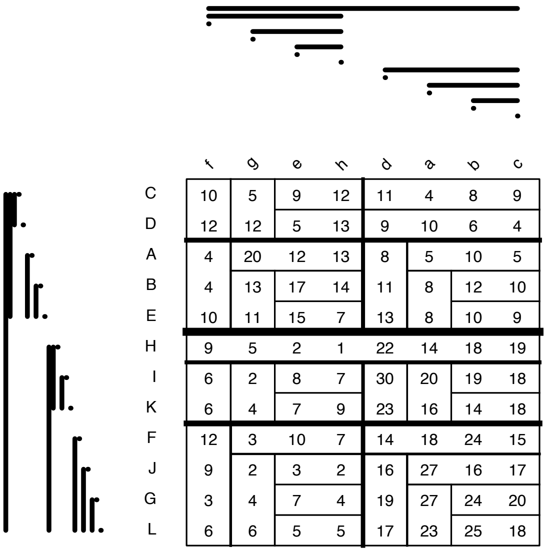
	f	g	e	h	d	a	b	c
C	10	5	9	12	11	4	8	9
D	12	12	5	13	9	10	6	4
A	4	20	12	13	8	5	10	5
B	4	13	17	14	11	8	12	10
E	10	11	15	7	13	8	10	9
H	9	5	2	1	22	14	18	19
I	6	2	8	7	30	20	19	18
K	6	4	7	9	23	16	14	18
F	12	3	10	7	14	18	24	15
J	9	2	3	2	16	27	16	17
G	3	4	7	4	19	27	24	20
L	6	6	5	5	17	23	25	18

Figure 7.3: Mondrianesque partition of M at the (default) significance level of 0.05.



	f	g	e	h	d	a	b	c
C	10	5	9	12	11	4	8	9
D	12	12	5	13	9	10	6	4
A	4	20	12	13	8	5	10	5
B	4	13	17	14	11	8	12	10
E	10	11	15	7	13	8	10	9
H	9	5	2	1	22	14	18	19
I	6	2	8	7	30	20	19	18
K	6	4	7	9	23	16	14	18
F	12	3	10	7	14	18	24	15
J	9	2	3	2	16	27	16	17
G	3	4	7	4	19	27	24	20
L	6	6	5	5	17	23	25	18

Figure 7.4: Mondrianesque partion of M at a significance level of 0.9.



	f	g	e	h	d	a	b	c
C	10	5	9	12	11	4	8	9
D	12	12	5	13	9	10	6	4
A	4	20	12	13	8	5	10	5
B	4	13	17	14	11	8	12	10
E	10	11	15	7	13	8	10	9
H	9	5	2	1	22	14	18	19
I	6	2	8	7	30	20	19	18
K	6	4	7	9	23	16	14	18
F	12	3	10	7	14	18	24	15
J	9	2	3	2	16	27	16	17
G	3	4	7	4	19	27	24	20
L	6	6	5	5	17	23	25	18

Figure 7.5: Mondrianesque partion of M at a significance level of 1; that is, fully resolved.

7.3 Data

A brief historical digression is necessary to provide a context for the type of data for which this analytical procedure is designed. The historical roots of ecology lie in what was at the time called botany but is now referred to as plant geography, phytogeography or phytosociology (McIntosh 1985). Basically, pre-twentieth century plant geographers or ecologists were interested in categorizing and explaining why different areas produced what seemed to be definite, predictable types of vegetation. Until the mid–late nineteenth century , most plant geography was descriptive and anecdotal (e.g. von Humboldt 1807; de Candolle 1855), and this general type of work has continued through the present, especially outside of mainstream Anglo-American plant ecology (Zohary 1973; Takhtajan 1986). As ecology constituted itself as a quantitative science, however, it became increasingly important to justify vegetation classification with an explicit methodology. Among many schools of vegetation classification that flourished around the end of the 19th through the first half of the 20th centuries, the best known is the Montpellier (or Zurich-Montpellier) School and the most characteristic version of its method was developed by the Swiss botanist Braun-Blanquet (1965). For extensive comparison of the different schools of vegetation classification, see Whittaker (1962), and Shimwell (1971).

The Braun-Blanquet method of vegetation classification seems very antiquated and subjective by the standards of modern anglophone plant ecology, but retains a data-centered and visual basis that is worth preserving. It is essentially an ordination procedure intended to operate on a sites-by-species matrix. The columns of the matrix represent a number of sites or localities (technically referred to as stands, *aufnahme,* or *relevés*), and the rows represent the species found at the sites. The Braun-Blanquet procedure consists of recording a list of species for each stand often with associated cover or abundance estimates, usually ranked on a 1 to 5 scale, but sometimes recording only presence or absence of species. The initial raw data is then recorded as in figure 7.6.

The species (rows) and sites (columns) are then reordered so as to produce blocks in the matrix of numbers that are uniformly high (or at least non-zero). This procedure seems always to have been subjective and based on trial-and-error. The species (rows) are constrained by prior knowledge about the (genetic) relationships among species as well as known patterns of ecology and life-form, while the sites (columns) were, at least in theory, supposed to be objectively classified by the patterns emerging from co-occurrence of groups of species. The final classification of sites (shown in figure 7.7) would then be represented by the re-ordered matrix, with formal names applied to the blocks of sites that had emerged from the analysis as characterized by the associated blocks of species.

The partioning procedure described in the previous sections represents an attempt to generalize and systematize a methodology similar to the Braun-Blanquet method, but more statistically justifiable and applicable to a broader group of scientific and statistical questions. To explore the

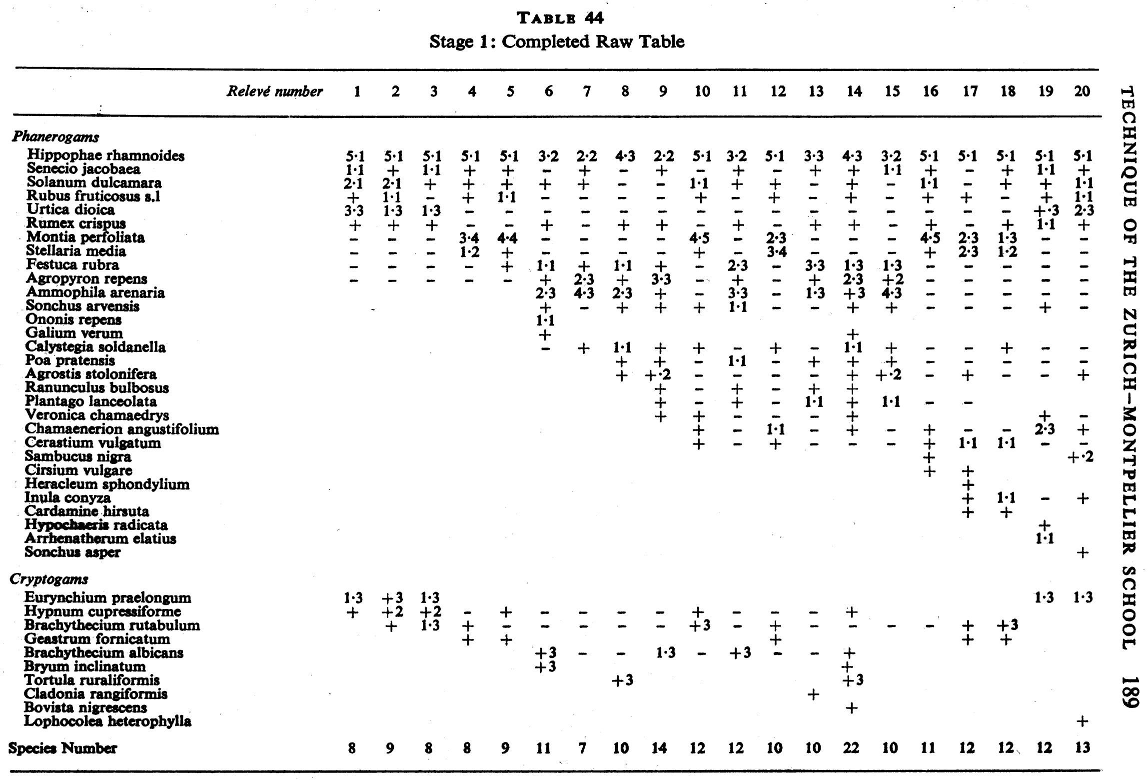
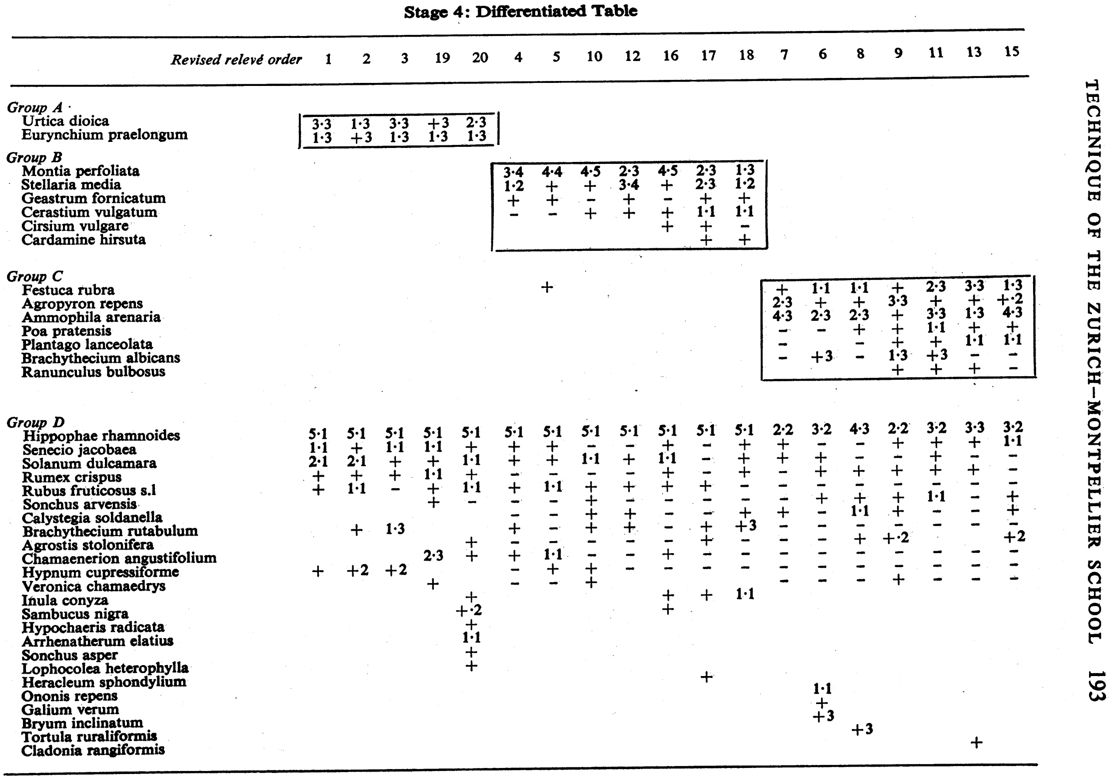


TABLE 44

Stage 1: Completed Raw Table

Relevé number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Phanerogams																				
Hippophae rhamnoides	5·1	5·1	5·1	5·1	5·1	3·2	2·2	4·3	2·2	5·1	3·2	5·1	3·3	4·3	3·2	5·1	5·1	5·1	5·1	5·1
Senecio jacobaea	1·1	+	1·1	+	+	−	+	−	+	−	+	−	+	+	1·1	+	−	+	1·1	+
Solanum dulcamara	2·1	2·1	+	+	+	+	+	−	−	1·1	+	+	−	+	−	1·1	−	+	+	1·1
Rubus fruticosus s.l	+	1·1	−	+	1·1	−	−	−	−	+	−	+	−	+	−	+	+	−	+	1·1
Urtica dioica	3·3	1·3	1·3	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+·3	2·3
Rumex crispus	+	+	+	−	−	+	−	+	+	−	+	−	+	+	−	+	−	+	1·1	+
Montia perfoliata	−	−	−	3·4	4·4	−	−	−	−	4·5	−	2·3	−	−	−	4·5	2·3	1·3	−	−
Stellaria media	−	−	−	1·2	+	−	−	−	−	+	−	3·4	−	−	−	+	2·3	1·2	−	−
Festuca rubra	−	−	−	−	+	1·1	+	1·1	+	−	2·3	−	3·3	1·3	1·3	−	−	−	−	−
Agropyron repens	−	−	−	−	−	+	2·3	+	3·3	−	+	−	+	2·3	+2	−	−	−	−	−
Ammophila arenaria						2·3	4·3	2·3	+	−	3·3	−	1·3	+3	4·3	−	−	−	−	−
Sonchus arvensis						+	−	+	+	+	1·1	−	−	+	+	−	−	−	+	−
Ononis repens						1·1														
Galium verum						+								+						
Calystegia soldanella						−	+	1·1	+	+	−	+	−	1·1	+	−	−	+	−	−
Poa pratensis								+	+	−	1·1	−	+	+	+	−	−	−	−	−
Agrostis stolonifera								+	+·2	−	−	−	−	+	+·2	−	+	−	−	+
Ranunculus bulbosus									+	−	+	−	+	+						
Plantago lanceolata									+	−	+	−	1·1	+	1·1	−	−			
Veronica chamaedrys									+	+	−	−	−	+					+	−
Chamaenerion angustifolium										+	−	1·1	−	+	−	+	−	−	2·3	+
Cerastium vulgatum										+	−	+	−	−	−	+	1·1	1·1	−	−
Sambucus nigra																+				+·2
Cirsium vulgare																+	+			
Heracleum sphondylium																	+			
Inula conyza																	+	1·1	−	+
Cardamine hirsuta																	+	+		
Hypochaeris radicata																			+	
Arrhenatherum elatius																			1·1	
Sonchus asper																				+
Cryptogams																				
Eurynchium praelongum	1·3	+3	1·3																1·3	1·3
Hypnum cupressiforme	+	+2	+2	−	+	−	−	−	−	+	−	−	−	+						
Brachythecium rutabulum		+	1·3	+	−	−	−	−	−	+3	−	+	−	−	−	−	+	+3		
Geastrum fornicatum				+	+							+					+	+		
Brachythecium albicans						+3	−	−	1·3	−	+3	−	−	+						
Bryum inclinatum						+3								+						
Tortula ruraliformis								+3						+3						
Cladonia rangiformis													+							
Bovista nigrescens														+						
Lophocolea heterophylla																				+
Species Number	8	9	8	8	9	11	7	10	14	12	12	10	10	22	10	11	12	12	12	13

Figure 7.6: The rough, starting matrix of a Braun-Blanquet vegetation classification. Each entry in the matrix consists of a pair of numbers, of which the first is the abundance (percentage cover) and the second the sociability (clumping), both on a scale of 1 to 5. This example is from Shimwell (1971).



Stage 4: Differentiated Table

Revised relevé order	1	2	3	19	20	4	5	10	12	16	17	18	7	6	8	9	11	13	15
Group A																			
Urtica dioica	3·3	1·3	3·3	+3	2·3														
Eurynchium praelongum	1·3	+3	1·3	1·3	1·3														
Group B																			
Montia perfoliata						3·4	4·4	4·5	2·3	4·5	2·3	1·3							
Stellaria media						1·2	+	+	3·4	+	2·3	1·2							
Geastrum fornicatum						+	+	–	+	–	+	+							
Cerastium vulgatum						–	–	+	+	+	1·1	1·1							
Cirsium vulgare										+	+	–							
Cardamine hirsuta											+	+							
Group C																			
Festuca rubra							+						+	1·1	1·1	+	2·3	3·3	1·3
Agropyron repens													2·3	+	+	3·3	+	+	+·2
Ammophila arenaria													4·3	2·3	2·3	+	3·3	1·3	4·3
Poa pratensis													–	–	+	+	1·1	+	+
Plantago lanceolata													–		–	+	+	1·1	1·1
Brachythecium albicans													–	+3	–	1·3	+3	–	–
Ranunculus bulbosus																+	+	+	–
Group D																			
Hippophae rhamnoides	5·1	5·1	5·1	5·1	5·1	5·1	5·1	5·1	5·1	5·1	5·1	5·1	2·2	3·2	4·3	2·2	3·2	3·3	3·2
Senecio jacobaea	1·1	+	1·1	1·1	+	+	+	–	–	+	–	+	+	–	–	+	+	+	1·1
Solanum dulcamara	2·1	2·1	+	+	1·1	+	+	1·1	+	1·1	–	+	+	+	–	–	+	–	–
Rumex crispus	+	+	+	1·1	+	–	–	–	–	+	–	+	–	+	+	+	+	+	–
Rubus fruticosus s.l	+	1·1	–	+	1·1	+	1·1	+	+	+	+	–	–	–	–	–	–	–	–
Sonchus arvensis				+	–	–	–	+	–	–	–	–	–	+	+	+	1·1	–	+
Calystegia soldanella						–	–	+	+	–	–	+	+	–	1·1	+	–	–	+
Brachythecium rutabulum		+	1·3			+	–	+	+	–	+	+3	–	–	–	–	–	–	–
Agrostis stolonifera					+	–	–	–	–	–	+	–	–	–	+	+·2			+2
Chamaenerion angustifolium				2·3	+	+	1·1	–	–	+	–	–	–	–	–	–	–	–	–
Hypnum cupressiforme	+	+2	+2			–	+	+	–	–	–	–	–	–	–	–	–	–	–
Veronica chamaedrys				+		–	–	+					–	–	–	+	–	–	–
Inula conyza					+					+	+	1·1							
Sambucus nigra					+·2					+									
Hypochaeris radicata					+														
Arrhenatherum elatius					1·1														
Sonchus asper					+														
Lophocolea heterophylla					+														
Heracleum sphondylium											+								
Ononis repens														1·1					
Galium verum														+					
Bryum inclinatum														+3					
Tortula ruraliformis															+3				
Cladonia rangiformis																		+	

Figure 7.7: The resulting Braun-Blanquet vegetation classification. This partitioning is obtained by reordering the rows and columns of the matrix in figure 7.6 by trial and error until blocks of values are discovered. Note that this example is conceptually the transpose of my procedure: more prior information is available about the rows and the new classification sought is of the columns.

effectiveness of this methodology when applied to actual questions of vegetation classification, we have collected two different sets of variables for encoding the morphological or architectural information in the leaves found at particular sites (these are referred to as *floras* instead of *stands* for contingent historical reasons).

The first set of data is based on a 31-variable method of coding leaves designed by Wolfe (1993). Strictly, the data that have been published are already percentages (of leaves having a particular characteristic) rather than true counts. The practice—strongly to be deprecated—has been not to publish the raw data but only this proportional digest. Nevertheless, sample size is constant enough that we can treat these percentages as if they were counts. We have an updated version of Wolfe's data, which was available for download as of April, 2005 at <http://www.open.ac.uk/earth-research/spicer/CLAMP/MET3AR.xls> (the data file Wolfe173.asc), a subset of this data set (Yakusugi.asc), and three other data sets published independently but using the same variables: Jacobs.asc (Jacobs 1999, 2002), Kowalski.asc (Kowalski 2002), and Gregory.asc (Gregory-Wodzicki 2000). Note that Jacobs.asc uses an earlier form of the data missing two variables.

The second set of variables is made up of 56 categories for which true counts are available. An as-yet unpublished matrix of counts is to be found in Green.asc and the particular subset of this data that we will examine here is given as Hawaii.asc (Green and Hickey 2004).

The *a priori* structures of these two sets of variables are stored as CLAMP.struct.asc and CIC.struct.asc respectively.

Calling `mondrian()` on the Hawaii.asc data with the appropriate column structure, CIC.struct.asc, produces figure 7.8. The only case in which a vertical split was made in preference to the horizontal splits occurs in a position that had already been identified in exploratory analysis as showing an important distinction: the compound leaves in the first three variables are known to be characteristic of the understory foliage of the bottom four floras.

7.4 Remaining Issues

First of all, the code attached in appendix B is poorly documented and does not interact smoothly; `part()`, which was written first, is unnecessarily slow and produces output that cannot be directly passed to `mondrian()` as a prior constraint.

Second, statistically the entries in the cells of the contingency table M have been treated like counts. In the CIC data (like the Hawaii.asc data set), they are in fact counts, but frequently they are so low and include so many zeros that they seldom or never produce statistically significant partitions. (For illustrative purposes, the counts shown in figure 7.8 were inflated so that statistically significant partitions would be made. In fact, the raw Hawaii.asc matrix is rejected by the algorithm

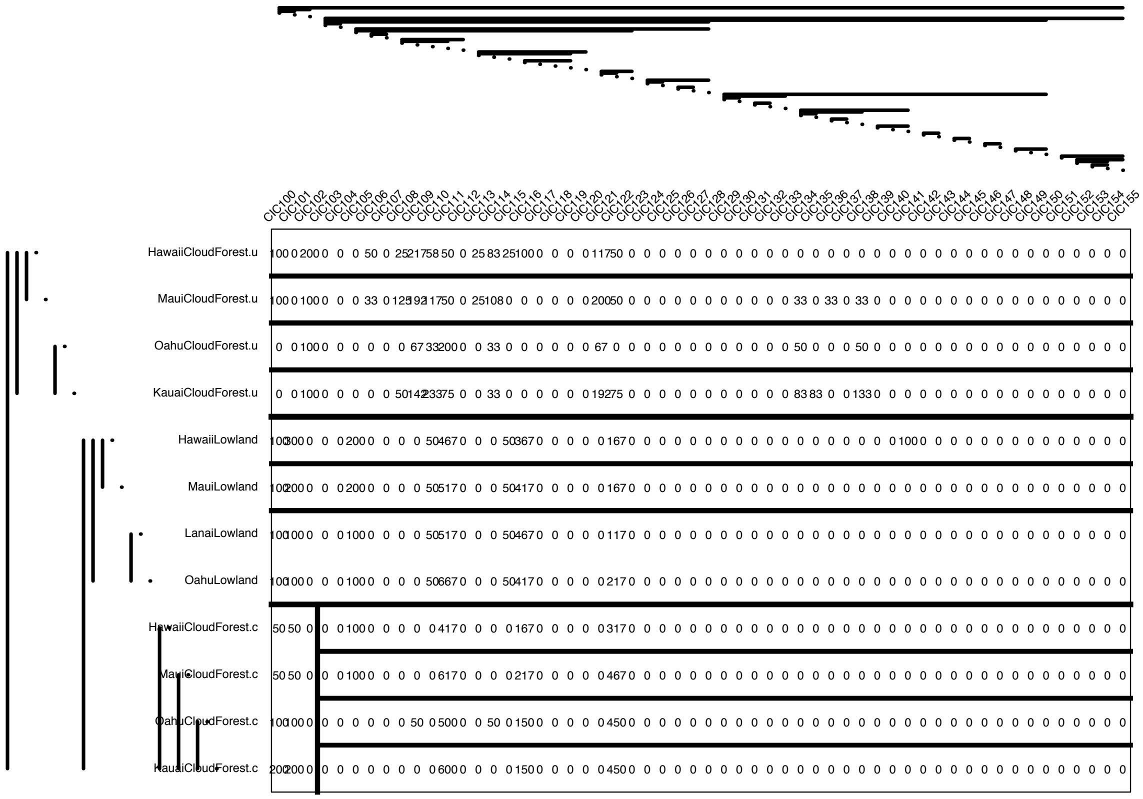


Figure 7.8: Mondrianesque partition of morphologically binned leaf counts from three forest types on four Hawaiian islands.

as having no significant structure at any reasonable significance level.

Also, the partitions made in the simulated data set are somewhat counter-intuitive. The split between the top 5 rows and bottom seven rows was made correctly, but then continued splitting failed to show any break in the bottom half (which was in fact made up of two distinct sections) but partitioned the top half more fully. Though this seems counter-intuitive, it is the the relative proportions of the counts that are being seen by the test statistic, so if there is a cluster of rows with real proportional similarity (as is the case when it is really made up from two different distributions, it is correct to refuse to partition it further.

Finally, there are three things that affect the degree of splitting: first is the actual total number of counts, second is the significance at which to compare *pll* with a chi-squared tail probability by the `null.check()` subroutine, and third is the type of correction used to penalize the test statistics when comparing row and column splits. The default coefficient of 1 is associated with an AIC-like penalty, but a larger penalty (*e.g.* 2 for BIC-like behavior) will affect the propensity for splitting long, thin blocks along their short axis or their long axis. The interactions between these three effects (overall number of counts, degree of splitting desired, and penalty for degrees of freedom) are not understood in detail.

7.5 Discussion

As is apparent from the variety of clustering methods found in the literature, there are a vast number of variant procedures that can be followed in partitioning a matrix. The choices that have been made here are so specific that the functions that have been described are only applicable to data of a very restricted type. In fact, this type is so restricted that the real data for which the procedures were designed may not meet the theoretical criteria for it to work properly. Further exploration of the algorithms applied to a variety of data sets is needed in order to constrain their exact regions of applicability.

Nevertheless, the good performance of the algorithms on the simulated data set indicate that if the assumptions are met exactly, correct results can be obtained. Naturally it remains to identify what modification to the algorithms are necessary for them to apply to the real data with the same degree of robustness, but there seems to be a reasonable chance that such tuning can be accomplished.

8 Conclusion

This conclusion has two complementary functions: one to summarize what has been shown in this dissertation, the other to point out what further work remains to be done, i.e. to highlight what has not been successfully demonstrated or accomplished here.

Chapters 1 introduced and contextualized leaf ecophenetics, the method developed in this dissertation. In chapter 2, I addressed the basic question asked in my thesis: whether leaf architectural data (as represented in a simple numerical coding scheme) contain extractible information about dicot forest ecosystems. Chapters 3 through 6 then provided examples of ways in which these data could be extracted and made to elucidate biological questions. Finally, chapter 7 proposed a new practical algorithm for classifying things like floras about which data is most naturally collected as structured factors.

More specifically, the basic leaf architectural data collected are organized in a sites-by-variables matrix where, for instance, the sites are represented by the rows and the variables by the columns. In chapter 2, I showed how two published descriptive systems (CICs and CLAMP), which provide different sets of variables, nevertheless produce similar classifications; and how the classifications produced are not obscured by the process of data collection and that they are reasonable in the light of prior knowledge about the ecosystems being classified. In chapter 3, I treated the rows as time periods and used a classification of the rows to argue that the effect of the Cretaceous/Tertiary boundary event on plant ecosystems in North America should be described as a mass death, not a mass extinction. This conclusion was based on the assumption that the paleobotanical literature (as represented by the *Compendium Index*) accurately reflects biological dynamics. I also pointed out the scale-dependence of extinctions, temporally, geographically, and taxonomically. In chapter 4, I performed a global meta-analysis of published CLAMP data and identified one major problem affecting it: inhomogeneous spatial sampling. In chapter 5, I illustrated how this problem can be addressed by using synthetic floras and identified a previously unknown non-linearity in the relationship between the percentage of entire leaves and mean annual temperature. Having dealt with temporal and geographical variation, in chapter 6, I turned to taxonomic relationships and showed how the leaf-architectural signal is not merely an artifact of phylogenetic relationships. Discussion of how to interpret a leaf architectural signal revealed problems with treating variables that are in fact structured as independent (an issue that had also arisen in chapter 4). Finally, in chapter 7, I proposed a new way of subdividing a matrix of counts based on such structured factors.

To describe the methods developed in chapter 2 with applications as shown in chapters 3–6, I have coined the term *ecophenetics,* in the hopes that putting a name on a method will make it easier for others to employ it. The method is characterized by a focus on ecology and environment (broadly construed), hence the prefix 'eco-'; and the quantitative or semi-quantitative description of the plant body or phenotype, hence 'phenetics'. No component of the method is new: quantitative architectural description of leaves, aggregating leaf descriptions by flora, the focus on classification and the particular graphical and algorithmic tools used for classification have all been used before. None of the components of the method are inalienable: it would be perfectly reasonable to use my new method for classification by structured factors, but apply it to flower and fruit morphology instead of leaf architecture; or to retain the focus on leaves but rely to a heavier degree on eigenvector methods of data summary. Although no individual component of the method is new, innovations include the application of such algorithmic or phenetic tools to old ecological questions about vegetation classification and the reliance on functionally important leaf architectural data.

A potential user of this new method might ask what else is needed to allow ecophenetics to enter textbooks as an established paleoecological methodology. Obviously it would benefit from continued development. I see two specific directions in which it could be improved: first, the variables that I rely on in this dissertation were not designed or chosen for the purpose to which I put them. CICs were intended for fossil identification; CLAMP variables for estimation of paleoclimates. Both are over-split in the sense that variation is spread out over too many variables for efficient description of leaves for ecological purposes, so reducing the number of variables would be a significant improvement.

An improved set of variables for ecological characterization of angiosperm leaves would be:

1. Size (e.g. area)
2. Aspect ratio (e.g. length/width)—measures perimeter-to-area-ratio
3. Moment—measures distance of photosynthetic area from its attachment to closest abscission
4. Margin dissection (e.g. perimeter to area ratio)—measures toothiness
5. Shade diameter (e.g. diameter of maximal included disk)—measures lobation
6. Organization—measures efficiency of vascularization and structural reinforcement
7. Drip tips
8. Compounding

These variables were not used in this dissertation because a large body of data is needed to characterize variation. The scope of a doctoral dissertation is not extensive enough to support inventing a new set of variables and then using the newly invented variables to compiling a large enough body of data to be interesting. Thus I was limited to the methods of description that had published data associated with them. This leads to the second improvement from which ecophenetics would benefit: data sharing.

As electronic data storage and automated methods for data collection become standard, paleontology, as well as other natural historical sciences, is undergoing a transition from having too few data to having too many. Reasonable scientific conclusions are no longer typically made from little information that must be extrapolated to cover missing cases. Instead the standard situation is to have too many data to determine what is meaningful. When only a few data points are available, a great deal of time is spent analyzing them. Any biases introduced by collection and analysis can at least be fully discussed, if not quantified. When vast amounts of easily-collected data are available, however, small biases in the collection process can easily swamp signals present in the data. Such biases or inconsistencies must be caught and fixed.

The solution I propose is analogous to the open-source software movement. Since the 1980s, a tremendous number of programmer-hours have been donated to the public on the grounds that in a non-zero-sum game (positive feedback machine) collaboration is a more advantageous strategy than competition, which in such a situation is fundamentally maladaptive. The production and analysis of eco-morphological data clearly depends on positive feedback: the more and better data others produce, the more valuable mine become. Therefore agreement among paleobotanists on common standards for data collection, combined with free and frequent data comparison is the second change that could dramatically increase the value of ecophenetics as a method. In systematic paleobotany, this fertilization effect has been noted and has led to the production of some more-or-less open and public databases like the Paleobiology Database (PBDB) and MorphoBank (Paleobiology Database; O'Leary and Kaufman 2007). While both of these attempts have serious problems that may prevent either one from implementing ecophenetic data-sharing, they provide examples on which such an ecophenetic data-exchange could be built.

Finally, ecophenetics will retain some inherent limitations even if these improvements can be made. No analysis, ecological or otherwise, can improve on poor data. The quality of the fossil record cannot be improved and biases due to taphonomy and collection procedures can only be explained and minimized, never eliminated. Moreover, forcing qualitatively different things into the same descriptive framework will always be difficult: a Carboniferous coal-swamp forest may share characteristics of a modern alpine meadow but discovering architectural or morphological measures to reflect those similarities accurately may not be possible. At some level, the only way to compare incomparables is with the most flexible of classificatory tools, language. Equifinality—when several different processes can lead to indistinguishable outcomes—provides a point beyond which quantitative analysis cannot penetrate. Another way of looking at this is that independent information is needed to deconvolute multiple signals: playing two different radio stations at the same time prevents either one from being heard properly. Leaf architecture, like many biological properties, may ultimately reflect many different causes interacting with too much complexity to be separated fully from knowledge of their effects alone.

In concert with existing complementary methods of data collection and analysis, ecophenetic examination of the plant fossil record has the potential to answer historical questions about what the planet's landscape looked like in the distant past, and to make predictions about what it is likely to look like in the remote future. Moreover it can identify regularities in the responses of plant ecosystems to their biotic and abiotic environments. Thus even when predictions about the future turn out to have been wrong, negative results will still improve our methods for making such predictions. Awaiting these potential improvements and noting inherent limitations, I believe ecophenetics can provide a useful addition to the toolbox of a plant paleoecologist.

References

J. M. Adams and H. Faure. Palaeovegetation maps of the earth during the last glacial maximum, and the early and mid Holocene as an aid to archaeological research. *Journal of Archaeological Science*, 24:623–647, 1997.

J. M. Adams, W. A. Green, and Y. Zhang. Leaf margins and temperature in the North american flora: recalibrating the paleoclimatic thermometer. *Global and Planetary Change*, in prep.

H. K. Airy Shaw. *Willis's Dictionary of the Flowering Plants and Ferns.* Cambridge University Press, Cambridge, 8 edition, 1973.

H. H. Allen. *Flora of New Zealand.* R. E. Owen, Government Printer, Wellington, New Zealand, 1961.

J. R. L. Allen. A review of the origin and characterisics of recent alluvial sediments. *Sedimentology*, 5:89–191, 1965.

W. D. Allmon and D. J. Bottjer, editors. *Evolutionary Paleoecology.* Columbia University Press, New York, 2001.

L. W. Alvarez, W. Alvarez, F. Asaro, and H. V. Michel. Extraterrestrial cause for the cretaceous-tertiary extinction: experimental results and theoretical interpretation. *Science*, 208:1095–1108, 1980.

H. N. Andrews. Index of generic names of fossil plants, 1820–1965. *United States Geological Society Bulletin*, 1300:1–354, 1970.

2003 Angiosperm Phylogeny Group. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society*, 141:399–436, 2003.

T. Armstrong. A Paleoclimatic Interpretation of Cenomanian Inland Flora from Schefferville, Quebec, Canada, 1993. Yale College Senior Thesis.

A Ash, B. Ellis, L. J. Hickey, K. Johnson, P. Wilf, and S. Wing. *Manual of Leaf Architecture.* Smithsonian Institution, Washinton, DC, 1999.

D. I. Axelrod. *The Pliocene Verdi Flora of Western Nevada.* University of California Press, Berkeley, 1958.

D. I. Axelrod. *The Pliocene Sequoiadendron Forest from Western Nevada.* University of California Press, Berkeley, 1962.

D. I. Axelrod. *The Pliocene Soboba Flora of Southern California.* University of California Press, Berkeley, 1966.

P. Baas. Ecological patterns in xylem anatomy. In Givnish (1986), pages 327–352.

I. W. Bailey and E. W. Sinnott. A botanical index of cretaceous and tertiary climates. *Science*, 41 (1066):831–834, 1915.

I. W. Bailey and E. W. Sinnott. The climatic distribution of certain types of angiosperm leaves. *American Journal of Botany*, 3(1):24–39, 1916.

K. F. Baker-Brosh and R. K. Peet. The ecological significance of lobed and toothed leaves in temperate forest trees. *Ecology*, 78:1250–1255, 1997.

K. E. Basford and J. W. Tukey. *Graphical Analysis of Multiresponse Data, Illustrated with a Plant Breeding Trial.* Chapman & Hall/CRC, Boca Raton, Fla., 1999.

N. C. W. Beadle. *The Vegetation of Australia.* Cambridge University Press, Cambridge, 1981.

H. F. Becker. *Oligocene Plants from the Upper Ruby River Basin, Southwestern Montana.* Number 82 in Memoirs. Geological Society of America, New York, 1961.

A. K. Behrenmeyer and R. W. Hook. Paleoenvironmental contexts and taphonomic modes. In Behrensmeyer et al. (1992), pages 15–136.

A. K. Behrensmeyer, J. D. Damuth, W. A. DiMichele, R. Potts, H. Sues, and S. L. Wing, editors. *Terrestrial Ecosystems through Time.* University of Chicago Press, Chicago, 1992.

R. A. Berner and Z. Kothavala. GEOCARB III, a revised model of atmospheric CO_2 over Phaneraozoic time. *American Journal of Science*, 301:182–204, 2001.

E. W. Berry. *The Lower Eocene Floras of Southeastern North America*, volume 91 of *United States Geological Survey Professional Papers.* Government Printing Office, Washington, 1916.

É. Boreau, editor. *Traitè de Palæobotanique.* Masson, Paris, 1964–.

E. L. Braun. *Deciduous Forests of Eastern North America.* Hafner Publishing Company, New York, 1950.

J. Braun-Blanquet. *Plant Sociology: the Study of Plant Communities*. Hafner Publishing Company, New York, 1965.

J. S. Bridge. Description and interpretation of fluvial deposits: A critical perspective. *Sedimentology*, 40:801–810, 1993.

N. Britton and A. Brown. *An Illustrated Flora of the Northern United States and Canada*. Scribner's, New York, 2nd edition, 1913.

N. L. Britton and P. Wilson. *Botany of Puerto Rico and the Virgin Islands*. New York Academy of Sciences, New York, 1923–1930.

R. W. Brown. Fossil plants from the Aspen Shale of southwestern Wyoming. *Proceedings of the United States National Museum*, 82(12):1–10, 1933.

V. K. Brown and J. H. Lawton. Herbivory and the evolution of leaf size and shape. *Philosophical Transactions of the Royal Society of London B*, 333:256–272, 1991.

R. J. Burnham, N. C. A. Pitman, K. R. Johnson, and P. Wilf. Habitat-related error in estimating temperatures from leaf margins in a humid tropical forest. *American Journal of Botany*, 88(6): 1096–1102, 2001.

M. Canny. What becomes of the transpiration stream? *New Phytologist*, 114:341–368, 1990.

D. J. Cantrill and I. Poole. Floristic change in the cretaceous to cenozoic vegetation of antarctica, 2005. Vienna, XVII International Botanical Congress Abstracts.

J. M. Chambers, W. S. Cleveland, B. Kleiner, and P. A. Tukey. *Graphical Methods for Data Analysis*. Wadsworth International Group, Belmont, California, 1983.

H. G. Champion and S. K. Seth. *A Revised Survey of the Forest Types of India*. Manager of Publications, Government of India Press, Delhi, 1968.

H. G. Champion, S. K. Seth, and G. M. Khattak. *Forest Types of Pakistan*. Pakistan Forest Institute, Peshawar, 1965.

M. E. J. Chandler. *The Lower Tertiary Floras of Southeast England*, volume 1–4. British Museum, London, 1961.

R. W. Chaney, M. K. Elias, E. Dorf, D. I. Axelrod, and C. Condit. *Miocene and Pliocene Floras of Western North America*. Number 476 in Carnegie Institute Publications. Carnegie Institute of Washington, Washington, DC, 1938.

R. W. Chaney, C. Condit, and D. I. Axelrod. *Pliocene Floras of California and Oregon.* Carnegie Institute of Washington, Washington, DC, 1944.

D.C. Christophel and B.P.M. Hyland. *Leaf Atlas of Australian Tropical Rain Forest Trees.* CSIRO, East Melbourne, 1993.

W. A. Clemens, J. D. Archibald, and L. J. Hickey. Out with a whimper not a bang. *Paleobiology*, 7 (3):293–298, 1981.

William S. Cleveland. *The Elements of Graphing Data.* Wadsworth Advanced Books and Software, Monterey, California, 1985.

M. E. Collinson, M. C. Boulter, and P. L. Holmes. Magnoliophyta. In M. J. Benton, editor, *The Fossil Record 2*, pages 809–841, London, 1993. Chapman and Hall.

W.B. Critchfield and Jr. Little, E.L. Geographic distribution of the pines of the world. *U.S. Department of Agriculture Miscellaneous Publications*, 991:1–97, 1966.

A. Cronquist. *An integrated system of classification of flowering plants.* Columbia University Press, New York, 1981.

E. J. Crosby. Classification of sedimentary environments. In J. K. Rigby and W. K. Hamblin, editors, *Recognition of Ancient Sedimentary Environments*, volume 16 of *Society of Economic Paleontologists and Mineralogists Special Publications*, pages 4–11, 1972.

P. Dansereau and P. F. Buell. *Studies on the Vegetation of Puerto Rico.* University of Puerto Rico, Mayagüez, Puerto Rico, 1966.

A. de Candolle. *Géographie Botanique Raisonée; ou Expostion des Faits Principaux et des Lois Concernant la Distribution Géographique des Plantes de l'Epoque Actualle.* V. Masson, Paris, 1855.

K. de Queirioz and J. A. Gautier. Phylogenetic taxonomy. *Annual Review of Ecology and Systematics*, 23:449–480, 1992.

W. A. DiMichele and S. L. Wing. Methods and applications of plant paleoecology. *Paleontological Society Special Publications*, 3, 1988.

W. A. DiMichele, W. E. Stein, and R. M. Bateman. Ecological sorting of vascular plant classes during the paleozoic evolutionary radiation. In W. D. Allmon and D. J. Bottjer, editors, *Evolutionary Paleoecology*, pages 285–335, New York, 2001. Columbia University Press.

J. R. Dodd and R. J. Stanton. *Paleoecology: Concepts and Applications.* Wiley, New York, 1990.

G. E. Dolph. Interrelationships among the gross morphological features of angiosperm leaves. *Bulletin of the Torrey Botanical Society*, 103:29–34, 1976.

G. E. Dolph. Variation in leaf size and margin type with respect to climate. *Cour. Forsch.-Inst.Senckenberg*, 30:153–158, 1978.

G. E. Dolph. Variation in leaf margin with respect to climate in costa rica. *Bulletin of the Torrey Botanical Society*, 106:104–109, 1979.

G. E. Dolph. Leaf form of the woody plants of Indiana as related to environment. In N. S. Margaris, M. Arianoustou-Farragitaki, and W. C. Oechel, editors, *Being Alive on Land: Tasks for Vegetation Science, vol. 13*, pages 51–61, Dr. W. Junk Publishers, 1984. The Hague.

G. E. Dolph and D. L. Dilcher. Variation in leaf size with respect to climate in costa rica. *Biotropica*, 12(2):91–99, 1980a.

G. E. Dolph and D. L. Dilcher. Variation in leaf size with respect to climate in the tropics of the western hemisphere. *Bulletin of the Torrey Botanical Society*, 107(2):104–109, 1980b.

E. Dorf. An illustrated catalogue of mesozoic and early cenozoic plants of north america. *Science*, 91(2368):1f., 1940.

L. S. Espinal and E. Montenegro. *Formaciones Vegetales de Columbia.* Republic of Columbia Department of Agriculture, 1963.

B. Everitt. *Cluster Analysis.* Wiley, New York, 1974.

C. E. Forest, J. A. Wolfe, P. Molnar, and K. A. Emanuel. Paleoaltimetry incorporating atmospheric physics and botanical estimates of paleoclimate. *Geological Society of America Bulletin*, 111(4): 497–511, 1999.

J. C. Gall. *Ancient Sedimentary Environments and the Habitats of Living Organisms.* Springer-Verlag, Berlin, 1983.

A. H. Gentry. *A Field Guide to the Families and Genera of Woody Plants of Northwest South America.* University of Chicago Press, Chicago, 1993.

T. J. Givnish. Ecologial aspects of plant morphology: leaf form in relation to environment. *Acta Biotheoretica*, 27:83–142, 1978.

T. J. Givnish, editor. *On the Economy of Plant Form and Function.* Cambridge University Press, Cambridge, 1986.

A. Gleason and A. Cronquist. *Manual of Vascular Plants of Northeastern United States and Adjacent Canada.* D. Van Nostrand Company, New York, 1963.

A. D. Gordon. *Classification.* Chapman and Hall / CRC, 2nd edition, 1999.

S. J. Gould. The disparity of the burgess shale arthropod fauna and the limits of cladistic analysis: why we must strive to quantify morphospace. *Paleobiology*, 17(4):411–423, 1991.

W. A. Green. Loosening the clamp: An exploratory graphical approach to the climate leaf analysis multivariate program. *Palaeontologia Electronica*, 9(2):9A, 2006. URL `http://www.palaeo-electronica.org/2006_2/clamp/clamp.pdf`.

W. A. Green and L. J. Hickey. Using leaf architecture to compare fossil and modern forests: preliminary results from hierarchical cluster analysis and graphical representation [abstract]. *Ecological Society of America 88th Annual Meeting Abstracts*, page 129, 2003.

W. A. Green and L. J. Hickey. If a forest falls, do the trees hear it? eco-morphological dynamics of the cretaceous-tertiary boundary event. *Botanical Society of America, Annual Meeting Abstracts*, page 84, 2004.

W. A. Green and L. J. Hickey. Leaf architectural profiles of angiosperm floras across the Cretaceous/Tertiary boundary. *American Journal of Science*, 305(10):983–1013, 2005.

D. R. Greenwood, P. Wilf, S. L. Wing, and D. C. Christophel. Paleotemperature estimation using leaf-margin analysis: Is australia different? *Palaios*, 19:129–142, 2004.

K. M. Gregory-Wodzicki. Relationships between leaf morphology and climate, bolivia: implications for estimating paleoclimate from fossil floras. *Paleobiology*, 26(4):668–688, 2000.

W. Grimm. *The Book of Shrubs.* The Stackpole Company, Harrisburgh, PA, 1957.

R. H. Groves, editor. *Australian Vegetation.* Cambridge University Press, Cambridge, 2 edition, 1994.

J. A. Hartigan. *Clustering Algorithms.* Wiley, New York, 1975.

A.B. Herman and R. A. Spicer. Palaeobotanical evidence for a warm cretaceous arctic ocean. *Nature*, 380:330–333, 1996.

V. H. Heywood and J. McNeill, editors. *Phenetic and Phylogenetic Classification.* Systematics Association Pub. No. 6. The Systematics Association, London, 1964.

L. J. Hickey. *Stratigraphy and Paleobotany of the Golden Valley Formation (Early Tertiary) of Western North Dakota.* Number 150 in Geological Society of America Memoirs. Geological Society of America, Boulder, CO, 1977.

L. J. Hickey. *Land plant change across the Cretaceous-Tertiary boundary*, volume 499 of *Lunar and PIanetary Institute Contributions*, pages 279–313. Lunar and PIanetary Institute, 1981.

L. J. Hickey. The Meeteetsee Flora. unpublished field notes, 1987–1992.

L. J. Hickey. The Englishtown Flora. unpublished field notes, 1999.

L. J. Hickey, L. S. Klise, and W. A. Green. The Yale-Princeton Compendium Index of North American Mesozoic and Cenozoic Type Fossil Plants, electronic version 1.0, 2006. URL `http://www.yale.edu/peabody/collections/pb/eCI/v1.0`.

J. C. Hickman, editor. *The Jepson Manual: Higher Plants of California.* University of California Press, Berkeley, 1993. URL `http://ucjeps.berkeley.edu/interchange/I_treat_indexes.html`.

J. D. Hooker. *Flora of British India*, volume 1–7. L. Reeve and Co., London, 1872–1897.

P. M. Huff, P. Wilf, and E. J. Azumah. Digital future for paleoclimate estimation from fossil leaves? preliminary results. *Palaios*, 18:266–274, 2003.

J. Hutchinson and J. M. Dalziel. *Flora of West Tropical Africa*, volume 1–3. Crown Agents for Oversea Governments and Administrations, London, 1954–1968.

B. F. Jacobs. Estimation of rainfall variables from leaf characters in tropical africa. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 145:231–250, 1999.

B. F. Jacobs. Estimation of low-latitude paleoclimates using fossil angiosperm leaves: examples from the miocene tugen hills, kenya. *Paleobiology*, 28(3):399–421, 2002.

B. F. Jacobs and A. L. Deino. Test of climate-leaf physiognomy regression models, their application to two Miocene floras from Kenya, and ^{40}Ar/^{39}Ar dating of the Late Miocene Kapturo site. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 123:259–271, 1996.

K. R. Johnson. The Puercan Flora and paleoenvironments of the Hunt Creek Sandstone Member of the Fort Union Formation, northern Bighorn Basin, Park County, Wyoming, and Carbon County, Montana. Master's thesis, University of Pennsylvania, 1985.

K. R. Johnson. *A high-resolution megafloral biostratigraphy spanning the Cretaceous-Tertiary boundary in the northern Great Plains.* PhD thesis, Yale University, 1989.

K. R. Johnson and L. J. Hickey. Megafloral change across the Cretaceous/Tertiary boundary in the northern Great Plains and Rocky Mountains, U.S.A. *Geological Society of America Special Papers*, 247:433–444, 1990.

G. J. Jordan. Uncertainty in palaeoclimatic reconstructions based on leaf physiognomy. *Australian Journal of Botany*, 45:527–547, 1997.

L. Kaufman and P. J. Rousseeuw. *Finding Groups in Data: An Introduction to Cluster Analysis.* Wiley, New York, 1990.

E. M. Kennedy. *Cretaceous and Tertiary Megafloras from New Zealand and their Climate Signals.* PhD thesis, Open University, 1998.

E. M. Kennedy, R. A. Spicer, and P. M. Rees. Quantitative palaeoclimate estimates from late cretaceous and paleocene leaf floras in the northwest of the south island, new zealand. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 184:321–345, 2002.

A. H. Knoll, K. Niklas, and B. H. Tiffney. Phanerozoic land plant diversity in north america. *Science*, 206:1400–1402, 1979.

F. H. Knowlton. Catalogue of the mesozoic and cenozoic plants of north america. *United States Geological Survey Bulletin*, 696:1–815, 1919.

W.-S. Kong and D. Watts. *The Plant Geography of Korea with an Emphasis of the Alpine Zones.* Kluwer Academic Publishers, Boston, 1993.

E. J. Kormondy. *Concepts of Ecology.* Prentice Hall, Englewood Cliffs, New Jersey, 1969.

W. L. Kovach and R. A. Spicer. Canonical correspondence analysis of leaf physiognomy: a contribution to the development of a new palaeoclimatological tool. *Palaeoclimates*, 1:125–138, 1995.

E. A. Kowalski. Mean annual temperature estimation base on leaf morphology: a test from tropical South America. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 188:141–165, 2002.

E. A. Kowalski and D. L. Dilcher. Warmer paleotemperatures for terrestrial ecosystems. *Proceedings of the National Academy of Sciences*, 100(1):167–170, 2003.

V. A. Krassilov. *Paleoecology of Terrestrial Plants: Basic Principals and Techniques.* Wiley, New York, 1975.

M. J. Krauss and T. M. Bown. Pedofacies analysis: A new approach to reconstructing ancient fluvial sequences. In *Paleosols and Weathering through Geologic Time: Principles and Applications*, volume 216 of *Geological Society of America Special Papers*, pages 143–152, 1988.

G. Krüssmann. *Manual of Cultivated Broad-leaved Trees and Shrubs*, volume 1–2. Timber Press, Portland, Oregon, 1985.

C. A. La Pasha and C. N. Miller. Flora of the Early Cretaceaous Kootenai Formation in Montana, paleoecology. *Palaeontographica Abteilung B*, 194:109–130, 1984.

C.C. Labandeira, K. R. Johnson, and P. Wilf. Impact of the terminal cretaceous event on plant-insect associations. *Proceedings of the National Academy of Sciences, USA*, 99:2061–2066, 2002.

H. S. Ladd. Treatise on marine ecology and paleoecology, vol. 2: Paleoecology. *Memoirs of the Geological Society of America*, 67, 1957.

R. N. Lakhanpal. *The Rujada Flora of West Central Oregon.* University of California Press, Berkeley, 1958.

R. S. Lamotte. Catalogue of the cenozoic plants of north america through 1950. *Memoirs of the Geological Society of America*, 51:1–381, 1952.

S.-C. Lee. *Forest Botany of China.* The Commercial Press, Shanghai, 1935.

L. Legendre and P. Legendre. *Numerical Ecology.* Elsevier, Amsterdam, 1983.

P. Legendre and E. D. Gallagher. Ecologically meaningful transformations for ordination of species data. *Oecologia*, 129:271–280, 2001.

H. Leon and H. Alain. *Flora de Cuba*, volume 1–5. P. Fernandez and Co., Havana, 1946–1962.

J. Lepš and P. Šmilauer, editors. *Multivariate Analysis of Ecological Data using CANOCO.* Cambridge University Press, Cambridge, UK, 2003.

S. Lidgard and P. R. Crane. Quantitative analysis of early angiosperm radiation. *Nature*, 331: 344–346, 1988.

E. M. Lind and M. E. S. Morrison. *East African Vegetation.* Langman Group, Ltd., London, 1927.

A. H. Liogier. *La Flora de la Española.* Universidad Central del Este, Santo Domingo, 1995.

E.L. Little Jr. Atlas of united states trees, volume 1, conifers and important hardwoods. *U.S. Department of Agriculture Miscellaneous Publication*, 1146, 1971.

E.L. Little Jr. Atlas of united states trees, volume 3, minor western hardwoods. *U.S. Department of Agriculture Miscellaneous Publication*, 1314, 1976.

E.L. Little Jr. Atlas of united states trees, volume 4, minor eastern hardwoods. *U.S. Department of Agriculture Miscellaneous Publication*, 1342, 1977.

E.L. Little Jr. Atlas of united states trees, volume 5, florida. *U.S. Department of Agriculture Miscellaneous Publication*, 1361, 1978.

E.L. Little Jr., R. O. Woodbury, and F. H. Wadsworth. Trees of puerto rico and the virgin islands, vol. 1. *U.S. Department of Agriculture Handbooks*, 249, 1964.

E.L. Little Jr., R. O. Woodbury, and F. H. Wadsworth. Trees of puerto rico and the virgin islands, vol. 2. *U.S. Department of Agriculture Handbooks*, 449, 1974.

D. J. Mabberley. *The Plant Book*. Cambridge University Press, Cambridge, 2nd edition, 1997.

H. D. MacGinitie. *Fossil Plants of the Florissant Beds, Colorado*, volume 599. Carnegie Institution of Washington, Washington, DC, 1953.

H. D. MacGinitie. *The Kilgore Flora*. University of California Press, Berkeley, 1962.

H. D. MacGinitie. *The Eocene Green River Flora of Northwestern Colorado and Northeastern Utah*. University of California Press, Berkeley, 1969.

H. D. MacGinitie. *An Early Middle Eocene Flora from the Yellowstone-Absaroka Volcanic Province, Northwestern Wind River Basin, Wyoming*. University of California Press, Berkeley, 1974.

P. F. Maycock. The ecology of beech (Fagus grandifolia Ehrh.) forests of the deciduous forests of southeastern North America, and a comparison with the beech (Fagus crenata) forests of Japan. In A. Miyawaki, K. Iwatsuki, and M. M. Grandtner, editors, *Vegetation in Eastern North America*, pages 351–407, Tokyo, 1994. University of Tokyo Press.

R. P. McIntosh, editor. *Phytosociology*. Number 6 in Benchmark Papers in Ecology. Dowden, Hutchinson and Ross, Inc., Stroudsburg, PA, 1978.

R. P. McIntosh, editor. *The Background of Ecology*. Cambridge Studies in Ecology. Cambridge University Press, Cambridge, 1985.

C. McMullan. *Flowering Plants of the Galápagos*. Cornell University Press, Ithaca, New York, 1999.

A. D. Miall. Description and interpretation of fluvial deposits: A critical perspective. *Sedimentology*, 42:379–384, 1995.

A. D. Miall. *The Geology of Fluvial Deposits.* Springer-Verlag, Berlin, 1996.

I. M. Miller, M. Brandon, and L. J. Hickey. Using leaf margin analysis to estimate the mid-Cretaceous (Albian) paleolatitude of the Baja BC block. *Earth and Planetary Science Letters*, 245(1–2):95–114, 2006.

R. C. Moore et al., editors. *Treatise on Invertebrate Paleontology.* Geological Society of America, New York, 1952–.

V. Mosbrugger and T. Utescher. The coexistence approach—a method for quantitative reconstructions of Tertiary terrestrial palaeoclimate data using plant fossils. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 134:61–86, 1997.

V. Mosbrugger, C. Traiser, and S. Klotz. 'synthetic floras'—a flexible tool in large scale vegetation analysis [abstract], 2005. XVII International Botanical Congress Abstracts.

D. Mueller-Dombois and F. R. Fosberg. *Vegetation of the Tropical Pacific Islands.* Number 132 in Ecological Studies. Springer, New York, 1998.

K. J. Niklas, B. H. Tiffney, and A. H. Knoll. Patterns in vascular land plant diversification: An analysis at the species level. In J. Valentine, editor, *Phanerozoic Diversity Patterns*, pages 97–127, Princeton, New Jersey, 1985. Princeton University Press.

Flora of North America Editorial Committee, editor. *Flora of North America North of Mexico, 7+ vols.*, volume 7+. New York, 1993–. URL `http://www.efloras.org/flora_page.aspx?flora_id=1`.

J. Oksanen. Ordination of boreal heath-like vegetation with principal component analysis, correspondence analysis and multidimensional scaling. *Vegetatio*, 52:181–189, 1983.

M. A. O'Leary and S. G. Kaufman. Morphobank2.5: Web application for morphological phylogenetics and taxonomy, 2007. URL `http://www.morphobank.org/`.

Paleobiology Database, 2007. URL `http://paleodb.org`.

A. R. Palmer and J. Geissman. 1999 geological time scale, 1999. URL `http://www.geosociety.org/science/timescale/timescl.pdf`.

S. E. Peters and M. Foote. Biodiversity in the phanerozoic: a reinterpretation. *Paleobiology*, 27(4): 583–601, 2001.

R. Preston. *North American Trees.* Iowa State University Press, Ames, IA, 1961.

G. R. Proctor. *Flora of the Cayman Islands.* Royal Botanic Gardens, Kew, London, 1984.

R Development Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria, 2005. URL `http://www.R-project.org`. ISBN 3-900051-07-0.

C. Raunkiaer. *The Life Forms of Plants and Statistical Plant Geography.* Clarendon Press, Oxford, 1934.

D. M. Raup. Taxonomic diversity during the phanerozoic. *Science*, 177(4054):1065–1071, 1972.

P. H. Raven, R. F. Evert, and S. E. Eichhorn. *Biology of Plants.* Worth Publishers, New York, 6th edition, 1999.

M. Rees, R. Condit, M Crawley, S. Pacala, and D. Tilman. Long-term studies of vegetation dynamics. *Science*, 293:650–655, 2001.

E. M. Reid and M. E. J. Chandler. *The Bembridge Flora. Catalogue of Cainozoic Plants in the Department of Geology.* The British Museum, London, 1926.

E. M. Reid and M. E. J. Chandler. *The Flora of the London Clay.* The British Museum, London, 1933.

G. J. Retallack. A paleopedological approach to the interpretation of terrestrial sedimentary rocks: The mid-Tertiary fossil soils of Badlands National Park, South Dakota. *Geological Society of America Bulletin*, 94:823–840, 1983.

G. J. Retallack and J. Germán-Heins. Evidence from paleosols for the geological antiquity of rain forest. *Science*, 265:499–502, 1994.

P.W. Richards. Ecological studies on the rain forest of southern nigeria: I. the structure and floristic composition of the primary forest. *Journal of Ecology*, 27(1):1–61, 1939.

J. S. Rodwell, editor. *British Plant Communities, Volume 1: Woodlands and Scrub.* Cambridge University Press, Cambridge, 1991.

D. L. Royer. Depth to pedogenic carbonate horizon as a paleoprecipitation indicator? *Geology*, 27: 1123–1126, 1999.

D. L. Royer and P. Wilf. Why do toothed leaves correlated with cold climates? Gas exchange at leaf margins provides new insights into a classic paleotemperature proxy. *International Journal of Plant Sciences*, 167(1):11–18, 2006.

D. L. Royer, P. Wilf, D. A. Janesko, E. A. Kowalski, and D.C.Dilcher. Correlations of climate and plant ecology to leaf size and shape: potential proxies for the fossil record. *American Journal of Botany*, 92(7):1141–1151, 2005.

C. S. Sargent. *Manual of the Trees of North America.* Houghton, Mifflin and Co., Boston, 1905.

P. Sarlin. *Bois et Forêts de la Nouvelle-Calédonie.* Centre Technique Forestier Tropical, Nogent-sur-Marne, France, 1954.

S. A. Schumm. Fluvial paleochannels. *Society of Economic Paleontologists and Mineralogists Special Publications*, 16:98–107, 1972.

J. J. Sepkoski. A compendium of fossil marine animal genera. *Bulletins of American Paleontology*, 363:1–560, 2002.

J. J. Sepkoski, R. K. Bambach, D. M. Raup, and J. Valentine. Phanerozoic marine diversity and the fossil record. *Nature*, 293(5832):435–437, 1981.

D. W. Shimwell. *The Description and Classification of Vegetation.* University of Washington Press, Seattle, 1971.

A. Smith. *Flora Vitensis Nova.* Pacific Tropical Botanical Garden, Lauai, Hawaii, 1979–1981.

E. E. Smith. *The Forests of Cuba.* Number 2 in Maria Moors Cabot Foundation Publications. Maria Moors Cabot Foundation, Cambridge, MA, 1954.

R. A. Spicer, A. B. Herman, and E. M. Kennedy. Foliar physiognomic record of climatic conditions during dormancy: Climate leaf analysis multivariate program (clamp) and the cold month mean temperature. *Journal of Geology*, 112:685–702, 2004.

R. A. Spicer, A. B. Herman, and E. M. Kennedy. The sensitivity of clamp to taphonomic loss of foliar physiognomic characters. *Palaios*, 20(5):429–438, 2005.

R.A. Spicer. Leaf physiognomy and climate change. In S.J. Culver and P. Rawson, editors, *Biotic Response to Global change: the Last 145 Million Years*, 1999.

C. Stace. *New Flora of the British Isles.* Cambridge University Press, Cambridge, 2nd edition, 1997.

P. F. Stevens. Angiosperm Phylogeny Website. version 7, May 2006, 2001–. URL `http://www.mobot.org/MOBOT/research/APweb/`.

A. C. Steward. *Catalogue of the Mesozoic Plants in the British Mueseum: The Wealden Flora*, volume 1–2. The British Museum, London, 1894.

A. L. Stoffers. *The Vegetation of the Netherlands Antilles*, volume 1 of *Studies on the Flora of Curacao and other Caribbean Islands*. Natuurwetenschappelijke Studiekring voor Suriname en de Nederlandse Antillen, Utrecht, 1956.

A. L. Stoffers, editor. *The Flora of the Netherlands Antilles*, volume 1–3. Natuurwetenschappelijke Studiekring voor Suriname en de Nederlandse Antillen, Utrecht, 1962–1984.

M. C. Stopes. *Catalogue of the Mesozoic Plants in the British Mueseum: The Cretaceous Flora*, volume 1–2. The British Museum, London, 1913.

L. Stranks. *Physiognomic and taphonomic studies in New Zealand and Australia: implications for the use of palaeobotany as a tool for palaeoclimate estimation*. PhD thesis, University of Oxford, 1996.

A. L. Takhtajan. *Floristic Regions of the World*. University of California Press, Berkeley, 1986.

A. L. Takhtajan. *Diversity and Classification of Flowering Plants*. Columbia University Press, New York, 1997.

V. Tamm. A biostratigraphical study of changes in lithology and paleofloral assemblages across the Lebo/Tongue River Member contact, in the Fort Union Formation of southeastern Montana, 1989. Amherst College Senior Thesis.

D. W. Taylor and L. J. Hickey. Phylogenetic evidence for the herbaceous origin of angiosperms. *Plant Systematic and Evolution*, 180:137–156, 1992.

D. W. Taylor and L. J. Hickey. Evidence for and implications of an herbaceous origin for angiosperms. In D. W. Taylor and L. J. Hickey, editors, *Flowering Plant Origin, Evolution and Phylogeny*, pages 232–266, New York, 1996. Chapman and Hall.

C.J.F. Ter Braak. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology*, 67:1167–1179, 1986.

R. S. Thompson, K. H. Anderson, and P. J. Bartlein. Atlas of relations between climatic parameters and distributions of important trees and shrubs in North America. *United States Geological Survey Professional Papers*, 1650a,b, 1999.

I. W. B. Thornton. *Krakatau; the destruction and reassembly of an island ecosystem*. Harvard University Press, Cambridge, Massachusetts, 1996.

D. Tilman. The ecological consequences of changes in biodiversity: A search for general principles. *Ecology*, 80(5):1455–1474, 1999.

C. Traiser. *Blattphysiognomie als Indikator für Unweltparameter: Eine Analyse rezenter und fossiler Floren.* PhD thesis, Eberhard-Karls Universität Tübingen, 2004.

C. Traiser, S. Klotz, D. Uhl, and V. Mosbrugger. Environmental signals from leaves—a physiognomic analysis of european vegetation. *New Phytologist*, 166:465–484, 2005.

E. Tufte. *The Visual Display of Quantitative Information.* The Graphics Press, Cheshire, Conn., 1983.

J. W. Tukey. *Exploratory Data Analysis.* Addison Wesley, Reading, Mass., 1977.

T. G. Tutin, V. H. Heywood, N. A. Burges, D. H. Valentine, S. M. Walters, and D. A. Webb, editors. *Flora Europaea.* Cambridge University Press, Cambridge, 1964–1983.

A. von Humboldt. *Essai sur la Géographie des Plantes.* F. Schoell, Paris, 1807.

W. L. Wagner, D. R. Herbst, and S. H. Sohmer. *Manual of the Flowering Plants of Hawai'i*, volume 1–2. University of Hawai'i Press, Honolulu, revised edition, 1999.

C.-W. Wang. *The Forests of China.* Number 5 in Maria Moors Cabot Foundation Publications. Maria Moors Cabot Foundation, Cambridge, MA, 1961.

L. F. Ward. *Sketch of Palaeobotany*, pages 357–452. United States Geological Survey, New York, 1883–1884.

P. Wardle. *Vegetation of New Zealand.* Cambridge University Press, Cambridge, 1991.

E. A. Wheeler and P. Baas. The potentials and limitations of dicotyledonous wood anatomy for climatic reconstructions. *Paleobiology*, 19:487–498, 1993.

R. H. Whittaker. Classification of natural communities. *Botanical Review*, 28:1–239, 1962.

R. H. Whittaker, editor. *Ordination and Classification of Communities.* Number 5 in Handbook of Vegetation Science. Dr. W. Junk Publishers, The Hague, 1973.

M. C Wiemann, S. R. Manchester, D. L. Dilcher, L. F. Hinojosa, and E. A. Wheeler. Estimation of temperature and precipitation from morphological characters of dicotyledonous leaves. *American Journal of Botany*, 85(12):1796–1802, 1998a.

M. C Wiemann, S. R. Manchester, and E. A. Wheeler. Paleotemperature estimation form dicotyledonous wood anatomical characters. *Palaios*, 14:459–474, 1998b.

M. C Wiemann, D. L. Dilcher, and S. R. Manchester. Estimation of mean annual temperature from leaf and wood physiognomy. *Forest Science*, 47(2):141–149, 2001.

P. Wilf. When are leaves good thermometers? a new case for leaf margin analysis. *Paleobiology*, 23 (3):373–390, 1997.

P. Wilf and K. R. Johnson. Land plant extinction at the end of the cretaceous; a quantitative analysis of the north dakota megafloral record. *Paleobiology*, 30(3):347–368, 2004.

P. Wilf, S. L. Wing, D. R. Greenwood, and C. L. Greenwood. Using fossil leaves as paleoprecipitation indicators: An eocene example. *Geology*, 26(3):203–206, 1998.

B. L. Williams. Megafloral relations within a Paleocene river system, southeastern Montana, 1988. Amherst College Senior Thesis.

G. B. Williamson. Driptips and splash erosion. *Biotropica*, 13:228–231, 1981.

G. B. Williamson, A. Romero, J. K. Armstrong, T. J. Gush, A. J. Hrushka, P. E. Klass, and J. T. Thompson. Driptips, drop size and leaf drying. *Biotropica*, 15:232–234, 1983.

S. L. Wing and D. R. Greenwood. Fossils and fossil climate—the case for equable continental interioars in the eocene. *Philosophical Transactions of the Royal Society of London B*, 341:243–252, 1993.

S. L. Wing, L. J. Hickey, and C. C. Swisher. Implications of an exceptional fossil flora for Late Cretaceous vegetation. *Nature*, 363:342–344, 1993.

L.M. Witmer. The extant phylogenetic bracket and the importance of reconstructing soft tissues in fossils. In J. J. Thompson, editor, *Functional Morphology in Vertebrate Paleontology*, pages 19–33, Cambridge, 1995. Cambridge University Press.

J. A. Wolfe. Temperature parameters of humid to mesic forests of eastern asia and relation to forests of other regions of the northern hemisphere and australasia. *United States Geological Survey Professional Paper*, 1106:1–37, 1979.

J. A. Wolfe. A method of obtaining climatic parameters from leaf assemblages. *United States Geological Survey Bulletin*, 2040:1–71, 1993.

J. A. Wolfe. Paleoclimatic estimates from tertiary leaf assemblages. *Annual Review of Earth and Planetary Sciences*, 23:119–142, 1995.

J. A. Wolfe and R. A. Spicer. Fossil leaf character states: Multivariate analysis. In T.P. Jones and N. P. Rowe, editors, *Fossil Plants and Spores: Modern Techniques*, pages 233–239, London, 1999. Geological Society.

J. A. Wolfe and G. R. Upchurch. Vegetation, climatic and floral changes at the cretaceous-tertiary boundary. *Nature*, 324(6093):148–152, 1986.

J. A. Wolfe and G. R. Upchurch. Leaf assemblages across the cretaceous-tertieary boundary in the raton basin, new mexico and colorado. *Proceedings of the National Academy of Sciences*, 84(15): 5096–5100, 1987.

D. W. Woodcock and C. M. Ignas. Prevalence of wood characters in eastern North America: What characters are most promising for interpreting climates from fossil wood? *American Journal of Botany*, 81:1243–1251, 1994.

Z.-Y. Wu and P. H. Raven, editors. *Flora of China*, volume 1–25. Missouri Botanical Garden Press, St. Louis, Missouri, 1994–.

C. Yu and Z. Chen. *Leaf Architecture of the Woody Dicotyledons from Tropical and Subtropical China*. Pergamon Press, Oxford, 1991.

M. Zohary. *Geobotanical foundations of the Middle East*. G. Fischer, Stuttgart, 1973.