Logistic regression in comparative wood anatomy: tracheid types, wood anatomical terminology, and new inferences from the Carlquist and Hoekman southern Californian data set

JULIETA A. ROSELL[1], MARK E. OLSON[1*], REBECA AGUIRRE-HERNÁNDEZ[2] and SHERWIN CARLQUIST FLS[3]

[1]*Instituto de Biología, Universidad Nacional Autónoma de México, 3er Circuito s/n CU, México DF 04510, Mexico*
[2]*Facultad de Medicina, Universidad Nacional Autónoma de México, Circuito Escolar s/n CU, México DF 04510, Mexico*
[3]*Santa Barbara Botanical Garden, 1212 Mission Canyon Road, Santa Barbara, CA 93105, USA*

Despite collecting copious amounts of data, wood anatomists rarely perform appropriate statistical analyses, especially in the case of categorical variables. Nevertheless, anatomists have succeeded in identifying strong ecological trends. We show that, with only a slightly more sophisticated analysis, the strength and significance of 'well-known' associations can be quantified, and new associations pinpointed. Using logistic regression to reanalyse the classic Carlquist and Hoekman data set for the southern Californian flora, we show strong support for the notion that true tracheid presence lowers vessel grouping; in contrast, vasicentric tracheids are associated with a diversity of vessel grouping strategies. We show that statistical models can refine anatomical interpretations by identifying unusual species. For example, *Fremontodendron californicum* and *Baccharis salicifolia* (= *B. glutinosa*) were identified as unusual in lacking vasicentric tracheids; a consultation of preparations revealed that they are indeed present. For purposes of ecological wood anatomy, anatomical terminology should reflect cell function; we suggest that terminological systems that yield better predictive power in statistical models such as ours are preferable. Finally, we make recommendations ranging from the statistical, e.g. the need to check assumptions and the need for the inclusion of phylogeny, to the biological, e.g. gathering data expressly designed to test functional hypotheses rather than all of the information in standardized lists.   © 2007 The Linnean Society of London, *Botanical Journal of the Linnean Society*, 2007, **154**, 331–351.

ADDITIONAL KEYWORDS: adaptation – categorical data – character correlation – contingency tables – data coding – ecological wood anatomy – evolution – relative risk – vasicentric tracheids.

INTRODUCTION

Comparative wood anatomists use laborious methods to generate large quantities of data. Despite the considerable work involved, statistical procedures have only rarely been used to guide data collection, and analyses of wood anatomical data have been restricted to descriptive statistics or the application of basic techniques that do not extract the maximum amount of information from the data (Burley & Miller, 1982). As a result, there are many opportunities for expanding the use of statistics in the analysis of anatomical information, including strategies for the collection of wood samples from wild populations and the sampling of anatomical information from preparations. However, perhaps the greatest opportunity for improving results while taking advantage of the prodigious amount of information already possessed by wood anatomists is data analysis. In one of the few works to treat statistics in comparative wood anatomy explic-

*Corresponding author. E-mail: explore@explorelifeonearth.org

itly, Burley & Miller (1982) highlighted a variety of techniques that could be used to analyse quantitative anatomical data. However, categorical data frequently arise in comparative wood anatomy. They include nominal (e.g. crystal type, presence/absence of true tracheids) and ordinal (e.g. vessel diameter coded as narrow, medium and wide) variables. Here we complement the work of Burley and Miller by providing an example of the analysis of qualitative data using logistic regression and other statistical methods for categorical data. We illustrate our novel recommendations with a reanalysis of a well-known ecological wood anatomy data set (Carlquist & Hoekman, 1985) to show that, with only a slightly more sophisticated analysis, new associations can be demonstrated, support for previously recognized trends can be quantified, and directions for further research can be identified

Wood identification, systematics, and ecological wood anatomy comprise the three main efforts that can be identified within comparative wood anatomy (Olson, 2005). Although improved statistical treatment could substantially contribute to each of these fields, as the one that relies the most on the statistical treatment of data, ecological wood anatomy stands to benefit most from more sophisticated analytical techniques. Ecological wood anatomy attempts to relate anatomical information with plant habit or environmental variables thought to be important in determining plant structural adaptations. Wood characters can vary remarkably, even between closely related species, in relation to habit and ecology. The repeated documentation of the same correlations between anatomical characteristics and environmental variables or habit (e.g. Carlquist, 1975, 2001; Carlquist & Hoekman, 1985; Arias & Terrazas, 2001; Jansen *et al.*, 2004; Lens *et al.*, 2004), and of the co-occurrence of given suites of anatomical characteristics (e.g. Carlquist, 1984) thus forms the basis of ecological wood anatomy. Whatever the specific variables used, the overall aim is to identify the occurrence of similar anatomical characteristics in similar environmental situations.

Although comparative wood anatomists have typically collected many qualitative data, the association between categorical variables has rarely been subjected to statistical analysis, and when it has, the treatment has been mainly descriptive (e.g. Alves & Angyalossy-Alfonso, 2000, 2002). One of our main aims was thus to highlight statistical methods that can be used with categorical data, and to illustrate their extraordinary utility in comparative wood anatomy with a reanalysis of the data set from Carlquist & Hoekman's (1985) ecological wood anatomy study of the flora of southern California. Although this study has been frequently cited because of the large quanti-

ties of ecological hypotheses and concepts it embodies, the original statistical analysis was, like those of most ecological wood anatomy studies, somewhat superficial. As a result, less obvious trends have remained hidden, but can be recovered with the appropriate statistical methods, such as logistic regression.

LOGISTIC REGRESSION MODELS AND COMPARATIVE WOOD ANATOMY

Although they are rarely used in ecological wood anatomy, methods are available for studying the association between categorical variables. For example, these associations can be examined using contingency tables (Sokal & Rohlf, 1995; Quinn & Keough, 2002), but this procedure is impractical for analysing more than three variables simultaneously. Furthermore, if continuous variables are to be included in the analysis, the contingency table approach cannot be used unless the numerical variables are categorized, a practice that is undesirable for both statistical and biological reasons (Olson, 2005; Royston, Altman & Sauerbrei, 2006). To describe more complex patterns of association between categorical variables, statistical models are available. Apart from modelling association, these models can also predict the probability of the presence or absence of one variable based on the values of other variables. They share many properties with linear regression and analysis of variance, and include logistic regression models.

Logistic models are the most widely used tools for studying the relationship between a binary trait of interest (the response variable) and a set of characteristics thought to influence this trait (the explanatory variables), which can be categorical and/or continuous (Kleinbaum, 2002). It is an ideal statistical technique for modelling the association between the presence or absence of a given type of cell and a set of environmental and anatomical variables. Because several explanatory variables can be included in a given model, logistic regression is ideal for discovering complex patterns of association that cannot be easily extracted with pairwise estimation of correlations, and much less with a descriptive analysis of the data (Kutner *et al.*, 2005). Suppose we are interested in modelling the way in which vessel grouping (*vg*) and the presence or absence of scalariform perforation plates (*sc*) predict the presence of vasicentric tracheids (*i*). We can use a logistic regression model with two explanatory variables (one continuous, *vg*, one categorical, *sc*), and a binary response variable (the presence or absence of *i*). Such a model has the following form:

$$\ln\left[\frac{\pi_i}{1-\pi_i}\right] = \beta_0 + \beta_1 vg + \beta_2 sc$$

The left part of the expression is called the logit transformation of π_i, that is, the logarithm of the ratio of the probability of the presence of vasicentric tracheids (referred to by statisticians as the 'probability of success', π_i) and the probability of absence (the 'probability of failure', $1 - \pi_i$). The probability of success varies from one species to another depending on the value of the two explanatory variables (*vg* and *sc*). This probability is expressed as a nonlinear function of the parameters β_0, β_1, and β_2, which are also known as regression coefficients. After a process of modelling, which is detailed below, a logistic model contains the explanatory variables that influence the response. As we will show, information regarding the direction and magnitude of the associations between each explanatory variable and the response can be extracted from the estimated parameters of the model. Logistic models can thus be used to unveil complex patterns of association between several variables, a property of evident utility in ecological wood anatomy.

REANALYSIS OF THE CARLQUIST AND HOEKMAN ECOLOGICAL ANATOMY DATA SET

Carlquist & Hoekman (1985) used the ecologically and phylogenetically diverse woody flora of southern California to examine variation in anatomical features associated with the conduction of water, such as vessel diameter, grouping, and perforation plate type. They included over 200 species in all of the woody genera then recognized in the region and examined the relationship of anatomical variables with habit (shrub, tree, vine, etc.) and environmental factors represented by habitat categories (e.g. alpine, riparian, woodland, etc.). With this large data set, Carlquist and Hoekman detected trends for the whole flora and illustrated numerous ecological hypotheses that still guide ecological wood anatomy studies to the present.

The original statistical analysis of Carlquist and Hoekman focused mainly on descriptive statistics (means and percentages), but two statistical tests (Mann–Whitney and chi-squared) were mentioned. The general approach was to divide the data into subsets according to the categories of a given qualitative variable. Then, means for the different subsets were calculated for continuous variables. These statistics were then compared with their global counterparts calculated using all of the data. The signs of the differences (positive or negative) were used to determine whether the woods of a given category tend to be xeromorphic or mesomorphic. The process was repeated for all of the qualitative variables in the data set. Although the main focus of the discussion of the paper centred on how mesomorphic or xeromorphic the wood of a given category appeared to be, the authors also highlighted associations between some variables.

Additionally, they commented on categories and species that did not follow the predicted patterns. Significant findings of Carlquist and Hoekman were that low vessel density, wide vessels, low vessel grouping, a lack of vasicentric tracheids, a lack of helical sculpture on vessel walls, and the absence of growth rings are common in woods in moister situations. Because the sign of these associations (positive or negative) has repeatedly been corroborated in many ecological wood anatomy studies, they are often regarded as well known. However, the magnitudes of these associations have rarely or never been quantified, making it impossible to establish comparisons between regions or taxa.

In our reanalysis, we chose to focus on the presence or absence of the different types of conductive imperforate tracheary elements included in the data set, both because they are a major focus of the original work and because the identification of these cells has been the object of a debate among wood anatomists (e.g. Bailey, 1936; Baas, 1986; Carlquist, 1986, 2001, etc.). Using Carlquist's classification (see Carlquist, 2001), the conductive imperforate tracheary elements in this study included true tracheids, vasicentric tracheids, and vascular tracheids. True tracheids have dense, fully bordered pits and are generally present as ground tissue. In general, true tracheids are assumed to be symplesiomorphic with the tracheids of gymnosperms, but this assumption is still controversial (e.g. Feild, Brodribb & Holbrook, 2002). Vasicentric tracheids are so called because they are found only in association with vessels, and are thought to provide a safer, if slower, alternative conductive path to the vessels with which they are associated. In contrast, vascular tracheids are never found in association with vessels and are instead typical of the latewood of soft-leaved shrubs whose leaves, and sometimes branches, desiccate in the summer dry season. Both vasicentric and vascular tracheids are thought to be derived evolutionarily from vessel elements, and are morphologically similar to one another. Despite the different evolutionary origins that these cell types are thought to have, for convenience we refer to them collectively as 'tracheid types' or 'conductive imperforate tracheary elements'.

The Carlquist and Hoekman data set is ideal for examining the associations of tracheid types with other variables because it is one of the few studies to examine these cells and because its sampling covered a wide variety of environmental conditions, anatomical variables, and plant families. In addition to allowing us to show that the analyses employed in traditional wood anatomical studies overlook important information, our reanalysis illustrates important issues in cell classification and data coding. Standard practice in comparative wood anatomy (e.g. IAWA

Committee, 1989) emphasizes the classification of cell types and wood features, e.g. growth ring types, imperforate tracheary element types, ray types, etc. How 'real' or biologically defensible these classifications are could be tested using approaches such as the one we propose. We also show how incorrect coding can lead to a loss of inferential power. Finally, we illustrate how the unusual species pointed out by logistic models as deviating from the general behaviour of the data could be revealing unrecognized anatomical strategies or could be indicating the need for new observations of preparations to verify that the information was collected correctly.

MATERIAL AND METHODS

ADJUSTMENTS TO THE CARLQUIST AND HOEKMAN DATA SET AND DEFINITION OF VARIABLES

All of the 207 species in the Carlquist and Hoekman data set were included in the analysis. The variables considered in this analysis are summarized in Table 1. Two variables required modification, whereas an additional two were deemed unsuitable for inclusion in logistic models and were eliminated. The 'average number of bars in scalariform perforation plate' was converted to a binary presence/absence character because only a very small fraction (5%) of the species had scalariform plates and maintaining it as a numerical variable would have produced a strongly right-

Table 1. Variables in the Carlquist and Hoekman data set that were included in our reanalysis. Variables included continuous vessel characteristics and categorical variables with two possible states (p/a: presence/absence) or with more than two states (growth ring and habit). Tracheid types (vasicentric, true, and vascular) were used as response variables in the analyses
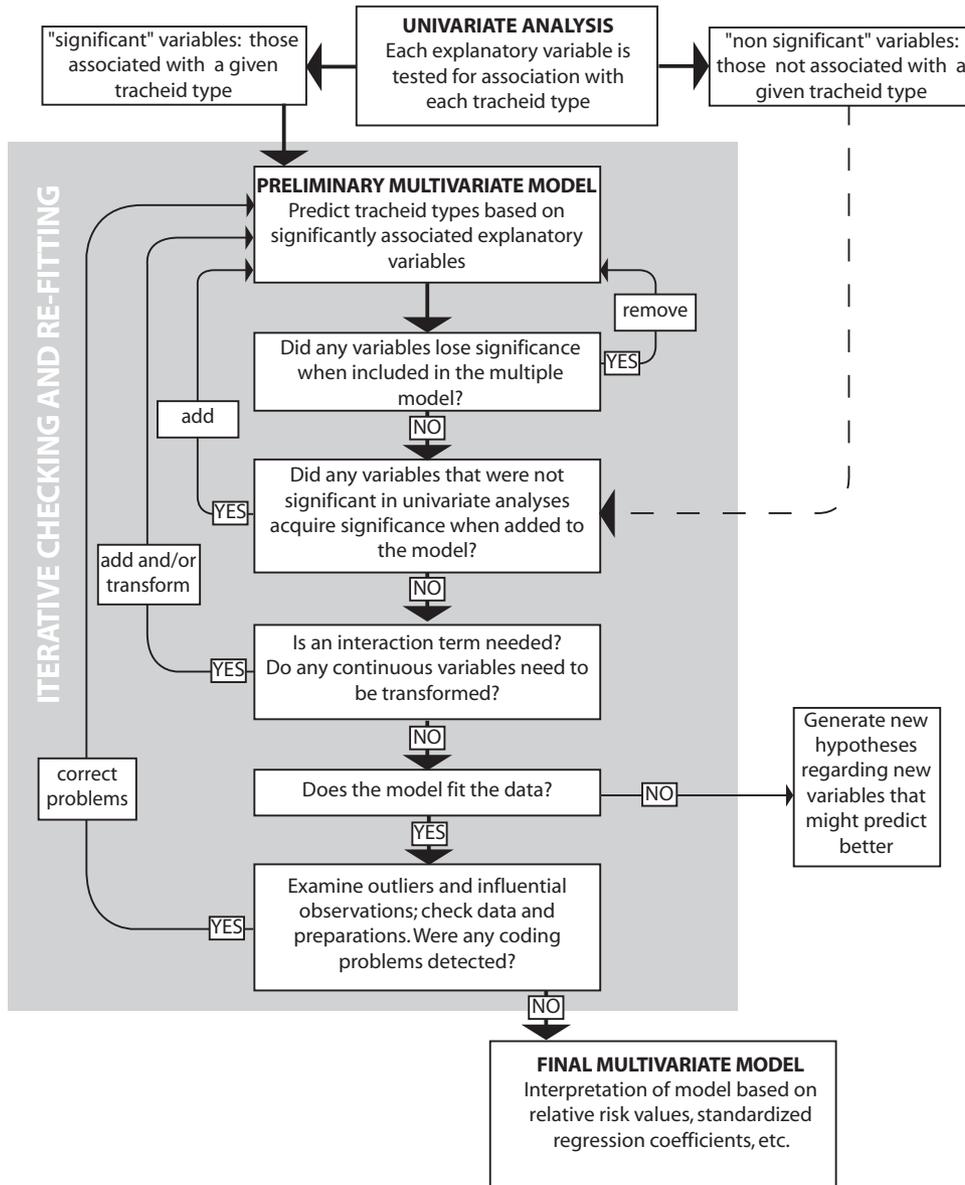
Variable	Type of variable
Vessels mm^{-2}	Continuous
Vessel element length (μm)	Continuous
Vessel diameter (μm)	Continuous
Vessels per group	Continuous
Vasicentric tracheids	Categorical (p/a)
True tracheids	Categorical (p/a)
Vascular tracheids	Categorical (p/a)
Scalariform perforation plates	Categorical (p/a)
Helical sculpture in earlywood	Categorical (p/a)
Helical sculpture in latewood	Categorical (p/a)
Growth ring type	Categorical (ring porous, semiring porous, and diffuse porous)
Habit	Categorical (shrubs, subshrubs, other)

skewed distribution of the data, with most of the species having zero bars. 'Helical sculpture' was separated into two variables: helical sculpture in earlywood and helical sculpture in latewood. 'Habitat' and 'mesomorphy index' were not included in the analyses. For 16 species the habitat was not assigned, and the distribution of the 181 remaining species among the 11 habitat categories was so uneven that the estimation of some parameters would probably have been inaccurate because there would have been too few data in some categories to serve as a basis for inference (Collett, 2003). The mesomorphy index was eliminated because it was computed based on continuous variables of the same data set and would therefore be a source of collinearity in the logistic models.

The original data set had a problem with the coding of vascular tracheids. When a given species had both vascular and vasicentric tracheids, it was coded by Carlquist and Hoekman as having vasicentric but not vascular tracheids. Therefore, some woods that have vascular tracheids are coded as lacking them. Although they did not mention how many species were coded in this way, Carlquist & Hoekman (1985) stated that they were few. Because apparently relatively few species were incorrectly coded and because the patterns of association of vascular tracheids and the other variables were also of interest, we did not discard vascular tracheids in the analyses.

FITTING OF LOGISTIC REGRESSION MODELS

The aim of our analyses was to obtain three models, one for each of the tracheid types. Each model attempts to predict the presence or absence of a tracheid type based on the rest of the nontracheid type variables. None of the tracheid types was used to predict the presence or absence of another one because, as coded by Carlquist and Hoekman, they are mutually exclusive, and such variables can cause numerical problems when the regression coefficients and their standard errors are estimated (Hosmer & Lemeshow, 2000). We outline the general strategy adopted for the fitting of the logistic models (see Fig. 1) before treating each step in greater detail. Broadly, the model-building process involved three main steps, which are very similar to those taken when fitting a multiple regression model: (1) univariate analyses; (2) construction of a preliminary multiple logistic model; and (3) construction of the final multivariate logistic model (Hosmer & Lemeshow, 2000; Quinn & Keough, 2002). The iterative nature of this process is designed to eliminate uninformative variables and identify and correct problems with the data or models, and is illustrated graphically in Figure 1. Each of these steps, which were carried out using the PC program STATA vs. 7 (Stata Corpora-

UNIVARIATE ANALYSIS
Each explanatory variable is tested for association with each tracheid type

"significant" variables: those associated with a given tracheid type

"non significant" variables: those not associated with a given tracheid type

ITERATIVE CHECKING AND RE-FITTING

PRELIMINARY MULTIVARIATE MODEL
Predict tracheid types based on significantly associated explanatory variables

Did any variables lose significance when included in the multiple model?

remove

YES

NO

add

Did any variables that were not significant in univariate analyses acquire significance when added to the model?

YES

NO

add and/or transform

Is an interaction term needed? Do any continuous variables need to be transformed?

YES

NO

correct problems

Does the model fit the data?

NO

Generate new hypotheses regarding new variables that might predict better

YES

Examine outliers and influential observations; check data and preparations. Were any coding problems detected?

YES

NO

FINAL MULTIVARIATE MODEL
Interpretation of model based on relative risk values, standardized regression coefficients, etc.

Figure 1. Flow chart illustrating the fitting process of logistic regression models, including the iterative checking and re-fitting procedure before arriving at the final multivariate logistic model.

tion, College Station, Texas, USA), is described in more detail below.

Initial steps in model construction: univariate analysis
The initial phase of model building used descriptive statistics, contingency tables, independence tests, and univariate logistic models to decide which variables should be included in the multiple logistic models and which should be excluded. To determine which variables were associated with the response variables, we applied tests of independence (or lack of association). Two variables are said to be independent if the value of one of them cannot be used to predict the value of the other one. The variables that were found to be associated with a tracheid type were entered into the preliminary multivariate models (see Fig. 1). These initial analyses are frequently skipped when fitting a multivariate model. This is unfortunate because they provide valuable information regarding the association between variables, as we will show for the Carlquist and Hoekman data set.

When the explanatory variable was continuous, we used univariate (simple) logistic regression models to test for independence, whereas if it was categorical we used contingency tables. For continuous explanatory variables, the null hypothesis of independence between the response and the explanatory variable was not rejected when regression coefficients equalled zero. The significance of the regression coefficients was tested using the Wald statistic (Quinn & Keough, 2002). When the explanatory variable was categorical, a logistic model yielded the same results as the contingency table approach. We preferred contingency tables for cases with categorical explanatory variables because they allow the examination of how frequent each combination of characters in the data set is. This is important because small frequencies can produce unreliable estimates for the logistic model. Based on contingency table results, we tested for independence using one of two methods. If some frequencies in the contingency table were small, the Fisher exact test was computed; otherwise a Pearson chi-squared test was used (see Sokal & Rohlf, 1995).

Rejecting the null hypothesis of independence suggests an association between the response and categorical explanatory variables, but cannot quantify the degree of association. Toward this end, we computed the relative risk (RR). This statistic is defined as $RR = p_1/p_2$, where p_1 and p_2 are the proportions (or probabilities) of success of two groups classified in different categories of an explanatory variable. RR indicates how much more probable it is to find a given trait in group 1 as compared with group 2. The terminology of 'risk' is of relevance in medical applications, and may sound odd when applied to wood, but the idea of the relative probability of the presence of a given variable between two groups is useful for our purposes. For example, the percentage of species that had vasicentric tracheids when the wood was ring porous was $p_1 = 45\%$ (41/91), whereas it was $p_2 = 25.3\%$ (24/95) when the wood was semiring porous. In this case, RR equalled $45\%/25.3\% = 1.78$. Therefore, the percentage of species with vasicentric tracheids is almost doubled when the wood is ring porous instead of semiring porous. A RR close to unity would have suggested that the probabilities of the presence of vasicentric tracheids were independent of the growth ring type.

Preliminary multivariate logistic model
All the explanatory variables individually associated with the response were incorporated simultaneously into multiple (multivariate) models (Fig. 1). In cases in which two or more explanatory variables were associated with one another, some that were significant in the univariate analysis became unimportant. Including associated explanatory variables did not provide additional information in the model and could have

caused problems of collinearity in the estimation procedure, so variables that no longer appeared to be significant were removed from the multiple model. We also incorporated one by one the variables that were unrelated to the response in the univariate analyses to see if they would become significant in the presence of others. Next, we determined whether a continuous variable needed to be transformed or if interaction terms needed to be included in the model (Hosmer & Lemeshow, 2000). To test whether the effect of an explanatory variable, an interaction term, or a transformation was significantly different from zero, we used the Wald and likelihood ratio tests. The latter test was used when the significance of two or more regression coefficients was tested simultaneously, whereas the Wald statistic was used when the significance of a single regression coefficient was tested.

We examined how well our preliminary logistic models described the data using goodness-of-fit statistics and a detailed examination of individual data. Goodness-of-fit statistics are based on the discrepancies between the observed values of the response variable and its estimated values from the fitted model. Because our models included one or more continuous explanatory variables, we used Hosmer and Lemeshow's goodness-of-fit statistic (Hosmer & Lemeshow, 2000). We also examined each data point to detect unusual observations that deviated from the general trend described by the models or that exerted a strong influence on model fit. In the first case, we looked for standardized residuals greater than three, irrespective of the sign, to detect atypical or outlying observations. The model was not considered invalid if 5% or fewer of the standardized residuals were greater than three (Collett, 2003). In addition, we detected influential observations using Pregibon's index. Influential observations have an undue impact on the estimated regression coefficients because of their distance from the remaining data in the values of their explanatory variables. Pregibon's index (Pregibon, 1981) identifies these observations by measuring the overall change in the coefficients when an observation is omitted from the data set. We directed special attention to influential observations to determine whether the anatomical preparations required reinterpretation, the data were recorded correctly, or the samples represented previously unrecognized anatomical modes or features. In the first two situations, the data were corrected and the model was fitted again (Fig. 1).

Final multivariate logistic model and interpretation
The iterative process of checking and refitting the models ended when the fit of the model could not be improved (Fig. 1). The models thus obtained are referred to as the final multivariate logistic models, and it is upon these that our interpretations were

based. These interpretations employed standardized regression coefficients, estimated probabilities, and RR based on them.

Within a given model, it is important to identify which variables are most strongly associated with the response variable. For example, is vasicentric tracheid presence more strongly associated with vessel diameter or vessel grouping? Traditionally, a measurement of the degree of association, known as the odds ratio, is computed directly from the estimated regression coefficients in the model. However, because the odds ratio is difficult to interpret, we preferred to examine the sign and magnitude of the regression coefficients and to calculate the estimated probabilities of finding each tracheid type. Regression coefficients reflected the degree and direction of association between an explanatory variable and the response. The sign indicated whether the association was positive or negative between the explanatory variable and the probability of the presence of a particular tracheid type, whereas the magnitude indicated the strength of the association. However, it would have been misleading to compare the magnitude of two regression coefficients to determine which is more strongly associated with the probability of presence when the explanatory variables were measured in different units. To achieve valid comparisons, we standardized the regression coefficients by multiplying each by the standard deviation of its corresponding explanatory variable (see Agresti, 2002).

We also interpreted the model using the estimated probabilities of finding each tracheid type for different species. For example, the probability of having vasicentric tracheids can be computed for two species with helical sculpture in latewood but differing in vessel grouping. These two probabilities were compared graphically and using RR.

RESULTS

This section is arranged in three main parts. First we report descriptive statistics for continuous vessel characteristics, highlighting species with extreme values. The second part reviews the occurrence and exclusive relationship between the three types of conductive imperforate tracheary elements. The final and main part of this section is devoted to the construction of the logistic regression models that predict the presence or absence of the different conductive imperforate tracheary elements based on the other variables. We have divided this last part into three subsections, one for each of the conductive imperforate tracheary element types. Each subsection is further subdivided into the three modelling steps described above.

DESCRIPTIVE STATISTICS FOR CONTINUOUS CHARACTERISTICS OF VESSELS

The histograms in Figure 2A–D show that all four continuous vessel characteristics had an asymmetrical distribution, with small values being more frequent than large ones. Vessel density for 75% of the species was less than 341 vessels mm^{-2} (Fig. 2A), and species such as *Romneya coulteri* and *Fremontodendron californicum* had an atypically large number of vessels mm^{-2} (1350 and 875 vessels mm^{-2}, respectively). The largest recorded average vessel diameter in the data set belonged to the liana *Vitis girdiana*, with 186 μm. The most highly asymmetrical distribution for vessel characters was found in vessel grouping (Fig. 2D). Nearly 77% of the species had between one and three vessels per group, with the largest value coming from *Romneya coulteri*, with 150 vessels per group. Additional summary statistics for continuous vessel characters are shown in Table 2.

Because of the non-normal distribution of the continuous vessel characteristics (Fig. 2A–D) and the presence of the extreme data mentioned, Spearman's rank correlation coefficient was used to measure the degree of association between each pair of continuous vessel characteristics (Table 3). Only the association between vessel diameter and vessel density was of noteworthy magnitude and statistically significant, showing that fewer vessels mm^{-2} are found when vessels are wider.

Table 2. Descriptive statistics for continuous vessel characteristics. The strong differences between medians and means were due to extreme values that influenced the means strongly

Vessel characteristic	Minimum	Maximum	Median	Mean	Standard deviation
Vessels mm^{-2}	10.8	1350	191	250.97	210.32
Vessel element length (μm)	53.1	643	208	233.17	110.3
Vessel diameter (μm)	12.5	186	29.4	35.19	20.22
Vessels per group	1	150	1.95	3.86	10.97
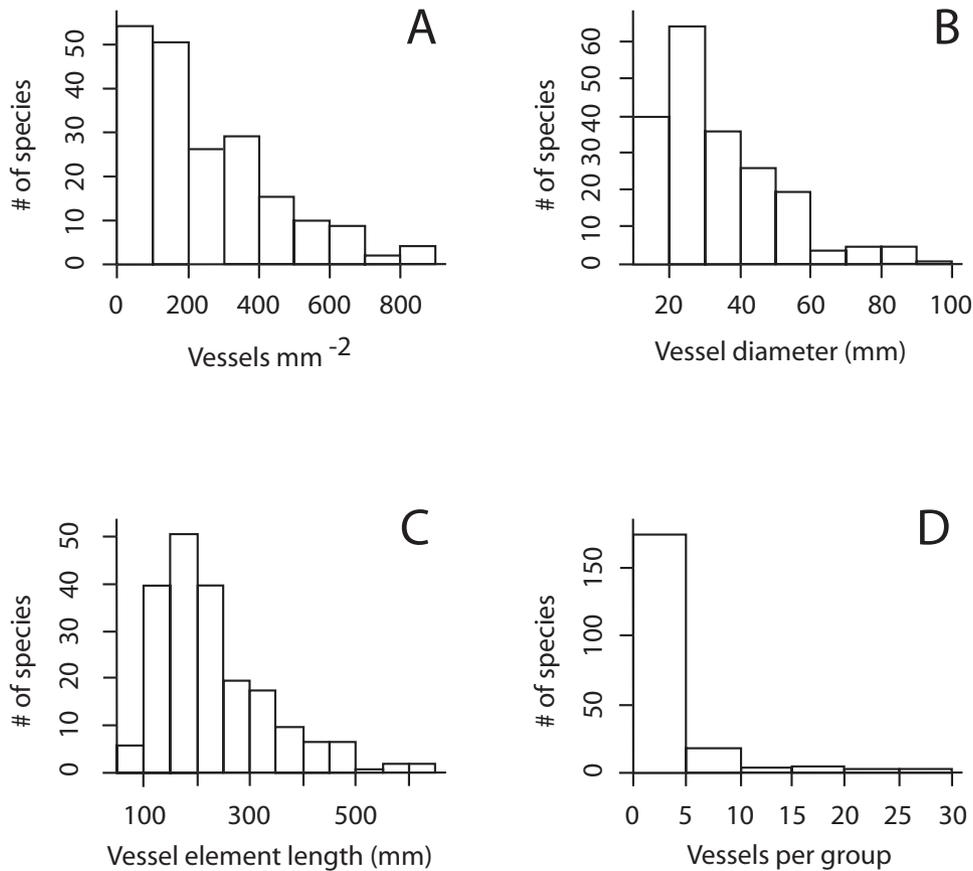
Figure 2. Histograms showing the strongly skewed distributions of continuous vessel characteristics. A, distribution of vessel density, with most species having relatively low density. B, distribution of vessel diameter, with most species having diameters smaller than 60 μm. C, distribution of vessel element length, showing that most species have vessel elements shorter than 300 μm. D, distribution of vessels per group, with its extremely skewed histogram in which the majority of species have fewer than five vessels per group. The following species with extreme values are not included in the histograms: for vessel density and vessels per group, *Romneya coulteri* (1350 vessels mm^{-2} and 150 vessels per group); for vessel diameter, *Vitis girdiana* (186 μm).

Table 3. Spearman's rank correlation coefficients between pairs of continuous vessel characteristics. All correlations were weak, except for that between vessel density and vessel element diameter

	Vessels mm^{-2}	Vessel element length (μm)	Vessel element diameter (μm)
Vessel element length (μm)	−0.13*		
Vessel element diameter (μm)	−0.75**	0.1	
Vessels per group	0.12*	−0.21**	−0.07

**$P < 0.005$; *$P < 0.10$.

OCCURRENCE AND RELATIONSHIPS BETWEEN IMPERFORATE TRACHEARY ELEMENT TYPES

Table 4 shows how the three types of conductive imperforate tracheary element were distributed among species. Vasicentric tracheids were the most common, with 33.3% (69/207) of the species, followed by true tracheids with 23.7% (49/207). Vascular tracheids were observed in 13% (27/207) of the species, but this was an underestimation caused by the coding problem mentioned earlier. Close to 30% (62/207) of the species examined had no subsidiary conduction system and rely exclusively on vessels. Because of the apparently small number of species in which vascular and vasicentric tracheids co-occur, even if vascular tracheids were correctly coded, the three types of cell would not be considered statistically independent,

because the presence of any tracheid type almost always implies the absence of the other two. We did not test for independence between the three tracheid types because several expected frequencies were lower than five, violating the assumptions of Pearson's chi-squared test. Therefore, we combined vasicentric and vascular tracheids, the first two columns of Table 4, into a single variable. We then tested for independence between this new variable and true tracheids. Pearson's test showed that the presence or absence of true tracheids depended strongly on the presence or absence of vascular–vasicentric tracheids ($\chi^2 = 55.10$, 2 d.f., $P < 0.001$).

LOGISTIC MODELS PREDICTING THE PRESENCE OR ABSENCE OF CONDUCTIVE IMPERFORATE TRACHEARY ELEMENT TYPES

The fitting of the logistic models for each type of conductive imperforate tracheary element is described in three subsections. Within each of these subsections,

Table 4. Distribution of conductive imperforate tracheary elements in the species of the data set. Vasicentric and vascular tracheids were combined into a single variable to yield a two-dimensional contingency table. The strong differences between observed and expected frequencies (shown in parentheses and calculated under the hypothesis of independence between variables) suggested a marked lack of independence between tracheid types

Vasicentric tracheids	Vascular tracheids	True tracheids Absence	Presence	Total
Absence	Absence	62 (84.7)	49 (26.3)	111
Absence	Presence	27 (20.6)	0 (6.4)	27
Presence	Absent/ present	69 (52.7)	0 (16.3)	69
Total		158	49	207

we first report the results of analyses of association between the response and explanatory variables, which guided decisions regarding which variables would be included in a preliminary multivariate logistic model. We start with the fitting of univariate logistic regression models for continuous variables and then report the results of contingency table analyses that detected association between the response and the categorical explanatory variables. In no case was it necessary to include interaction terms or to transform continuous variables. On the basis of the preliminary model, we calculated Pregibon's index, which permitted the identification of species with a strong influence on the estimated regression coefficients. In many cases we referred to the original preparations, and in some instances changed the coding of these species for the final analyses, which are described at the end of each subsection. These changes in interpretation of anatomical preparations guided by the logistic models represent one of the most noteworthy results from our analysis, and are treated further in the discussion section. The last subsection indicates how the final logistic multivariate model was interpreted.

Vasicentric tracheids as the response variable
Initial steps in model construction: univariate analysis: The simple logistic regression models fitted with continuous vessel characteristics as explanatory variables showed that vasicentric tracheids were more likely to be found as the number of vessels per group or vessel density increased. Conversely, it was less likely to find them as vessel diameter or length increased (Table 5).

All of the categorical explanatory variables were strongly associated with vasicentric tracheids. First, with regard to perforation plate type, vasicentric tracheids were found in 35.2% (69/196) of the species with simple perforation plates, whereas they were absent in all of the 11 species with scalariform plates. The strong negative association between vasicentric trac-

Table 5. Univariate logistic models for each tracheid type (vasicentric, true, and vascular) as response variables and the continuous vessel characteristics as explanatory variables

Explanatory variable	Vasicentric tracheids Coefficient (SE)	*P*-value	True tracheids Coefficient (SE)	*P*-value	Vascular tracheids Coefficient (SE)	*P*-value
Vessels per group	0.11 (0.04)	0.011	–3.07 (0.55)	< 0.001	0.002 (0.017)	0.885
Vessel diameter	–0.04 (0.01)	< 0.001	–0.05 (0.02)	< 0.001	0.02 (0.01)	0.057
Vessel element length	–0.003 (0.002)	0.057	0.01 (0.001)	0.001	–0.004 (0.002)	0.075
Vessels mm^{-2}	0.001 (0.001)	0.096	0.003 (0.001)	< 0.001	–0.002 (0.001)	0.184

SE, standard error of the logistic regression coefficient.

heids and scalariform perforation plates was statistically significant (one-sided Fisher exact test with $P = 0.01$).

With regard to habit, trees, herbs, and vines were combined into a single habit group to avoid categories with small frequencies. Shrubby species were most abundant, accounting for 66.67% (138/207). Vasicentric tracheids were common in smaller life forms, being found in 39.13% of shrubs (54/138) and 31.25% of subshrubs (10/32). Only 13.51% (5/37) of the other types of habit had such cells. Thus, vasicentric tracheids were more probable in shrubs (RR = 2.89) and subshrubs (RR = 2.31) than in the rest of the habits. The association between habit and the presence of vasicentric tracheids was corroborated by Pearson's chi-squared test ($\chi^2 = 8.69$, 2 d.f., $P = 0.013$). Carlquist and Hoekman's data set contained two vines, 16 herbs, and 19 trees. They represented 18% of the studied habit types, but just 6% of them had conductive imperforate tracheary elements. Therefore, although these habits are obviously distinct, for the purposes of our analysis they are virtually indistinguishable. To avoid table cells with very low values, we thus decided to combine them into a single category before examining the association between habit type and the response variables.

Growth ring type was also closely associated with the presence or absence of vasicentric tracheids. These cells were present in 45% (41/91) of the species with ring porous wood, whereas only 25.3% (24/95) and 19.0% (4/21) of those with semiring porous or diffuse porous wood had them. Vasicentric tracheids were approximately twice as probable in species with ring porous wood than in species with semiring porous (RR = 1.78) or diffuse porous wood (RR = 2.36). The association between growth ring type and the presence or absence of vasicentric tracheids was supported by Pearson's chi-squared test ($\chi^2 = 10.34$, 2 d.f., $P = 0.006$).

A variable very strongly associated with vasicentric tracheids was the presence of helical sculpture, both in early- and latewood. In turn, helical sculpture in earlywood and latewood was also strongly associated with one another ($\chi^2 = 128.79$, 1 d.f., $P < 0.001$). There was no case in which helical sculpture was present in earlywood without being present in latewood. Vasicentric tracheids were present in 50% (41/82) of species with helical sculpture in latewood, whereas just 22.4% (28/125) of those without helical sculpture presented such cells. Thus, the percentage of species with vasicentric tracheids was doubled in the presence of helical sculpture (RR = 2.23). Pearson's chi-squared test was significant for the association between vasicentric tracheids and helical sculpture in latewood ($\chi^2 = 16.97$, 1 d.f., $P < 0.001$). Likewise, the 11 species with scalariform perforation plates lacked both helical sculpture

(except *Philadelphus microphyllus*, which had latewood sculpture) and vasicentric tracheids.

Preliminary multivariate logistic model: Because of the strong association between the two types of helical sculpture, little was gained when both were used to predict the presence or absence of vasicentric tracheids. In fact, the inclusion of both variables could have led to collinearity and the consequent unstable regression estimates, and therefore only one was included. We chose helical sculpture in latewood because all species with helical sculpture showed it in latewood, but not all had it in earlywood, thereby including all species with helical sculpture.

Growth ring type was not highly significant after helical sculpture in latewood was included as an explanatory variable (likelihood ratio test statistic = 5.80, 2 d.f., $P = 0.055$). When the average vessel diameter was added to the model, growth ring type lost all predictive contribution and was eliminated from the model (likelihood ratio test statistic = 3.87, 2 d.f., $P = 0.144$). Finally, the number of vessels per group was added to the model.

The preliminary multiple logistic model predicted the presence or absence of vasicentric tracheids based on helical sculpture in latewood, vessel diameter, and the number of vessels per group. Pregibon's index detected two species that strongly influenced the model fit, *Baccharis salicifolia* and *Fremontodendron californicum*. Both deviated from the general trend of the data in having very high vessel grouping but lacking vasicentric tracheids. All species with more than 11 vessels per group exhibited vasicentric tracheids, except for *Baccharis salicifolia* and *Fremontodendron californicum*, which despite having 29 and 26.7 vessels per group were reported as lacking them. Before continuing with the fit of the definitive model, we re-examined the preparations of these species, which revealed the presence of vasicentric tracheids in both. When we refitted the model with these species recoded as having vasicentric tracheids, a non-significant effect of vessel diameter on the response variable was reported (Wald statistic = −1.57, 1 d.f., $P = 0.116$), and this variable was therefore not included in the final model.

Final multivariate logistic model: The final model included vessels per group and helical sculpture in latewood as explanatory variables (Table 6), with updated coding for *Baccharis salicifolia* and *Fremontodendron californicum*. The Hosmer–Lemeshow test indicated that the model described the data well (Hosmer–Lemeshow statistic = 13.10, 8 d.f., $P = 0.11$), and just six of 207 species (3%) had a standardized residual greater than three in absolute value. The positive regression coefficients (Table 6) indicated that the

Table 6. Final multivariate logistic model predicting the presence or absence of vasicentric tracheids based on the number of vessels per group and the presence or absence of helical sculpture in latewood. Variables are listed in order of decreasing association with the response, according to the magnitude of their standardized coefficients

	Coefficient	SE	*P*-value	Stand coeff
Constant	−1.92	0.30	< 0.001	−0.49
Vessels per group	0.25	0.07	0.001	2.72
Helical sculpture in latewood	1.18	0.32	< 0.001	0.58

SE, standard error of the logistic regression coefficient; Stand coeff, standardized logistic regression coefficient, calculated as the original regression coefficient multiplied by the standard deviation of its corresponding explanatory variable.

probability of finding vasicentric tracheids is greater in woods with more vessels per group or with helical sculpture in latewood. Figure 3 presents the two curves showing how the estimated probability of finding vasicentric tracheids increases with the number of vessels per group. Species with helical sculpture in latewood, in the upper curve, have a higher probability of presenting vasicentric tracheids. The smaller the number of vessels per group, the greater the difference between the probabilities of finding vasicentric tracheids in species with and without helical sculpture in latewood.

The standardized regression coefficients (Table 6) implied that the number of vessels per group was the variable most strongly related to the presence of vasicentric tracheids. To give an idea of the magnitude of the effect of vessel grouping, consider a plant with helical sculpture in latewood. Based on the model, the probability of finding vasicentric tracheids is 62.3% if the plant has an average of five vessels per group and 38.05% if the average is one vessel per group (these cases are highlighted in Figure 3). Thus, an increase of four units in the average number of vessels per group produces an increase of 24.31% in the probability of finding vasicentric tracheids, and amounts to RR = 1.64.

True tracheids as the response variable
Initial steps in model construction: univariate analysis: Simple logistic regression models indicated that all the continuous characteristics of vessels were strongly associated with true tracheids (Table 5). They tended to be present in species with a large number of vessels mm^{-2} or with long vessel elements. On the contrary, the wider the vessels were or the greater the
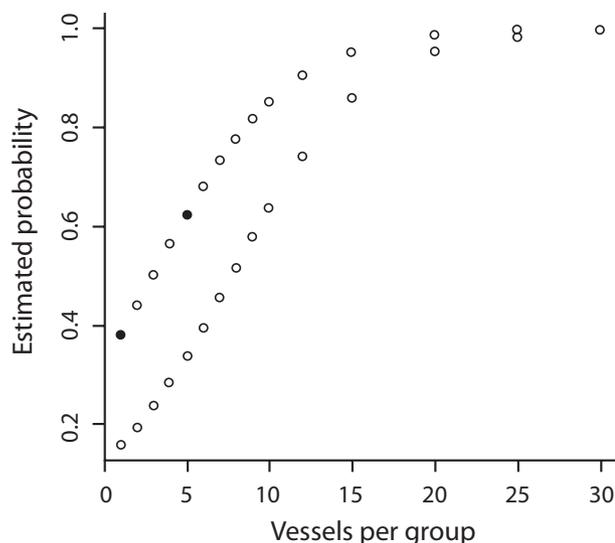
Figure 3. Estimated probability of the presence of vasicentric tracheids as vessel grouping is varied in the presence of helical sculpture in latewood (upper curve) and in the absence of helical sculpture in latewood (lower curve). The estimates are based on the final multivariate model predicting the presence or absence of vasicentric tracheids. Based on the model, the probability of finding vasicentric tracheids is 62.3% if a plant has an average of five vessels per group and 38.05% if the average is one vessel per group. These two cases are highlighted in the graph.

number of vessels per group, the less probable it became to find true tracheids.

The only categorical variable strongly associated with true tracheids was perforation plate type. Only 11 species in the data set had scalariform perforation plates, and all of them had true tracheids, except for *Alnus rhombifolia* and *Betula fontinalis*. Of the 196 remaining species that had simple plates, only 20.41% (40/196) had true tracheids. This clear association between the presence of scalariform perforation plates and the presence of true tracheids was confirmed by the Fisher exact test (*P* < 0.0001). The RR value indicated that the probability of finding true tracheids in species with scalariform perforation plates was four times that of species having simple plates (RR = 4.01).

Preliminary multivariate logistic model: A preliminary model was fitted to predict the presence of true tracheids using all the continuous vessel characters as explanatory variables. Vessel diameter was then excluded because, although it was associated with true tracheids, it was not significant in the presence of the other continuous variables (Wald statistic = −0.76, 1 d.f., *P* = 0.445). Perforation plate type was not included in any multivariate model because the low number of species with scalariform perforation plates

and lacking true tracheids did not enable us to estimate some regression coefficients precisely.

As with vasicentric tracheids, Pregibon's index was calculated to detect influential observations, with *Symphoricarpos mollis* having the largest index. When this species was removed from the data and the model fitted again, the estimated coefficient for vessels per group changed considerably, corroborating the influence of *Symphoricarpos mollis*, which had unusually high vessel grouping for a species with true tracheids. Upon the examination of preparations of *Symphoricarpos mollis*, we found that nonconductive axial cells formed the ground tissue in earlywood, along with grouped vessels, and that vessels were nearly solitary when embedded in a matrix of latewood true tracheids. Therefore, we considered only the grouped earlywood vessels, where tracheids are absent, and recoded this species as lacking true tracheids rather than having them before proceeding with the final version of the model.

Final multivariate logistic model: The final model included vessel density and grouping, and vessel element length as explanatory variables (Table 7). The model fitted the data quite well (Hosmer–Lemeshow statistic = 2.31, 8 d.f., $P = 0.97$). Only three species had a standardized residual whose absolute value ranged between 3 and 3.5. Vessel density and vessel element length were positively associated with the presence of true tracheids, whereas vessel grouping reduced the probability of finding true tracheids (Table 7). The standardized regression coefficients of these variables in the multiple model (Table 7) indicated that vessel grouping and vessel density were the two variables most strongly related to the presence of true tracheids. For illustrative purposes, consider a plant with 200 vessels mm^{-2} and an average vessel element length of 200 μm. If the average number of vessels per group was one, the probability of finding true tracheids would be 66.8%, whereas this probability would be nearly zero (0.01%) if there were three per group.

Thus, this difference of two vessels per group produces a drastic decrease in the probability of finding true tracheids.

Vascular tracheids as the response variable
Just 13% of the species in the data set had vascular tracheids. This percentage is underestimated because vascular and vasicentric tracheids were coded as mutually exclusive when they can actually be present in the same plant. Therefore, the results in this section should be interpreted with caution until they can be repeated with the correct coding.

Initial steps in model construction: univariate analysis: Univariate logistic models indicated that vessel diameter and vessel element length were associated with vascular tracheids, albeit weakly (Table 5). Neither vessel grouping nor vessel density were associated with the presence of vascular tracheids (Table 5). Likewise, none of the categorical variables was associated with vascular tracheids. *Betula fontinalis* was the only species with both vascular tracheids and scalariform perforation plates.

Preliminary multivariate logistic model: A multivariate logistic model with vessel diameter and vessel element length as explanatory variables was fitted to the data. *Sambucus mexicana* and *Betula fontinalis* were the two species with the largest influence on the model. These two species affected the estimated regression coefficient for vessel element length, because the 37 species with a vessel element length greater than 328 μm lacked vascular tracheids, except *Sambucus mexicana* and *Betula fontinalis* (vessel element length 438 and 426 μm, respectively).

Final multivariate logistic model: The final model included vessel element diameter and length as explanatory variables (Table 8), and fitted the data well (Hosmer–Lemeshow statistic = 10.41, 8 d.f., $P = 0.237$). Only five species had a standardized

Table 7. Final multivariate logistic model predicting the presence or absence of true tracheids based on vessels per group, vessel density, and vessel element length. Conventions as for Table 6

	Coefficient	SE	*P*-value	Stand coeff
Constant	1.71	1.320	0.195	−12.479
Vessels per group	−4.90	0.996	< 0.001	−53.763
Vessels mm^{-2}	0.011	0.002	< 0.001	2.387
Vessel element length	0.01	0.002	0.001	0.896

Table 8. Final multivariate logistic model predicting the presence or absence of vascular tracheids based on vessel element length and vessel diameter. Conventions as for Table 6

	Coefficient	SE	*P*-value	Stand coeff
Constant	−1.588	0.607	0.009	−2.015
Vessel element length	−0.005	0.002	0.055	−0.512
Vessel diameter	0.019	0.009	0.050	0.377

residual with an absolute value between 3 and 5. Species with wide vessels had a higher probability of presenting vascular tracheids, but, as indicated by the Wald test (Table 8), the association was weak. Similarly, a weak association was observed between vessel element length and the presence of vascular tracheids. The low proportion of species with vascular tracheids and the problems with coding may be responsible for the low power of the tests to detect association.

DISCUSSION

In many cases, our results were congruent with well-established concepts in ecological wood anatomy, such as the negative relationship between vessel density and vessel diameter (e.g. Baas, Werker & Fahn, 1983; Carlquist & Hoekman, 1985; Carlquist, 2001). Nevertheless, the present study represents the first time that the magnitude of many of these 'well-known' associations has ever been documented and assigned a *P*-value. However, rather than focus on widely recognized associations, in this section we treat less commonly discussed aspects of ecological wood anatomy highlighted by our work, and novel directions for further research that our analysis suggested. Among these are compelling results regarding the functional interplay between vessel grouping and imperforate tracheary element type, the detection of unusual species for further study, and the identification of suites of character combinations that may be of functional significance. We also discuss lessons pointed to by our analyses regarding the limits of standard comparative wood anatomical sampling and data handling, and ways that statistical analyses could be used to evaluate schemes for classifying wood features.

VESSEL GROUPING, DROUGHT RESISTANCE, AND IMPERFORATE TRACHEARY ELEMENT TYPE

In a keenly insightful application of the comparative method, Carlquist (1984) proposed that, if vessel grouping is driven by selective pressure favouring redundancy of conductive streams and consequent drought resistance, then it should be lessened in the presence of subsidiary conductive cells such as true tracheids or vasicentric tracheids. Building on this notion, he suggested that the conductive nature of imperforate tracheary elements could be diagnosed by vessel grouping, with lowered grouping indicating conductive imperforate elements (Carlquist, 1984, 2001). The strong pressure differential between an embolized vessel and an adjacent water-filled one appears to put the non-embolized vessel at risk (Sperry & Hacke, 2004), so there may in fact be selection for decreased vessel grouping in the presence of sufficient subsidiary conductive cells, rather than simply relaxed selection. Although this reasoning provides the anatomist with a comparative means of inferring the functional nature of the imperforate tracheary elements in a given taxon, it appears not to have been incorporated into functional comparative wood anatomical analyses generally. However, our findings strongly support Carlquist's interpretation.

True tracheids and vessel grouping

Vessel grouping showed a remarkable ability to predict true tracheid presence (Tables 5 and 7). Figure 4 shows that, as Carlquist (1984) noted, vessels are almost exclusively solitary when true tracheids are present, with no instances of substantial vessel grouping having been observed in the presence of true tracheids. Likewise, the differences in vessel grouping between woods with true tracheids and those without them suggested by Figure 4 were strongly statistically significant (Mann–Whitney $U = 757.5$, $P < 0.001$). That vessel grouping only occurs when true tracheids are absent implies that true tracheids are indeed conductive, and that they provide greater conductive safety than vessel grouping (cf. Carlquist, 2001).
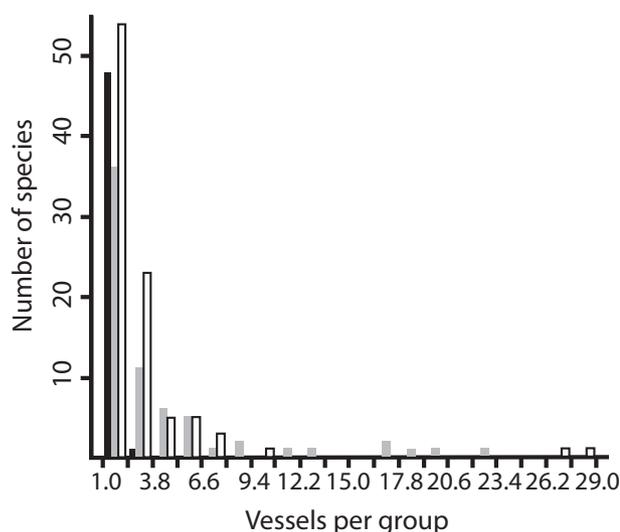
Figure 4. Difference in the distribution of vessel grouping in species with true tracheids (black bars), vasicentric tracheids (grey bars), and vascular tracheids or no conductive imperforate tracheary elements (white bars). There is a significant difference between vessel grouping in taxa with true tracheids and those with vasicentric tracheids ($P < 0.001$), and also between species with vasicentric tracheids and no conductive imperforate tracheary elements ($P < 0.001$).

Vasicentric tracheids and the functional significance of vessel grouping type

Vasicentric tracheids are thought to be subsidiary conductive cells of high safety. Just as true tracheid presence is associated with solitary vessels, lower vessel grouping could be predicted in a wood with vasicentric tracheids as compared with one with only nonconductive imperforate tracheary elements. In contrast to these expectations, our logistic regression model showed that vasicentric tracheid presence is actually *more* probable when vessels are grouped (Tables 5 and 6). Figure 4 illustrates this result graphically. If vasicentric tracheid presence lowered vessel grouping in a way analogous to that of true tracheids, then the distribution of vessel grouping in species with vasicentric tracheids should overlap that of species with true tracheids. To the contrary, the distribution of species with vasicentric tracheids appears distinct from that of species with true tracheids, a difference that is supported statistically (Mann–Whitney $U = 487.5$, $P < 0.001$).

Although some species with vasicentric tracheids do have very low vessel grouping (Table 9), Figure 4 shows that many species have remarkably large vessel groups. That this range of vessel grouping is observed in species with vasicentric tracheids suggests that plants with these cells employ an array of functional strategies, in contrast to the consistent lowering of vessel grouping observed when true tracheids are present (Fig. 4; Carlquist, 1984, 1985). Carlquist (1985, 2001) illustrated the radially extensive, often diagonal, aggregations of vessels that typify many species with vasicentric tracheids, and suggested that interconnection between these massive aggregations probably unites most of the vessels in a given stem into a single highly redundant conductive system. That large aggregations of vessels have evolved in a wide array of lineages, producing so strong a statistical association with vasicentric tracheids, suggests that they are one of the key strategies for conductive safety that arise in sclerophyllous plants of drylands. From a statistical point of view, questions that cannot be addressed in a given study might be accessible by examining additional variables. In the case of vessel grouping types and vasicentric tracheids, a study explicitly designed to identify association between qualitative and quantitative aspects of vessel grouping type, vasicentric tracheid distribution and abundance, and variables such as phenology and libriform fibre characteristics (cf. Jacobsen *et al*., 2005), could help to identify the selective pressures associated with different vessel grouping phenomena in the presence of vasicentric tracheids.

Vasicentric tracheids are consistently associated with high vessel grouping to the point where the logistic regression singled out *Baccharis salicifolia* and *Fremontodendron californicum* because of their high

Table 9. Species in the data set with vasicentric tracheids and low vessel grouping (fewer than three vessels per group) to show that a wide range of vessel grouping is observed in groups with vasicentric tracheids. This range is probably associated with a range of vasicentric tracheid abundance. Family names are given to show that many families are involved

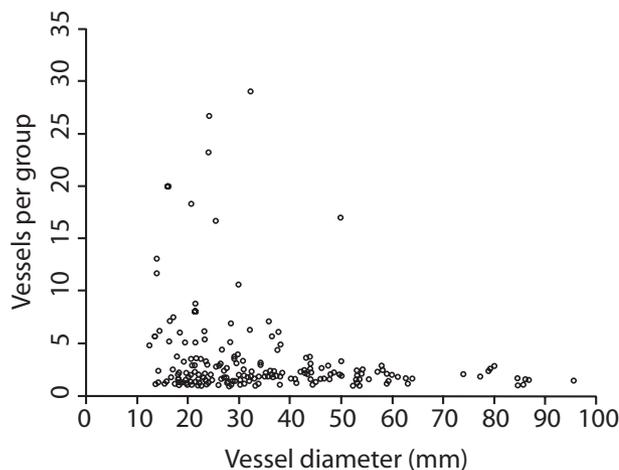| Species	Family	Vessels per group
Lithocarpus densiflora	Fagaceae	1
Quercus chrysolepis	Fagaceae	1
Quercus dumosa	Fagaceae	1
Larrea tridentata	Zygophyllaceae	1
Comarostaphylis diversifolia	Ericaceae	1.08
Fagonia laevis	Zygophyllaceae	1.08
Chrysolepis sempervirens	Fagaceae	1.12
Frankenia grandifolia	Frankeniaceae	1.12
Thamnosma montanum	Rutaceae	1.2
Arctostaphylos patula	Ericaceae	1.28
Trichostema lanatum	Lamiaceae	1.34
Condalia globosa	Rhamnaceae	1.36
Ornithostaphylos oppositifolia	Ericaceae	1.43
Holodiscus microphyllus	Rosaceae	1.44
Cassia armata	Leguminosae	1.45
Ribes cereum	Glossulariaceae	1.56
Ferocactus acanthodes	Cactaceae	1.6
Lippia wrightii	Verbenaceae	1.6
Arctostaphylos glauca	Ericaceae	1.64
Holodiscus discolor	Rosaceae	1.64
Arctostaphylos glandulosa	Ericaceae	1.65
Buddleja utahensis	Buddlejaceae	1.76
Echinocactus polycephalus	Cactaceae	1.8
Xylococcus bicolor	Ericaceae	1.8
Solanum xantii	Solanaceae	1.8
Celtis reticulata	Ulmaceae	1.8
Colubrina californica	Rhamnaceae	1.85
Beloperone californica	Acanthaceae	1.96
Salazaria mexicana	Lamiaceae	2
Mammillaria dioica	Cactaceae	2.01
Bernardia incana	Euphorbiaceae	2.07
Arbutus menziesii	Ericaceae	2.1
Monardella linoides	Lamiaceae	2.16
Haplopappus squarrosus	Asteraceae	2.17
Ceratoides lanata	Chenopodiaceae	2.24
Adolphia californica	Rhamnaceae	2.32
Rhamnus crocea	Rhamnaceae	2.55
Zizyphus parryi	Rhamnaceae	2.55
Gutierrezia microcephala	Asteraceae	2.7
Asclepias albicans	Apocynaceae	2.85
Echinocereus englemannii	Cactaceae	2.95
Rhamnus californica	Rhamnaceae	2.95

vessel grouping but lack of vasicentric tracheids. After *Romneya coulteri*, with its extreme value of 150 vessels per group, these two species had the highest values for vessel grouping. Above ten vessels per group, all species had vasicentric tracheids, until reaching values of 26.7 (*Fremontodendron californicum*) and 29 (*Baccharis salicifolia*). This result raised several questions. For example, did these two species represent a strategy otherwise unique in the data set that obviated the need for these subsidiary conductive cells? Alternatively, could vasicentric tracheids, which are often difficult to identify, actually have been present in these species but were not observed in the original study? A re-examination of the slides of *Baccharis salicifolia* and *Fremontodendron californicum* in preparation for the present paper revealed that in fact both of these species do have vasicentric tracheids. Thus, a simple statistical procedure helped guide, to a remarkable degree, our interpretation of these preparations, and led to a refined understanding of the association between vasicentric tracheids and vessel grouping.

Vessel grouping, vessel diameter, and vessel density
Many anatomists have noted that the number of vessels per unit of transection is clearly related to how wide the vessels are, which in turn affects how many could be expected to be found in groups (e.g. Sidiyasa & Baas, 1998; Carlquist, 2001). We present levels of statistical significance for these associations (Table 3). The significant negative correlation between vessel diameter and density (Spearman's correlation coefficient = −0.75, $P < 0.001$) is an intuitively obvious result, because fewer larger vessels can fit into a given space than small ones.

Likewise, the positive correlation between vessel density and the average number of vessels per group is to be expected, given that the more abundant vessels are, the more probable it is that they will come into contact with one another. In contrast to the correlation between vessel diameter and density, this correlation is in fact a rather weak one (Table 3), although it is significant at the 10% level and is thus worth calculating in other studies. Figure 5 shows the scatter of species based on vessel diameter and vessels per group and shows that most species have low vessel grouping but are distributed across a wide range of vessel diameter. The region of the graph in which diameter is wide and grouping is low is predicted to be associated with areas of low water stress, and species in this part of the graph comply with this expectation, or have wood with an internally mesic environment (e.g. *Cornus nuttallii* from shady canyons or the succulent tree *Bursera microphylla*).

Understanding the functional significance of vessel grouping could be refined by generating predictions

Figure 5. Scatterplot of the species in the data set according to their values of vessel diameter and vessel grouping. Species tend to have few vessels per group, but can vary their vessel diameter greatly.

regarding how many vessels should be found in groups if the probability of vessels being in contact in a given sample is random. Given the range of diameters and density of vessels in a given species, a certain number of vessels will be in contact if their distribution is determined at random. Vessel grouping that is lower than this figure for a given species would suggest selection for solitary vessels, whereas higher grouping could suggest selection favouring grouping of vessels.

IDENTIFICATION OF UNUSUAL SPECIES

One of the most fruitful results of the models we present was that they singled out certain species as being in some way unusual. 'Unusual' in this case refers to species with combinations of characteristics that differ from the rest of the species. One possibility for such a result is that subtleties in the original anatomical preparations remained undiscovered. As was the case for *Baccharis salicifolia* and *Fremontodendron californicum*, the models drew attention to readily overlooked details that might never have been noticed otherwise, leading to a refined understanding of the anatomy of these species. We detail an additional case here.

An example of a plant whose anatomical interpretation was refined as a result of our models is *Symphoricarpos mollis*, a finely branched evergreen caprifoliaceous understory shrub. The imperforate tracheary elements of this species were originally classed as true tracheids. However, vessels in this species were coded as grouped rather than having the solitary vessels that would be expected if true tracheids were present. Examination of the data showed that,

with the exception of *Symphoricarpos mollis*, true tracheids were only found in species with less than 2.36 vessels per group; all 75 species with more than 2.36 vessels per group lacked true tracheids except for *Symphoricarpos mollis*, which had true tracheids and 3.25 vessels per group. This raised several questions: Are the 'tracheids' of *Symphoricarpos mollis* truly conductive? What implications would this observation have for the notion that true tracheid presence should lower vessel grouping (Carlquist, 1984)? What features of the imperforate tracheary elements made them appear conductive if they were not? Resolution of this issue required the consultation of new preparations.

A re-examination of sections of *Symphoricarpos mollis* prepared from liquid-preserved material showed that true tracheids are indeed present, as are grouped vessels, but in distinct portions of growth rings. Earlywood vessels – but only for a short period of a season's growth – are grouped. They occur in a background of what could either be called living fibres (filled with starch in this case) or axial parenchyma that has not subdivided into strands. Using either designation, they are nonconductive cells, and thereby vessel grouping is an adaptive strategy for increasing conductive safety in earlywood. Most of each growth ring consists of latewood in which vessels are solitary, and the background tissue consists of true tracheids; the starch-filled living fibres (or fibriform parenchyma cells) are sparsely distributed among the tracheids. Thus, vessel grouping is not an adaptive strategy for increasing conductive safety in latewood. The example of *Symphoricarpos mollis* demonstrates the usefulness of the logistic models in highlighting taxa that require special consideration. In this case, vessel grouping and imperforate tracheary element type could be considered in early- and latewood separately.

UNUSUAL CHARACTER COMBINATIONS AND LACK OF CORRELATIONS

One of the keystones of the use of the comparative method for inferring the functional roles of wood anatomical features has been the detection of associations or the lack thereof between anatomical characters. In this approach, the functional nature of cells is inferred via examination of the association between characters of many species and not on the functional study of one or a few species. An approach such as the one we present is ideal for revealing such patterns of association. A special case of association would be represented by combinations of characters that are unusual or would be if they existed. *Philadelphus microphyllus* provides an example of a plant with such a combination, being the only species with true tracheids and helical vessel sculpture. Outside of the southern

California flora area, there is a small number of other genera in which true tracheids co-occur with vessels with helical thickening, such as *Buxus* (Carlquist, 1982) and *Ilex* (Baas, 1973). Both of these genera also have helical thickenings in tracheids (IAWA Committee, 1989). *Buxus*, *Ilex*, and *Philadelphus* may represent predominantly mesophytic clades in which a few genera (which have scalariform perforation plates as a symplesiomorphy) have been able to adapt to seasonal cold or drought (neither of a severe sort) with the aid of a few modifications, such as helical thickenings. Further reinforcing the idea that helical sculpture is somehow of value in drought survival, if helical thickenings are present in the vessels of woods from seasonal climates, they are always more pronounced in latewood, as in *Philadelphus*.

The absence of correlations is equally informative. For example, the lack of correlation between vessel element length and vessel diameter (Table 3) suggests that these characteristics can respond independently to selection. This was previously shown by Carlquist (1966) in a summary of woods of Asteraceae. Woods of Asteraceae are relatively uniform with respect to qualitative differences. Therefore, the family is like a template that reveals sensitively adaptive changes in quantitative vessel features to shifts in ecology. When subdivided into ecological categories ('mesic, dry, desert'), Asteraceae show a marked decrease in vessel diameter and vessel element length with increased xeromorphy accompanied by more numerous vessels per group. The absence of correlation among these three features strongly suggests that morphogenetically they are not related. Vessel element length is traditionally thought to represent about the same length as fusiform cambial initials in any given specimen, whereas vessel diameter is the result of radial and tangential cell enlargement during maturation. Finally, the number of vessels per group is a function of the number of vessels derived from the cambium in series or simultaneously, uninterrupted by other kinds of axial cell. The identification of which characters are correlated is a vital first step in understanding the types of evolutionary change that are possible (cf. Olson & Rosell, 2006).

STATISTICS AND WOOD ANATOMICAL TERMINOLOGY

All systems of classification to some extent gloss over biological realities. Although it is not our aim to advocate one set of wood anatomical terminology over another, our study does offer an example of how explicit statistical approaches can help to indicate terminological systems that are more closely aligned with biological realities than others. The terminology of imperforate tracheary elements has long been a subject of debate, much of which has consisted of

expressions of personal taste on the part of anatomists and little effort to identify biological discontinuities along which terminological lines could be drawn. For example, the distinction between vasicentric and vascular tracheids used here is not universally accepted by anatomists (e.g. IAWA Committee, 1989). If the distinction we used were truly arbitary, then it would be unlikely that a statistical model could predict effectively the presence of either of these cell types. However, vasicentric tracheid presence was readily predicted by the model based on other variables (Table 6). If definitions are taken as hypotheses, then our results strongly support the hypothesis of vasicentric tracheids as conceived by Carlquist (1984, 1985, 2001).

The cell types conceived here thus do appear to have distinct biological functions, and their distinctness has been inferred by previous authors. Metcalfe & Chalk's (1950) list of families that have vasicentric tracheids was incomplete, although the concept of vasicentric tracheids in that work seems entirely consistent with our usage. An intensive search for vasicentric tracheid presence (Carlquist, 1985) greatly expanded the list of Metcalfe & Chalk (1950), but did not change the concept. The concept was changed in the IAWA Committee's (1989) listings. The IAWA Committee (1989) recognized vasicentric tracheids only in two families where they are abundant, Fagaceae and Myrtaceae. Although restriction of the concept of vasicentric tracheid to those two families may simplify identification work, it does not recognize the functional role evident in the earlier definition of the term as used by Metcalfe & Chalk (1950) and Carlquist (1985). Furthermore, by assigning the term vascular tracheid to imperforate tracheary elements with numerous bordered pits (in woods that also contain fibre-tracheids or libriform fibres), the IAWA Committee (1989) has blurred the distinction between vasicentric tracheids and vascular tracheids. The term vascular tracheid, as earlier used, applied only to last-formed latewood tracheary elements that lack perforation plates, as in *Sambucus* or *Betula*. In the traditional definition of vascular tracheids, by being embolism-resistant cells adjacent to the cambium when growth ceases, these cells protect the cambium but not the three-dimensional conductive pathways of a stem. Thus, vascular tracheids aid in protecting a stem (but not its leaves) in cold or dry periods when vessels in a stem may have embolized. Vascular tracheids are thus an adaptive mechanism in plants with drought-deciduous leaves (*Aesculus*, *Toxicodendron*) or winter-deciduous leaves (*Betula*). Vasicentric tracheids and true tracheids, by contrast, protect conductive pathways by remaining embolism free when nearby vessels embolize, and characterize the evergreen chaparral elements, such as *Adenostoma*, *Arctostaphylos*, *Ceanothus*, etc. For the purposes of ecological wood anatomy, terminology reflecting cell function clearly influences the inferences that can be derived from a given data set.

The methods we suggest provide a way of comparing and choosing between classification systems. The 'true tracheids' of the present study are referred to as 'fibres with distinctly bordered pits' by many anatomists, following the wood identification guidelines of the IAWA Committee (1989), which does not distinguish based on cell function and thus includes both conductive and nonconductive cells in this category. However, the present study does suggest a very marked boundary between conductive tracheids and nonconductive imperforate tracheary elements, supporting the validity of the 'tracheid' category as conceived here. It would thus be possible to see how well the presence of 'fibres with distinctly bordered pits' is predicted by other variables in a data set constructed using the IAWA Committee's (1989) classification vs. a data set based on the same samples but classified using the definitions employed here. The classification system that results in better predictive power would be preferred as reflecting biological reality better than the other.

We must emphasize that our primary interest is not the classification of cell types or the invention of terminology. Indeed, this effort has to some extent hindered the progress of evolutionary comparative wood anatomy, because debate has at times focused more on what name to use to lump together superficially similar cells across dicots rather than the inference of their function or their status as homologies associated with individual clades. For studies of ecological wood anatomy, collecting quantitative data rather than classing cells should be considered, because the use of classes, e.g. 'vessels large, vessels small', etc., inevitably leads to a loss of statistical power vs. what can be achieved analytically with continuous data (e.g. Royston *et al.*, 2006). Moreover, these classes are often arbitrary divisions of continua, and all inferences based on them are subject to error to the extent that the classes are arbitrary (see Olson, 2005). For studies of ecological wood anatomy, effort is much better directed to gathering data that can be used to test a specific functional hypothesis than to gathering all of the traditional wood identification data typically collected by anatomists (e.g. IAWA Committee, 1989). Although terminology is not an end in itself, from the standpoint of ecological wood anatomy, terminology that reflects function is at the very least convenient. At best, and as our analysis shows, when classes apparently coincide with biological reality, they can be used to build models with excellent predictive ability, and consequently can be a rich source of biological information.

Every data set has its limits as a natural consequence of sampling. Identifying the limits imposed by the gathering and treatment of data in a given study can serve to widen the envelope of informativeness of future investigations. In this section we discuss specific examples of analyses that remained inaccessible due to such limitations. By identifying currently unanswerable questions, our models provide explicit directions for future work.

The coding of variables directly determines the inferences that can be drawn. In some cases, simplifying these codings can be a detriment to the conclusions that can be made. For example, Carlquist & Hoekman (1985) attempted to simplify the interpretation of the data by coding a single type of conductive imperforate tracheary cell per species. The species in which both vasicentric and vascular tracheids were found were coded as having exclusively vasicentric tracheids. Possibly as a result, the logistic regression model was unable to predict the presence or absence of vascular tracheids dependably based on the other variables (Table 8). Coding the presence of both vasicentric and vascular tracheids when they co-occur should increase the explanatory and predictive abilities of the model.

Anatomists often break continua into categories, thereby losing crucial information. One example of the use of categories instead of continuous variables includes recording presence/absence rather than abundance data. In the current study, we found that the probability of finding vasicentric tracheids increased with increasing numbers of vessels per group (Tables 5 and 6; Fig. 3). This finding is significant because the presence of subsidiary conductive cells could be expected to lower, rather than raise, vessel grouping (Carlquist, 1984). Carlquist & Hoekman (1985) mentioned that vessel grouping is lowered only in the presence of very abundant vasicentric tracheids. This would clearly appear to be the case in families such as Fagaceae and Myrtaceae, in which very abundant vasicentric tracheids are found in many species (Metcalfe & Chalk, 1950). Consistent with expectations, vessels are solitary in species that appear to have abundant vasicentric tracheids (Table 9). At the other extreme, vasicentric tracheids are relatively sparse in other taxa, and a range of vessel grouping is observed. Between these extremes, a threshold of vasicentric tracheid abundance is probably crossed, below which selection for vessel grouping becomes a significant force. To identify this threshold, abundance data for vasicentric tracheids would be required. However, not only is vasicentric tracheid abundance not recorded in the Carlquist and Hoekman data set, but it is doubtful that this information has ever been gathered. Yet without abundance data it is impossible to test the hypothesis that vessel grouping is lowered in the presence of abundant vasicentric tracheids quantitatively and to describe the nature and magnitude of the association.

Similarly, rather than categorizing imperforate tracheary elements (tracheid, fibre-tracheid, etc.), a better test of Carlquist's (1984) hypotheses would be to compare vessel grouping with quantitative features of imperforate tracheary elements, such as pit dimensions and density. Such an approach could help to identify the threshold range of continuous features of imperforate tracheary elements over which the change from solitary to grouped vessels occurs. The identification of a quantitative threshold of pit size and density below which imperforate tracheary elements are apparently nonconductive would be an invaluable aid to inferring the adaptive responses of woody plants to environmental challenges.

To make such inferences valid, analyses must simultaneously take into account biological and statistical considerations. It should go without saying that statistical models should never be used to make inferences without checking whether they adequately fit the data, yet such diagnostics are seldom performed in practice. However, some of the most important results of our study, such as the identification of unusual species, emerged in exactly this phase. Likewise, it is crucial to heed the need for a sufficient number of observations to obtain reliable estimates for the standard deviations of the estimated regression coefficients. Thus, although worthwhile, simply obtaining a large overall sample size does not guarantee that standard deviations can be precisely estimated. In our case, it was also necessary to examine how many species with and without each tracheid type were found for different values of the explanatory variables. For example, all species in the Carlquist and Hoekman data set with scalariform perforation plates lacked vasicentric tracheids. Thus, the presence of scalariform perforation plates perfectly predicted the absence of vasicentric tracheids. This would have resulted in a zero observed frequency in the contingency table defined by the two variables, yielding estimates of regression coefficients with standard deviations so large as to be practically meaningless. Similar data patterns are common in the Carlquist and Hoekman data (e.g. the presence of any tracheid type is generally associated with the absence of the other two; the presence of helical sculpture in earlywood implies the presence of helical sculpture in latewood, etc.). Such patterns may be expected in any ecological wood anatomy study and it is therefore crucial that their effect on the analysis be taken into account.

Finally, a large number of strong ecological wood anatomical correlations, some of which form the foundations of the current ecological wood anatomy paradigm, are almost certainly indirect, correlating with plant size rather than directly with environmental variables. The positive correlations between vessel diameter and water availability (e.g. Baas *et al.*, 1983) and between tracheary element length and latitude (Van der Graaff & Baas, 1974; Lens *et al.*, 2004) are probably in this category. Nevertheless, we are not aware of anatomical studies in which vessel diameter is corrected for stem size before seeking ecological correlations. Results that may be similarly indirect were also recovered in the present analysis. For example, vessel diameter tends to be smaller when vasicentric tracheids are present (Table 5). This finding is consistent with the common ecological wood anatomy assumptions that both narrow vessels and vasicentric tracheids are associated with dry habitats, and thus both could be interpreted as adaptations to drought. However, RR values show that vasicentric tracheids are more common in shrubs and subshrubs than in trees or other habit types. Likewise, a narrower vessel diameter is found in smaller stems (Gartner, 1995; Carlquist & Grant, 2005; Olson & Rosell, 2006). It is thus unclear whether both narrow vessels and vasicentric tracheids are directly associated with drought or indirectly via an association with plant size, which is in turn associated with dry conditions. Does natural selection act on vessel dimensions independently of plant size, or does selection act on plant size, with vessel size changing accordingly (or vice versa)? If smaller habits tend to have narrower vessels *and* have vasicentric tracheids, the association between vessel diameter and vasicentric tracheid presence could be an indirect one. Also, it is not clear why vasicentric tracheids are so rare in trees. Yet another example is the association between the presence of growth rings and the presence of vasicentric tracheids. RR values highlight that the association is most important in ring porous woods, and is also strong, although less so, in semiring porous woods. Are vasicentric tracheids related directly to growth rings, or do both represent independent responses to drought resistance? Many more examples could be cited, clearly seeming to belie the notion that these ecological wood anatomical associations are 'well known'. In fact, much work remains to disentangle the chain of causality leading to most such associations.

CONCLUSIONS

Comparative wood anatomy has succeeded in identifying patterns, despite employing practices that date from the founding of the field. Many of these results have been repeatedly confirmed, which may make it seem that there is no reason to change these traditional methods. However, although the application of basic statistics has allowed the identification of strong patterns, more subtle associations and the identification of causality have clearly remained undiscovered. Also, it is possible that outright incorrect use of statistics (e.g. applying analyses of variance without checking assumptions) has led to erroneous conclusions. Very simple measures would substantially improve statistical practice in comparative wood anatomy. Specific recommendations stemming from our analysis naturally include the use of analytical techniques that take advantage of the considerable amounts of data locked away in comparative anatomical data sets, including the analysis of categorical data. Our analysis also points to situations in which gathering quantitative data (e.g. abundance of vasicentric tracheids, pit size, and density) would be preferable to traditional categorical data (e.g. the presence or absence of vasicentric tracheids, tracheids vs. fibretracheids). Gathering data regarding the abundance of cell types would likewise be extremely useful and, avoiding arbitrary or overly simplistic codings, is imperative.

Many associations between anatomical characters, and between characters and environmental variables, are restricted to certain taxonomic groups. For example, in this study, all instances of trees with vasicentric tracheids come from Fagaceae. Likewise, both species with scalariform perforation plates but lacking true tracheids are Betulaceae. A final example is the presence of helical sculpture on vessel walls in plants with diffuse porous wood, with the only two examples coming from Rosaceae. These within-family resemblances almost certainly do not represent independent occurrences, but instead characters inherited from a common ancestor. This tendency for closely related species to be more similar to one another than to distantly related ones is present in all comparative data. If the observed association arose only once in the evolutionary history of the group, all species should not be considered independent observations when describing the strength of the association in statistical terms. Nevertheless, although some anatomists may mention the apparent effect of phylogeny, all ecological wood anatomy studies to date have employed standard, nonphylogenetic statistical analyses. This practice is equivalent to considering data to be independent and translates into the implicit assumption that wood characters respond evolutionarily instantaneously to climatic conditions.

Statistical methods that deal with the non-independent nature of comparative data are available (e.g. Felsenstein, 1985; Harvey & Pagel, 1991; Martins & Hansen, 1997; Paradis & Claude, 2002), but have been applied in only a few anatomical studies (e.g.

Olson & Rosell, 2006; Preston, Cornwell & DeNoyer, 2006). Reanalysing the Carlquist and Hoekman data set with these methods would test whether the association patterns detected with the standard logistic models in this paper are maintained and would inform the degree of phylogenetic effect that is present in anatomical characters. It would thus provide a valuable guide as to just how much or how little neglecting of phylogeny distorts the inferences drawn by wood anatomists.

Carefully planned sampling and appropriate comparative methods complement functional studies because they can examine a greater number of species and ecological situations than could ever be studied using experimental approaches (e.g. Osborn, 1915; Bailey, 1920; Mayr, 1982). The analysis we present demonstrates that new information does remain even within a supposedly well-known data set. Furthermore, carrying out the suggestions for further work pointed to by this analysis would greatly refine the knowledge of wood cell function across families. We suggest that the future of comparative ecological wood anatomy would appear to depend on the eagerness of anatomists to abandon traditional methods that lead to a loss of analytical possibilities by refining sampling strategies and adopting more powerful analytical approaches designed to test specific functional hypotheses.

ACKNOWLEDGEMENTS

REFERENCES

Agresti A. 2002. *Categorical data analysis*, 2nd edn. New York: John Wiley & Sons.

Alves ES, Angyalossy-Alfonso V. 2000. Ecological trends in the wood anatomy of some Brazilian species. 1. Growth rings and vessels. *International Association of Wood Anatomists Journal* **21:** 3–30.

Alves ES, Angyalossy-Alfonso V. 2002. Ecological trends in the wood anatomy of some Brazilian species. 2. Axial parenchyma, rays and fibres. *International Association of Wood Anatomists Journal* **23:** 391–418.

Arias S, Terrazas T. 2001. Variación en la anatomía de la madera de *Pachycereus pecten-aboriginum* (Cactaceae). *Anales del Instituto de Biología, UNAM, Serie Botánica* **72:** 157–169.

Baas P. 1973. The wood anatomical range in *Ilex* (Aquifoliaceae) and its ecological and phylogenetic significance. *Blumea* **21:** 193–258.

Baas P. 1986. Terminology of imperforate tracheary elements—in defense of libriform fibres with minutely bordered pits. *International Association of Wood Anatomists Bulletin* **7:** 82–86.

Baas P, Werker E, Fahn A. 1983. Some ecological trends in vessel characters. *International Association of Wood Anatomists Bulletin* **4:** 141–159.

Bailey IW. 1920. The cambium and its derivative tissues II. Size variations of cambial initials in gymnosperms and angiosperms. *American Journal of Botany* **7:** 355–367.

Bailey IW. 1936. The problem of differentiation and classifying tracheids, fibre-tracheids, and libriform wood fibres. *Tropical Woods* **45:** 18–23.

Burley J, Miller RB. 1982. The application of statistics and computing in wood anatomy. In: Baas P, ed. *New perspectives in wood anatomy*. The Hague: Martinus-Nijhoff, 223–242.

Carlquist S. 1966. Wood anatomy of Compositae: a summary, with comments on factors controlling wood evolution. *Aliso* **6:** 25–44.

Carlquist S. 1975. *Ecological strategies of xylem evolution*. Berkeley, California: University of California Press.

Carlquist S. 1982. Wood anatomy of Buxaceae: correlations with ecology and phylogeny. *Flora* **172:** 463–491.

Carlquist S. 1984. Vessel grouping in dicotyledon wood: significance and relationship to imperforate tracheary elements. *Aliso* **10:** 505–525.

Carlquist S. 1985. Vasicentric tracheids as a drought survival mechanism in the woody flora of southern California and similar regions: review of vasicentric tracheids. *Aliso* **11:** 37–68.

Carlquist S. 1986. Terminology of imperforate tracheary elements. *International Association of Wood Anatomists Bulletin* **7:** 75–81.

Carlquist S. 2001. *Comparative wood anatomy. Systematic, ecological and evolutionary aspects of dicotyledon wood*, 2nd edn. Berlin: Springer.

Carlquist S, Grant JR. 2005. Wood anatomy of Gentianaceae, tribe Helieae, in relation to ecology, habit, systematics, and sample diameter. *Brittonia* **57:** 276–291.

Carlquist S, Hoekman DA. 1985. Ecological wood anatomy of the woody southern Californian flora. *International Association of Wood Anatomists Bulletin* **6:** 319–347.

Collett D. 2003. *Modelling binary data*. Boca Raton, Florida: Chapman & Hall.

Feild TS, Brodribb T, Holbrook NM. 2002. Hardly a relict: freezing and the evolution of vessels wood in Winteraceae. *Evolution* **56:** 464–478.

Felsenstein J. 1985. Phylogenies and the comparative method. *American Naturalist* **125:** 1–15.

Gartner BL. 1995. Patterns of xylem variation within a tree and their hydraulic and mechanical consequences. In: Gartner BL, ed. *Plant stems: physiology and functional morphology*. San Diego, California: Academic Press, 125–149.

Harvey PH, Pagel MD. 1991. *The comparative method in evolutionary biology*. Oxford: Oxford University Press.

Hosmer DW, Lemeshow S. 2000. *Applied logistic regression*, 2nd edn. New York: John Wiley & Sons.

IAWA Committee. 1989. IAWA list of microscopic features for hardwood identification. *International Association of Wood Anatomists Bulletin* **10:** 219–332.

Jacobsen AL, Ewers FW, Pratt RB, Paddock WA, III, Davis SD. 2005. Do xylem fibres affect vessel cavitation resistance? *Plant Physiology* **139:** 546–556.

Jansen S, Baas P, Gasson P, Lens F, Smets E. 2004. Variation in xylem structure from tropics to tundra: evidence from vestured pits. *Proceedings of the National Academy of Sciences of the USA* **101:** 8833–8837.

Kleinbaum DG. 2002. *Logistic regression: a self-learning text*. New York: Springer.

Kutner MH, Nachtsheim CJ, Neter J, Li W. 2005. *Applied linear statistical models*, 5th edn. New York: McGraw-Hill.

Lens F, Luteyn JL, Smets E, Jansen S. 2004. Ecological trends in the woody anatomy of Vaccinioideae (Ericaceae s.1.). *Flora* **199:** 309–319.

Martins EP, Hansen T. 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *American Naturalist* **149:** 646–667.

Mayr E. 1982. *The growth of biology thought*. Harvard, Massachusetts: Harvard University Press.

Metcalfe CR, Chalk L. 1950. *Anatomy of the dicotyledons*. Oxford: Clarendon.

Olson ME. 2005. Homology, typology, and homoplasy in comparative wood anatomy. *International Association of Wood Anatomists Journal* **26:** 507–522.

Olson ME, Rosell JA. 2006. Using heterochrony to infer modularity in the evolution of stem diversity in *Moringa* (Moringaceae). *Evolution* **60:** 724–734.

Osborn HF. 1915. Origin of single characters as observed in fossil and living animals and plants. *American Naturalist* **49:** 193–239.

Paradis E, Claude J. 2002. Analysis of comparative data using generalized estimating equations. *Journal of Theoretical Biology* **218:** 175–185.

Pregibon D. 1981. Logistic regression diagnostics. *Annals of Statistics* **9:** 705–724.

Preston KA, Cornwell WK, DeNoyer JR. 2006. Wood density and vessel traits as distinct correlates of ecological strategy in 51 California Coast Range angiosperms. *New Phytologist* **170:** 807–818.

Quinn GP, Keough MJ. 2002. *Experimental design and data analysis for biologists*. Cambridge: Cambridge University Press.

Royston P, Altman DG, Sauerbrei W. 2006. Dichotomizing continuous predictors in multiple regression: a bad idea. *Statistics in Medicine* **25:** 127–141.

Sidiyasa K, Baas P. 1998. Ecological and systematic wood anatomy of *Alstonia* (Apocynaceae). *International Association of Wood Anatomists Journal* **19:** 207–229.

Sokal RS, Rohlf FJ. 1995. *Biometry. The principles and practice of statistics in biology research*, 3rd edn. New York: Freeman.

Sperry JS, Hacke UG. 2004. Analysis of circular bordered pit function I. Angiosperm vessels with homogenous pit membranes. *American Journal of Botany* **91:** 369–385.

Van der Graaff NA, Baas P. 1974. Wood anatomical variation in relation to latitude and altitude. *Blumea* **22:** 101–121.