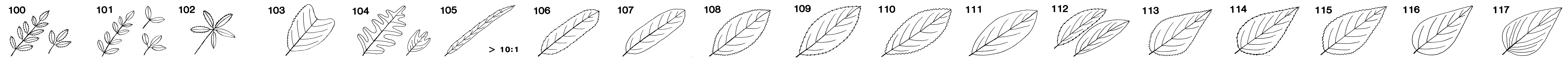
Cartoons shown around margin represent Compendium Index categories



Using Leaf Architecture to Compare Fossil and Modern Forests

Preliminary Results from Hierarchical Cluster Analysis and Graphical Representation

W. A. Green and L. J. Hickey

Department of Geology and Geophysics, Yale University, P. O. Box 208109, New Haven, Connecticut 06520

ABSTRACT

We have developed a methodology for numerically describing dicot forests that allows direct comparison of fossil and modern stands. We use readily observed characteristics of the form and venation of angiosperm leaves as proxies for environment because comparable leaf-architectural data are easily obtainable from imperfectly preserved or insufficiently described fossil and modern floras, and because a strong prima facie case has been made for the presence of an ecological signal in leaf morphology. Descriptions of forest stands based on both the two main published systems of leaf architectural analysis—Climate Leaf Analysis Multivariate Program (CLAMP), and Compendium Index Categories (CICs)—confirm expected relationships among modern forests and allow semi-quantitative analogies to be drawn between, for instance, the Puryear flora from the Lower Eocene of Tennessee and a dry tropical forest currently growing in lowland Cuba. We are working to expand this strategy into a standard methodology for reporting paleoecological data on leaf litter assemblages that (1) provides a readily visualized way of comparing forest ecosystems on a meso- or macroscopic spatial scale, (2) allows us to track vegetational changes through geological time, and (3) enables data from the leaf fossil record to inform explanations of modern vegetation patterns like the similarity of mixed mesophytic forests in Asia and North America.

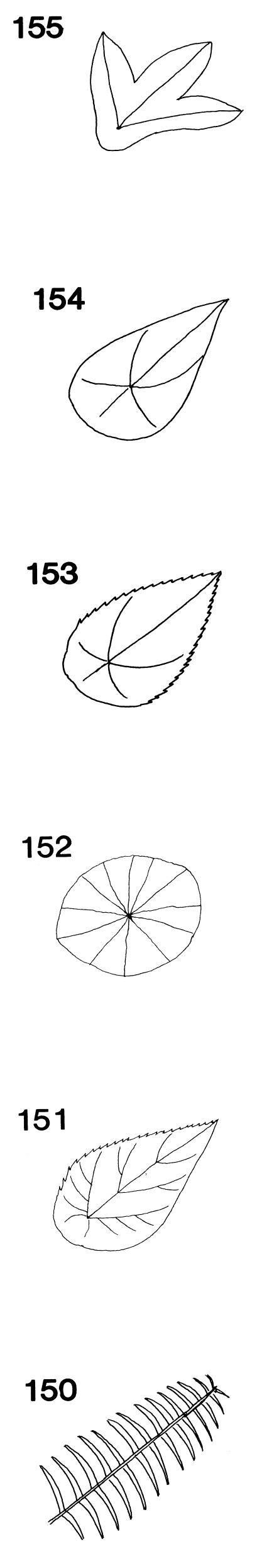
CONCLUSIONS

More detailed conclusions await the full analysis of our data, but even at this the early stage, we can conclude:

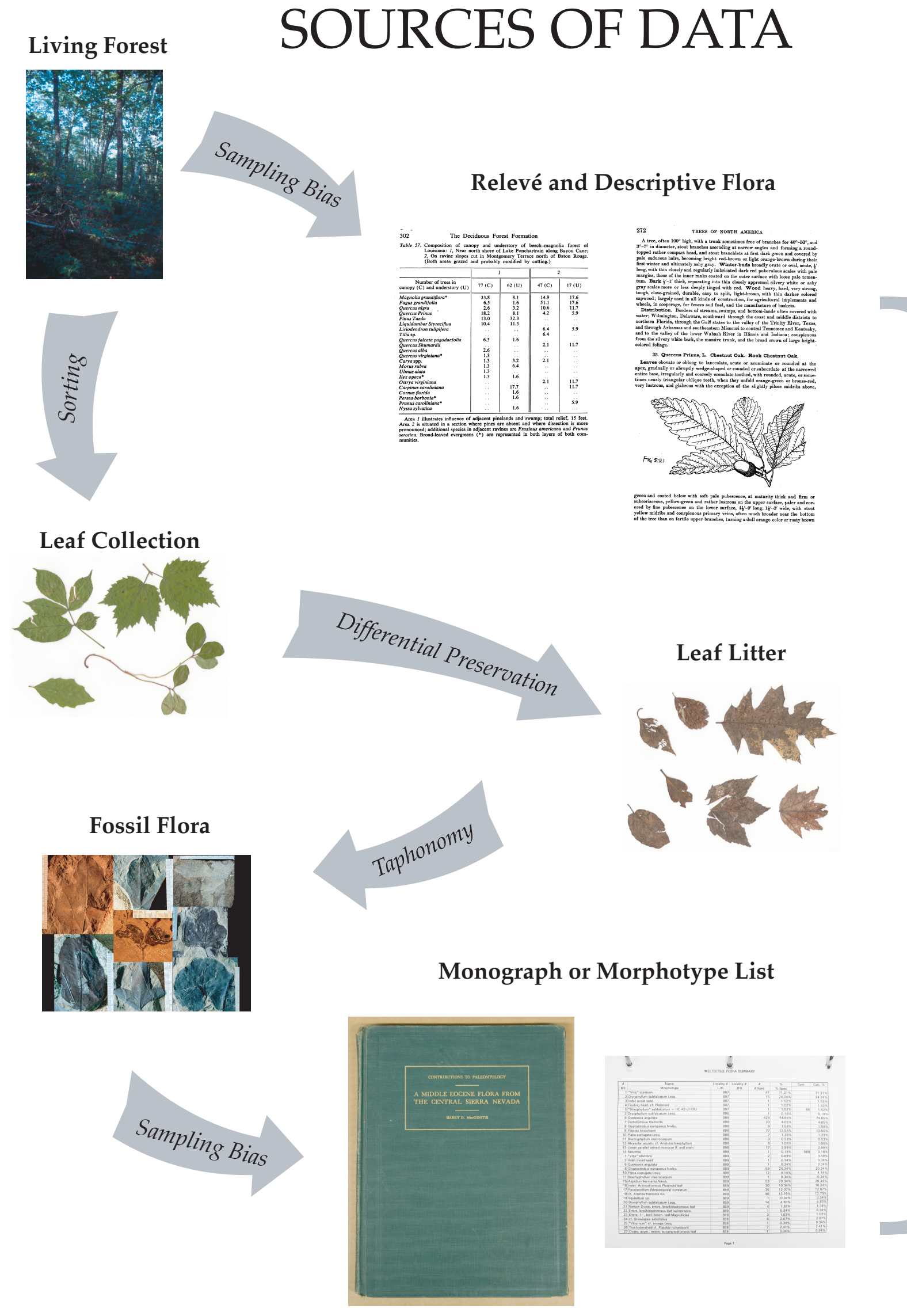
1. A vegetation classification produced by relying entirely on leaf architecture and ignoring traditional taxonomy, nevertheless encapsulates a significant proportion of the variation that is important for large scale vegetaion analysis. This provides an alternative, semi-quantitative method of vegetation analysis that is less subject to the subjective biases of Braun-Blanquet (Montpellier School) phytosociology and related methods.

2. The presence of a strong ecological signal in leaf architecture is verified by our ability to see meaningful biological patterns in leaf architectural data coded and analyzed in different ways. Since the exact method of encoding the leaf architectural data is not critical to the method of analysis; the choice of a coding system can depend on practical concerns of whether data of a particular type can be efficiently obtained.

3. The application of appropriate coding systems allows the direct comparison of fossil and modern floras. For the first time, this provides us with the ability to look at large scale changes in vegetation through geologic time and across continental distances without relying on subjective taxonomic judgements.



SOURCES OF DATA



Living Forest

Sampling Bias

Relevé and Descriptive Flora

Sorting

Leaf Collection

Differential Preservation

Leaf Litter

Taphonomy

Fossil Flora

Monograph or Morphotype List

Sampling Bias

One of two systems is used to assign an n-dimensional vector to each flora (stand): either the species are coded according to CLAMP (the Climate Leaf Analysis Multivariate Program, Wolfe 1993), giving a 31-dimensional vector, or each species is assigned to a CIC (Compendium Index Category, Dorf 1940, Ash et al. 1999), giving a 56 dimensional vector. Cartoons representing these categories are shown around the margin of the poster. In either case, each stand (relevé) or fossil collection locality gets a vector describing the leaf shapes found at that location in space and time. This can be seen as a semi-quantitative elaboration of what Shinwell (1971) called the symbolic tradition in vegetation classification.

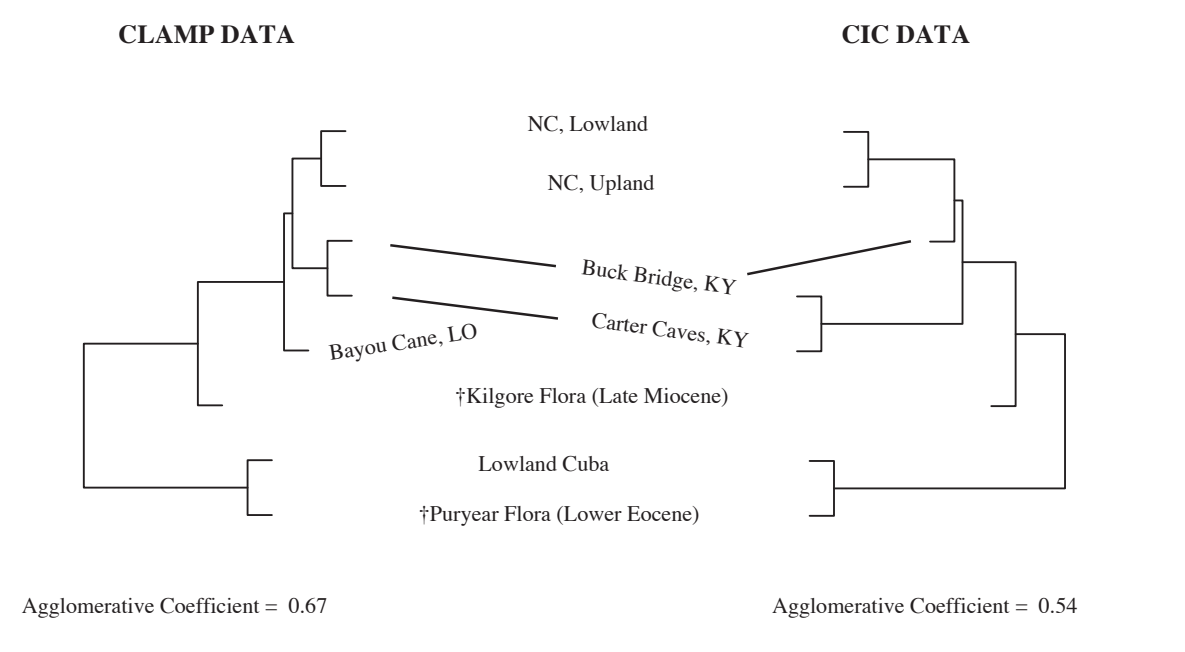
Coding

Note that of the two available (published) systems, CLAMP variables are easier to interpret in ecological terms, but take an order of magnitude longer to obtain than the CIC categories. As is shown to the right, the two sets of variables seem to provide similar classifications indicating that the ecological signal in leaf architecture is strong enough that the choice of coding method is not critical. This being the case, we have focussed on collecting CIC data because it is easier to obtain.

Raw Data Matrix

```
"HawaiiLowland" 1 3 0 0 0 2 0 0 0 0 0.5 4.67 0 0 0 0.5 3.67 0 0 0 0 0 1.67 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
"MauiLowland" 1 2 0 0 0 2 0 0 0 0 0.5 5.17 0 0 0 0.5 4.17 0 0 0 0 0 1.67 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
"LanaiLowland" 1 1 0 0 0 1 0 0 0 0 0.5 5.17 0 0 0 0.5 4.67 0 0 0 0 0 1.17 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
"OahuLowland" 1 1 0 0 0 1 0 0 0 0 0.5 6.67 0 0 0 0.5 4.17 0 0 0 0 0 2.17 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
"HawaiiCloudForest.c" 0.5 0.5 0 0 0 1 0 0 0 0 0 4.17 0 0 0 0 1.67 0 0 0 0 0 3.17 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
"MauiCloudForest.c" 0.5 0.5 0 0 0 1 0 0 0 0 0 6.17 0 0 0 0 2.17 0 0 0 0 0 4.67 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
"OahuCloudForest.c" 1 1 0 0 0 0 0 0 0 0 0.5 0 5 0 0 0.5 0 1.5 0 0 0 0 0 4.5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
"KauaiCloudForest.c" 2 2 0 0 0 0 0 0 0 0 0 6 0 0 0 0 1.5 0 0 0 0 0 4.5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
"HawaiiCloudForest.u" 1 0 2 0 0 0 0.5 0 0.25 2.17 0.58 0.5 0 0.25 0.83 0.25 1 0 0 0 0 1.17 0.5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
"MauiCloudForest.u" 1 0 1 0 0 0 0.33 0 1.25 1.92 1.17 0.5 0 0.25 1.08 0 0 0 0 0 0 2 0.5 0 0 0 0 0 0 0 0 0 0 0 0.33 0 0.33 0 0.33 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
"OahuCloudForest.u" 0 0 1 0 0 0 0 0 0 0 0.67 0.33 2 0 0 0.33 0 0 0 0 0 0 0.67 0 0 0 0 0 0 0 0 0 0 0 0 0.5 0 0 0 0 0.5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
"KauaiCloudForest.u" 0 0 1 0 0 0 0 0 0.5 1.42 2.33 0.75 0 0 0.33 0 0 0 0 0 0 1.92 0.75 0 0 0 0 0 0 0 0 0 0 0 0.83 0.83 0 0 1.33 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

Root-Sum-Square Normaliazation so that each row is a unit vector

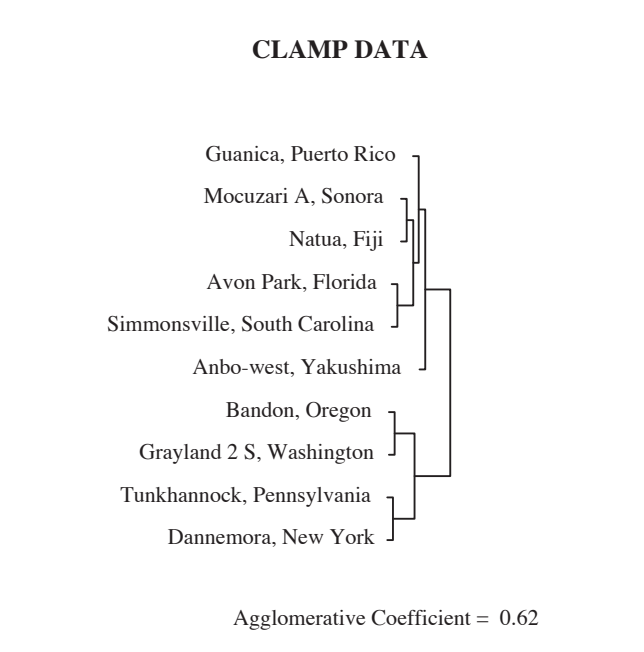


In a pilot set of six modern and two fossil stands (the latter marked with a †), coded by both methods (CLAMP and CICs), almost the same topology was obtained. Note that the fossil Puryear Flora, which has been likened to a dry tropical flora, does indeed appear most similar to a modern dry tropical flora from Cuba. The other floras show approximately the relationships their their geographical proximity would suggest, with the exception of the mis-classification of the Bayou Cane flora from Louisiana in the right-hand tree.

HIERARCHICAL CLUSTER ANALYSIS

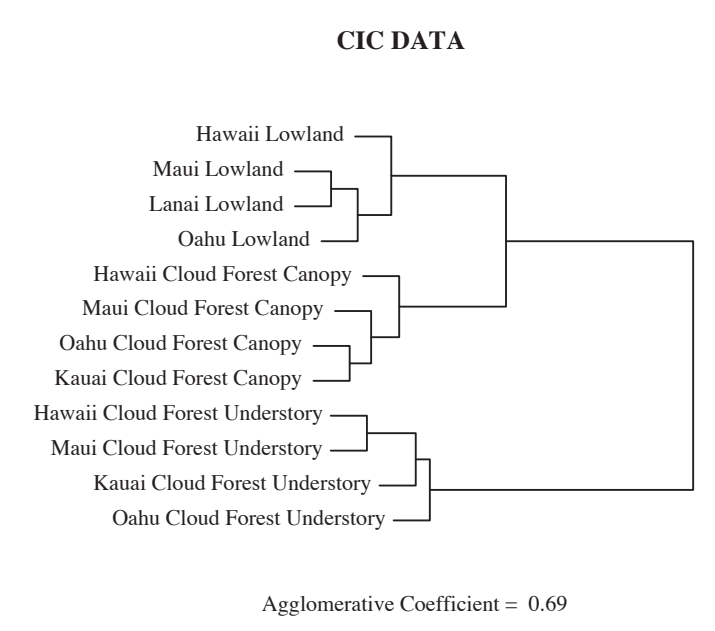
Stands are clustered using one of a number of algorithms to produce dendrograms showing overall 'distance' between stands in *n*–space. All dendrograms shown here were produced using root-sum-square normalization of the raw data and the default settings (euclidean distance metric, average linkage) in the agglomerative hierarchical clustering algroithm AGNES (Kaufmann and Rousseau, 1995) as implemented in the open source statistics and graphics package R (see http://www.r-project.org for more information).

Such trees can show geographical distance:

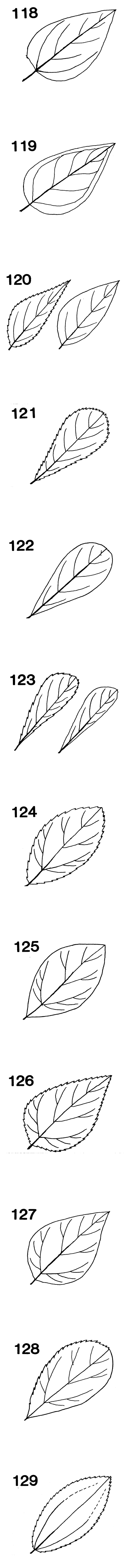


Note how the leaf architectural similarity (as shown by the topology of the dendrogram) correlates approximately with the geographic distance between the coded floras. CLAMP data collected by J. A. Wolfe, which is accessible in electronic format from: http://tabitha.open.ac.uk/spicer/CLAMP//Clampset1.html is used here to indicate that similar analysis can be performed on any multivariate data set that encapsulates infromation about leaf architecture. In general CIC data have been preferred because they are much easier to obtain.

or different ecological roles, independent of geographical distance:

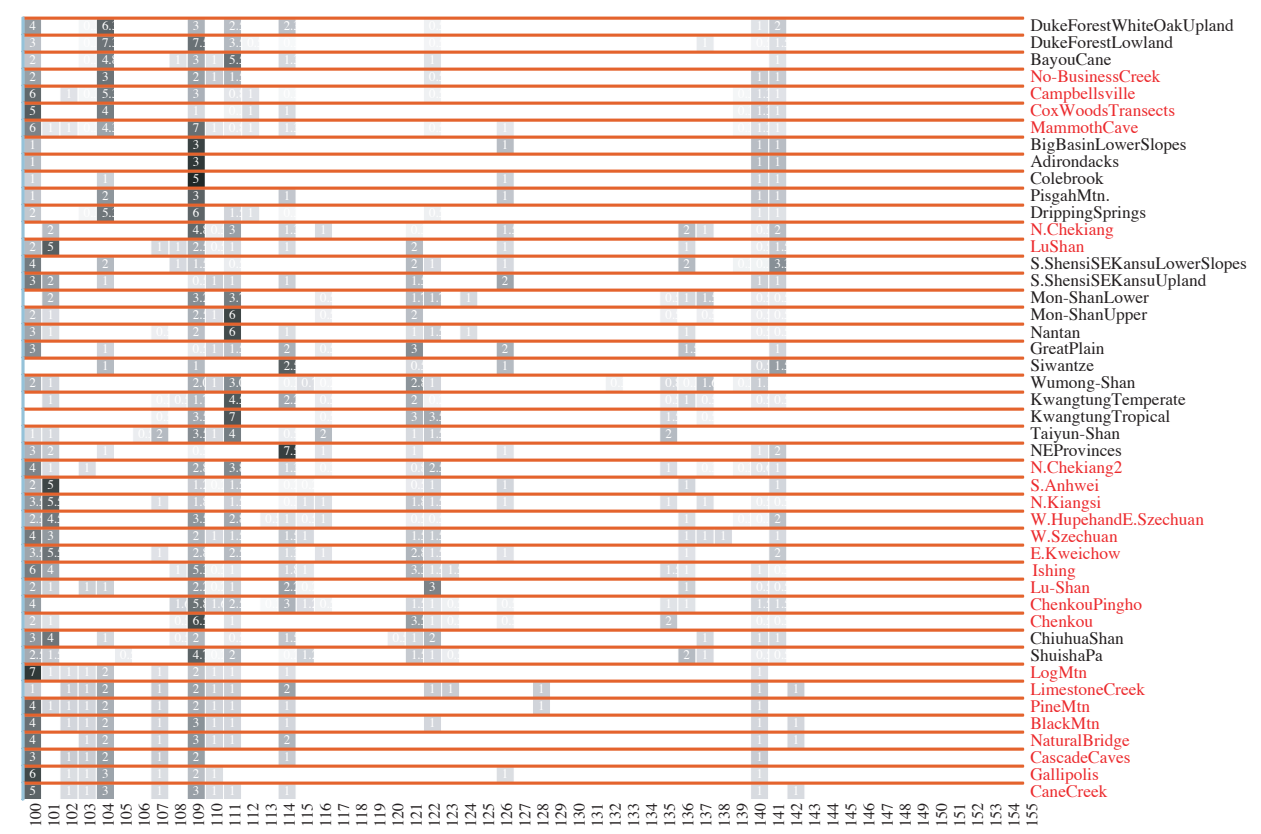


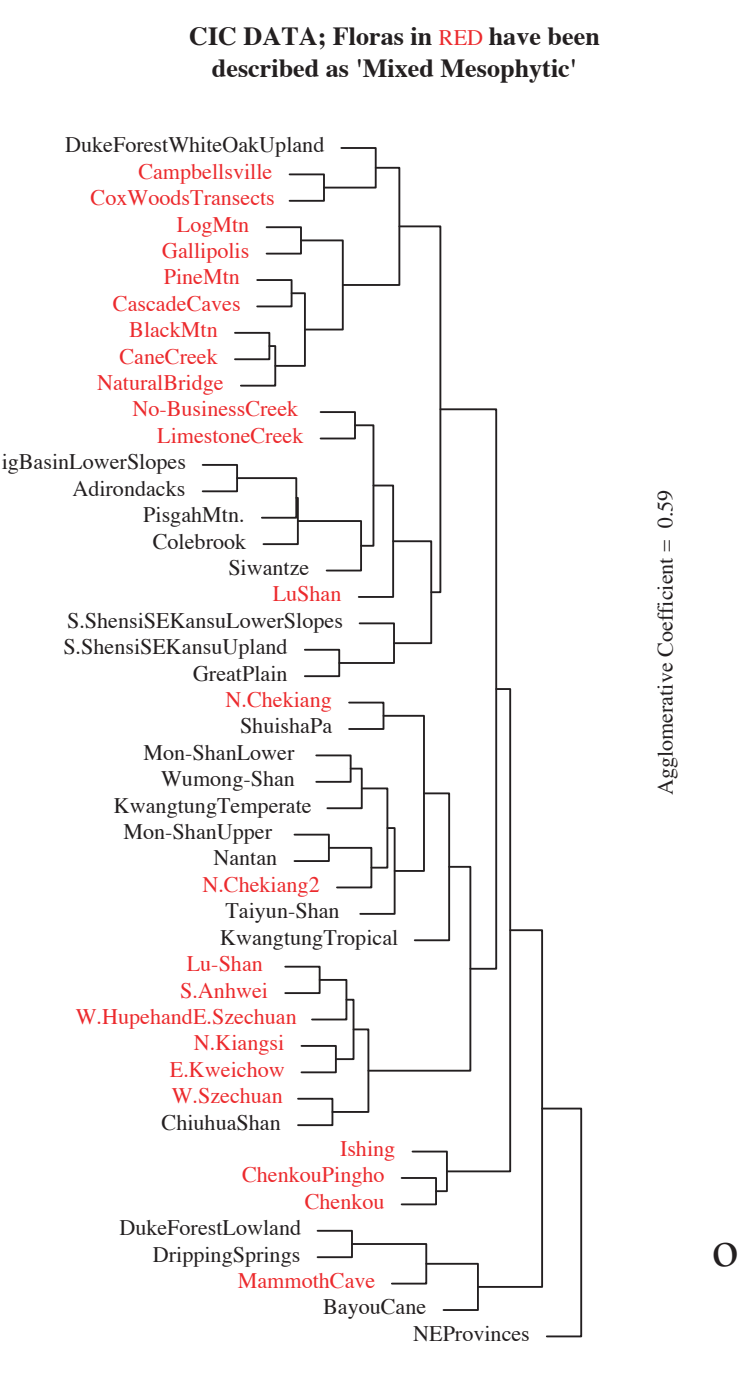
Here, the lowland, upland canopy, and upland understory components of Hawaiian forests are separated in the dendrogram, independent of which island the assemblage came from. Hence, we can conclude that there are cases in which leaf architecture reflects canopy position rather than any edaphic variable. Future investigation should reveal when such cases occur and what leaf characteristics are indicative of canopy position rather than climate.



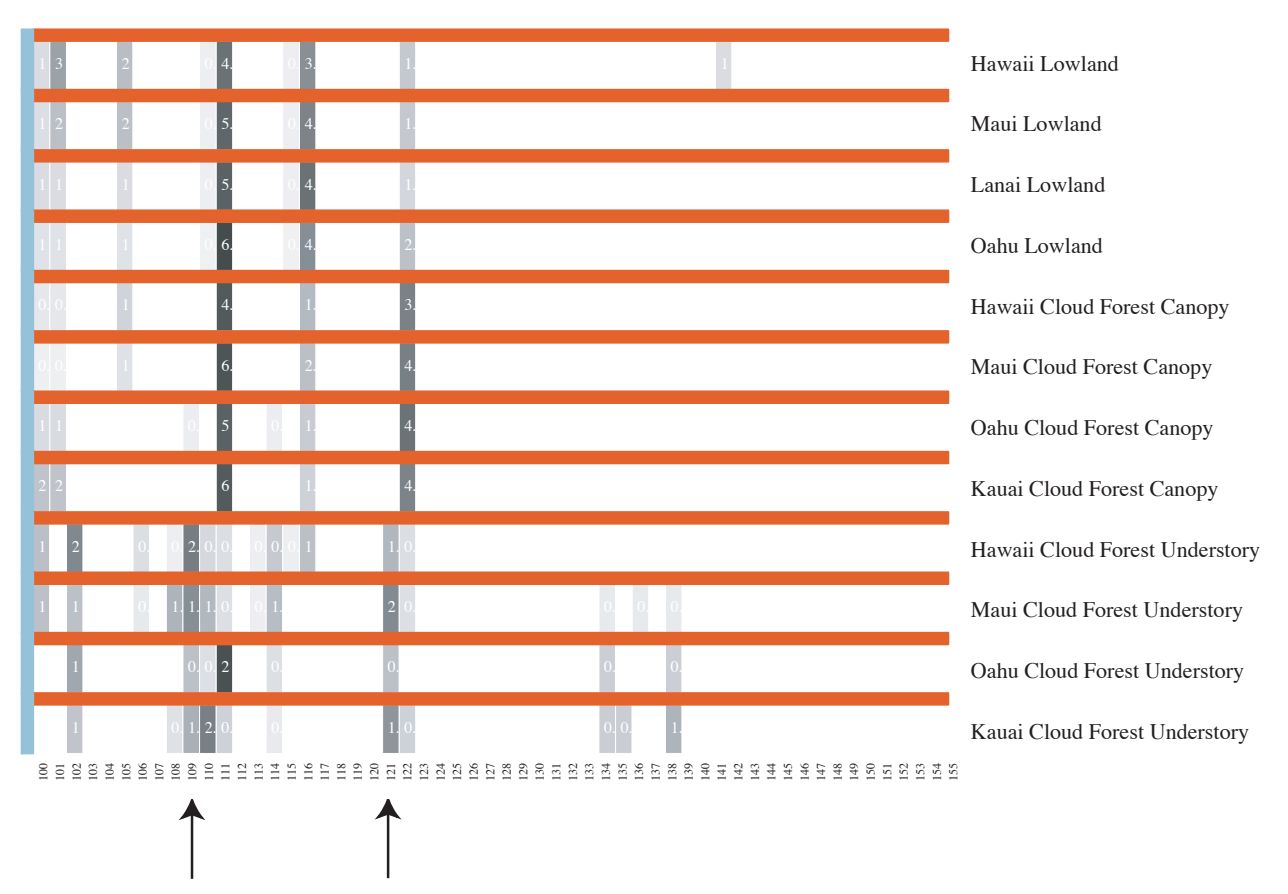
GRAPHICAL REPRESENTATION

The number of species in each flora that represent a given leaf architecture is shown as a block of grey, where black indicates that the form composes a large proportion of the flora and white indicates the absence of that form from the flora. The superposed white figures indicate the absolute number of species counted in each category.





At this preliminary stage, it cannot be determined whether the Mixed Mesophytic forests of northern China and the eastern United States are remnants of a formerly continuous ancient forest or whether they are convergent adaptaion to similar edaphic conditions. It will be possible, however, to separate the leaf architectural characters that unite Mixed Mesophytic Forests, regardless of their geographic position from those that separate Asian from American forests.

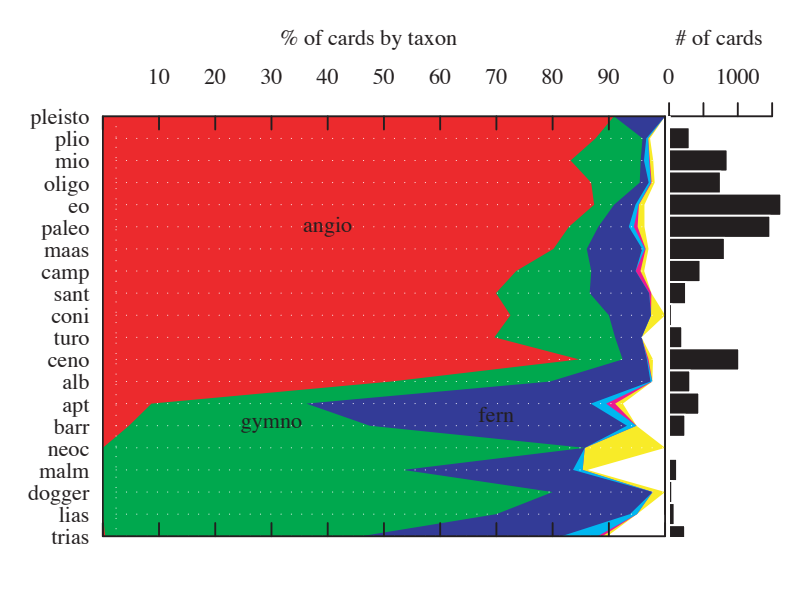


e.g. CIC 102 (*Rubus* spp.) and 121 (*Broussaisia, Vaccinium, Clermontia*) unite the understories

Compare results from graphical analysis and hierarchical cluster analysis: the same groups are apparent. The dendrogram gives a precise, description of the relationships among the floras, but is difficult to interpret; the the relationships apparent from the graphical display are less precise, but are immediately interpretable in terms of the leaf architectural characters that make up the floras.

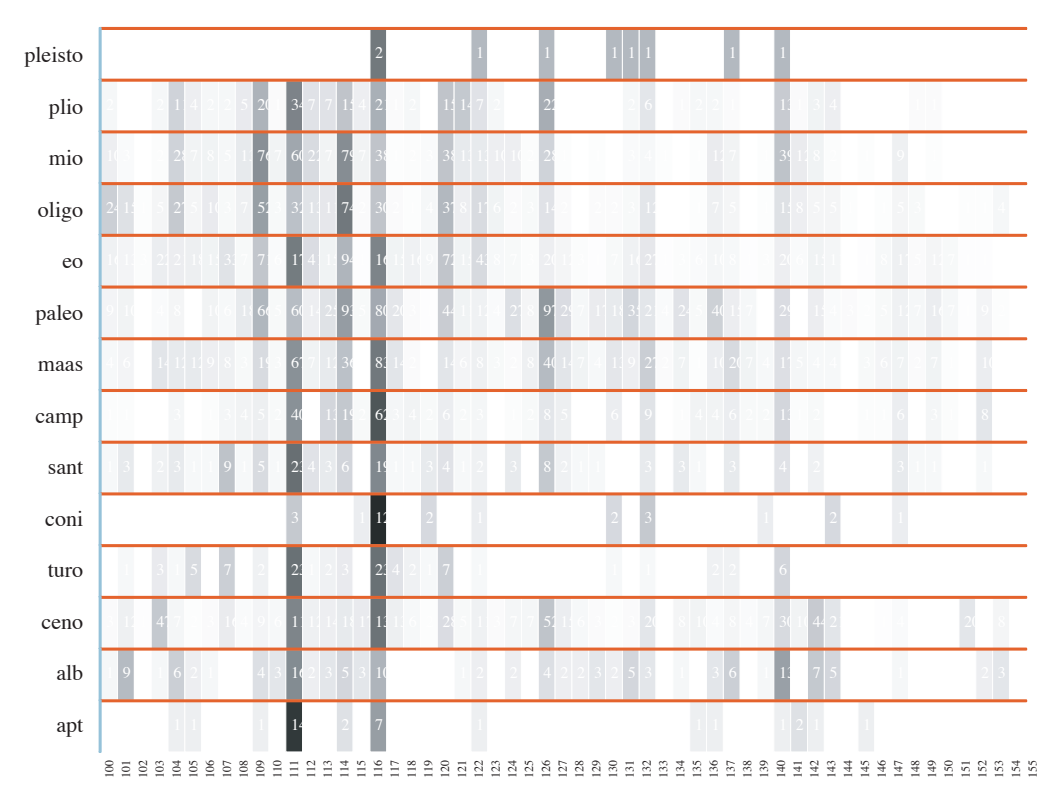
DIACHRONIC CHANGE

The data plotted below come from the Compendium Index, a taxonomic database now housed at the Yale Peabody Museum, to organize which the Compendium Index Categories (CICs) were originally designed (Dorf, 1940).



To the left is a mountian plot and histogram of the numbers of entries in the Compendium Index in each taxonomic Class by age or epoch. At this scale, phylogenetic groups correspond to leaf architectural categories, so a picture of floral change through time can be obtained comparable to that provided by Nikas et al. (1985).

To the right, we can see a progressive increase through the Late Cretaceous and Cenozoic in the importance of the leaf architectural characters characteristic of the modern North American temperate flora. This probably reflects the trend from tropical and subtropical climates, towards the temperate climates that now prevail.



REFERENCES:

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

Dorf, E. (1940) An illustrated Catalogue of Mesozoic and Early Cenozoic Plants of North America *Science* 91(2368):1:2.

Kaufman, L and P. J. Rousseeuw (1990) *Finding Groups in Data* New York: Wiley.

Nikas, K. J., B. H. Tiffney and A. H. Knoll (1985) Patterns in vascular land plant diversification: An analysis at the species level Pp. 97–127 in J. Valentine (ed.) *Phanerozoic Diversity Patterns* Princeton: Princetion Univ. Press.

Shimwell, D. W. (1971) *The Description and Classification of Vegetation* Seattle: Univ. of Washington Press.

Wolfe, J. A. (1993) A method of obtaining climatic parameters from leaf assemblages *USGS Bulletin* 2040.

Thanks to Ms. Emily Schenker for assistance in data collection.
This work was supported in part by a Yale University Fellowship and by NSF grant DEB-0235736