American Journal of Botany

American Journal of Botany 97(8): 1247–1259. 2010.

A QUANTITATIVE AND STATISTICALLY ROBUST METHOD FOR THE DETERMINATION OF XYLEM CONDUIT SPATIAL DISTRIBUTION[1]

MAURIZIO MENCUCCINI[2,3,6], JORDI MARTINEZ-VILALTA[3], JOSEP PIÑOL[3], LASSE LOEPFE[3], MIREIA BURNAT[3], XAVIER ALVAREZ[4], JUAN CAMACHO[4], AND DEBORA GIL[5]

[2]University of Edinburgh, School of GeoSciences, Crew Building, West Mains Road, EH9 3JN Edinburgh, UK; [3]Centre for Ecology Research and Forestry Applications (CREAF), Universitat Autonoma de Barcelona, Bellaterra, Cerdanyola, Barcelona, Spain; [4]Department of Physics, Universitat Autonoma de Barcelona, Bellaterra, Cerdanyola, Barcelona, Spain; and [5]Centre for Computer Vision, Universitat Autonoma de Barcelona, Bellaterra, Cerdanyola (Barcelona, Spain)

- *Premise of the study*: Because of their limited length, xylem conduits need to connect to each other to maintain water transport from roots to leaves. Conduit spatial distribution in a cross section plays an important role in aiding this connectivity. While indices of conduit spatial distribution already exist, they are not well defined statistically.
- *Methods*: We used point pattern analysis to derive new spatial indices. One hundred and five cross-sectional images from different species were transformed into binary images. The resulting point patterns, based on the locations of the conduit centers-of-area, were analyzed to determine whether they departed from randomness. Conduit distribution was then modeled using a spatially explicit stochastic model.
- *Key results*: The presence of conduit randomness, uniformity, or aggregation depended on the spatial scale of the analysis. The large majority of the images showed patterns significantly different from randomness at least at one spatial scale. A strong phylogenetic signal was detected in the spatial variables.
- *Conclusions*: Conduit spatial arrangement has been largely conserved during evolution, especially at small spatial scales. Species in which conduits were aggregated in clusters had a lower conduit density compared to those with uniform distribution. Statistically sound spatial indices must be employed as an aid in the characterization of distributional patterns across species and in models of xylem water transport. Point pattern analysis is a very useful tool in identifying spatial patterns.

Key words: Geyer; hydraulic conductivity; point pattern analysis; Ripley; Spatstat; vessel clusters; xylem anatomy; xylem network.

Water transport is commonly recognized as one of the most important functions performed by wood. It is now clear that plant hydraulics is heavily dependent upon wood anatomical properties such as conduit size and density or, at a smaller scale, size and characteristics of interconduit pits and pit membranes (e.g., Zwieniecki et al., 2001; Choat et al., 2006, 2008; Christman et al., 2009; Jansen et al., 2009). However, for interconduit pits to exist in the first place, conduits must be positioned in such a way that contact among them is possible. Hence, it may be hypothesised that tissue-level properties, such as the spatial distribution of conduits in a cross section or their vertical orientation and angle, will affect how frequently and extensively vessel walls come into contact with each other.

Plant anatomists have long highlighted that conduit grouping within a xylem growth ring varies significantly across species groups and habitats and have suggested that these spatial patterns may have an adaptive function (e.g., Carlquist, 1984, 2001). A similar argument may be made for the presence of isolated conduits, dendritic patterns, radial multiples, and diagonal or tangential files of conduits, although less is known about them.

The occurrence of these grouping patterns is frequently associated with the presence of other anatomical elements whose function is poorly understood. For example, one study based on the InsideWood database showed that species with vessels grouped in clusters and those with xylem arrangement based on solitary vessels appear to possess vascular and/or vasicentric tracheids (Wheeler et al., 2007). Similarly, a reanalysis of the wood anatomy of the southern Californian flora (Carlquist and Hoekman, 1985), using logistic regression, showed that solitary vessels were regularly associated with true tracheids, whereas vasicentric tracheids were more abundant with vessels grouped in clusters (Rosell et al., 2007). Although these two studies adopted different definitions for the various tracheid types, they concur in identifying the importance of conduit spatial arrangement in affecting the abundance of other elements (Carlquist, 1984).

Other lines of investigation have highlighted the importance of water transport sectoriality (Zanne et al., 2006; Schenk et al., 2008) and showed that this syndrome occurs more frequently under particular environmental conditions. Loepfe et al. (2007) used graph theory to model xylem properties as a network of interconnected elements. They employed their model to assess the significance of conduit connectivity on

[1] Manuscript received 22 September 2009; revision accepted 11 June 2010.

The authors thank the Catalan government (AGAUR-Generalitat de Catalunya) for a visiting Professorship in UAB to M.M. This study was supported by the Spanish Ministry of Education and Science via competitive project CGL2007-28784-E to J.M.-V. D.G. has been supported by The Ramon y Cajal Program. Thanks are due to Frederic Lens, Anna Jacobsen, and two anonymous reviewers for reading a draft of the manuscript and suggesting useful improvements, to D. Sol for help with the R software for the analysis of phylogeny, and to A. Baddeley and R. Turner for clarifications regarding Spatstat and for suggestions on estimating confidence intervals for spatial model parameters.

[6] Author for correspondence (e-mail: m.mencuccini@ed.ac.uk)

doi:10.3732/ajb.0900289

both xylem conductivity and P50, the pressure at which 50% of the conductive capacity is lost because of xylem embolism (cf. Tyree et al., 1994).

Despite this importance, the spatial distribution of xylem conduits is a poorly defined anatomical property, and no statistical tools are available to quantify it objectively and accurately. For example, Carlquist (1984) and Carlquist and Hoekman (1985) quantified grouping by counting the mean number of vessels physically touching each other in a section. The International Association of Wood Anatomists (IAWA) Committee on the use of Microscopic Features for Hardwood Identification (IAWA Committee, 1989, p. 219) adopted the definition that a vessel is "exclusively solitary when 90% or more of the vessels are completely surrounded by other (nonvessel) elements" and that "clusters" include "groups of three or more vessels having both radial and tangential contact and of common occurrence".

While these are useful indices for anatomical investigation, there are good reasons for exploring new ones. First, the enormous progress made in automatic image analysis and statistical spatial analysis now allows a quantitative approach to determine wood spatial features, employing sophisticated software. Second, any one of the indices mentioned only quantifies one component of the conduit spatial distribution (i.e., either grouping or vessel isolation), whereas arguably they all represent components of the same continuum, which spans from conduit aggregation (in its various forms) to spatial randomness and to conduit uniformity. This continuum should and can in fact be translated into quantitative, not simply categorical, variables to enhance the power of the statistical analyses. Third, spatial aggregation and uniformity are dependent on the spatial scale over which the analysis is carried out. The existing indices focus entirely on the smallest possible spatial scale, the one at which conduits either touch each other or do not. However, patterns occurring at larger scales (for example, dendritic groupings with large in-between areas with very low conduit densities or patterns caused by the presence of parenchyma rays effectively constraining the conduits into regular radial strips) may also carry biological significance. A statistical approach capable of identifying spatial "proximity" at different spatial scales may be better suited than one based on conduit physical contact as an input to three-dimensional hydraulic models, because neighboring conduits are more likely to be connected to each other compared to distant conduits, even though they do not make contact within the cross section analyzed (e.g., Loepfe et al., 2007).

Here we present a statistical method, point pattern analysis (PPA), which is useful for the study of the spatial properties of xylem cross-sections. The theory of PPA has been presented in several textbooks (Ripley, 1988; Diggle, 2003; Møller and Waagepetersen, 2003; Illian et al., 2008) and articles, including applications to ecology (e.g., Wiegand and Moloney, 2004; Wiegand et al., 2006; Lynch and Moorcroft, 2008). We determined the degree by which conduit distribution varied across species and wood types and compared the performance of various spatial indices. A direct comparison with the existing indices mentioned already could not be made, because the definitions are entirely different and the measuring approaches incompatible. As already noted by Carlquist (1984; cf. Rosell et al., 2007), complete spatial randomness in a cross section would still result in a certain fraction of the conduits being in contact with each other by chance. Our method automatically accounts for this issue, but other methods cannot. In addition, the higher the density and the wider the conduits, the more likely that some conduits will touch each other by chance. Hence, a vessel grouping index (sensu Carlquist) value = 1 will indicate perfect uniformity (disregarding overlapping vessel endings), whereas a value >> 1 cannot be interpreted to mean that the conduit spatial distribution shows a tendency toward clustering and across-species comparison is difficult, unless conduit size and density are simultaneously accounted for using a randomization procedure.

In addition to developing statistically based indices for the quantification of the distributional properties of xylem conduits, we also adopted a phylogenetic approach, and we asked the question of whether the phylogenetic patterns emerging using these statistical indices of spatial distribution were biologically meaningful and whether they broadly agreed with our prior knowledge of the presence and abundance of isolated vessels and vessel groupings, as found in previous studies on different floras (e.g., Carlquist and Hoekman, 1985; Baas and Carlquist, 1985; Herendeen et al., 1999).

MATERIALS AND METHODS

Anatomical images—Xylem anatomical images were downloaded from the website http://www.woodanatomy.ch (Wood Anatomy of Central European Species) (Schoch et al., 2004), which updates the identification key and the number of images available from a previously published wood anatomical textbook (Schweingruber, 1990). The database contains macroscopic and microscopic information for about 130 woody (tree and shrub) species, of which 105 were of sufficient quality to be used for further quantitative analysis. The complete list of species, with a synthetic list of the main features of the images analyzed and of the main anatomical characters of each cross section are given in Appendix S1 (see Supplementary Data at http://www.amjbot.org/cgi/content/full/ajb.0900289/DC1). For each species, the best cross-sectional image was selected out of the 1–6 available. For all images, the analyzed area effectively spanned one growth ring selected among those present in the image.

Image processing—Images were analyzed using two approaches. The first one employed a standard shareware software (ImageJ v.1.40, available from http://rsb.info.nih.gov/ij/), while the second made use of a bespoken software developed in our laboratory. The first approach required choosing image-by-image values for the gray scale and the minimum threshold for element size separating conducting vessels from other elements. Image-by-image checking and manual corrections were necessary to exclude nonvascular elements or include vessels that had not been selected (except obviously for the tracheid-bearing conifers). For all conduits in each image, we determined the x and y coordinates of the centers of their lumina as well as their respective lumen areas. The second approach was developed to allow semiautomated image analysis and is based on the experience in the analysis of biomedical images of one of the authors (Gil et al., 2006, 2009). The fundamental feature of this semiautomatic approach is the initial smoothing of the gray levels to enhance the quality of the binary image. Smoothing of the gray levels preserves the structure of the main conducting elements while reducing the visibility of the remaining matrix, thereby increasing the contrast between the two. The subsequent binary image was then subjected to two morphological filters (based on area and eccentricity) to remove nontubular structures. This approach still required calibration for each image, but the manual work required to cross-check them was drastically reduced. Using this second approach, we obtained the binary images of the finite-sized conduits, which were then used with the Programita software (see later).

Point pattern analysis—Point pattern analysis (PPA) is frequently used to detect whether a certain spatial pattern significantly departs from randomness. However, fitting more sophisticated nonrandom models and simulating the resulting spatial patterns is becoming more common thanks to available software (e.g., Spatstat; Baddeley and Turner, 2005), albeit at the cost of increased mathematical complexity and length of the learning curve.

One needs to distinguish between local variability in conduit density (as caused, e.g., by increases in conduit size, which almost inevitably reduce the density and cause the process to be spatially heterogeneous, otherwise referred

to as first-order variability of the local intensity) and the overall variability in conduit density (caused by interpoint interactions, i.e., conduit aggregation in clusters or lower conduit densities caused by conduit "repulsion", otherwise referred to as second-order variability of the intensity).

Our strategy to describe and interpret patterns of vessel distribution in xylem cross sections was as follows: (1) Fit Poisson models to the point pattern obtained from the analyzed images to test for significant departures from a random distribution, while at the same time accounting for local heterogeneity in conduit density. (2) Fit non-Poisson models accounting for the presence of point-to-point interactions, either of an aggregatory or of a repulsory nature and comparison with Poisson models. (3) Extract qualitative and quantitative indices of conduit aggregation/repulsion.

Poisson models—Conduit density can vary for different portions of an anatomical image. Assume that the expected number of conduits falling in a small region of area du around a location u is equal to $\Lambda(u)\,du$. Then $\Lambda(u)$ is the locally varying "intensity function" of the process [i.e., the expected number of conduits per unit of area, while the total number of conduits is $\int \Lambda(u)\,du$]. In the case of a spatially homogeneous process, this simplifies to $\Lambda \times$ area. We used an isotropic Gaussian kernel (the spatial equivalent of a moving average filter for time series) with a large bandwidth to estimate variations in conduit density due primarily to changes from early- to latewood. Comparisons between spatially homogeneous and heterogeneous Poisson models were carried out using values of the Akaike information criterion (AIC) for the two models.

A test of complete spatial randomness (CSR) was obtained by testing for departure from a Poisson null model. If the process obeys CSR, the conduits are independent of each other and have the same propensity to be found at any location (apart from variations due to local spatial heterogeneity). Hence, under Poisson conditions, the mean number of points for all regions equals the variance in the number of points across all regions. The probability density function of a Poisson model is:

$$f(x) = \alpha\beta^n \tag{Eq. 1a}$$

under homogeneous conditions and

$$f(x) = \alpha\prod_{i=1}^{n} b(x_i) \tag{Eq. 1b}$$

under heterogeneous conditions. Here, α is a normalizing constant (itself dependent on n), $\Lambda(u) = \beta$ is the expected number of points per unit of area under homogeneous conditions, $\Lambda(u) = b(x_i)$ is the locally varying vector for the expected number of points per unit of area under heterogeneous conditions, and n is the total number of points in the pattern (Møller and Waagepetersen, 2003; Baddeley et al., 2005).

Maximum likelihood estimators of Λ for Poisson processes were calculated using the algorithms of Berman and Turner (1992) implemented in the program Spatstat (Baddeley and Turner, 2005), for both the homogeneous and the heterogeneous case.

Nearest neighbor distances and Ripley's K and L functions—We computed the mean of the empirical distribution of Euclidean distances to the nearest neighbor, d_{nn}, for all conduits in a cross section and compared it against the expected nearest neighbor distance. For a homogeneous Poisson point process of intensity Λ, the expected nearest-neighbor distance is simply (e.g., Krebs, 1999):

$$d_{nn}^{E} = \frac{1}{2\sqrt{\Lambda}}. \tag{Eq. 2}$$

A simple index of uniformity is then:

$$R_{nn} = \frac{d_{nn}}{d_{nn}^{E}} \tag{Eq. 3}$$

If the spatial pattern follows Poisson, $R_{nn} = 1$. If clumping occurs, R_{nn} tends to zero, but if uniformity occurs, R_{nn} approaches an upper limit of about 2.15. A simple test of significance of the deviation from randomness is also available (e.g., Krebs, 1999). An edge correction was applied by clipping 10% of the window margins for the d_{nn} data (Baddeley, 1998).

We also computed a summary function known as the K of Ripley (and its main modification, the L function). Ripley's reduced second moment function K of a stationary point process is defined so that $\Lambda K(r)$ equals the expected number of additional points within a distance r of the typical random point of x. Because the expected $K_{pois} = \pi r^2$, measured values lower than K_{pois} indicate repulsion, whereas values larger than K_{pois} indicate aggregation. The L function is a simple modification of the K function:

$$L = \sqrt{\frac{K(r)}{\pi}} - r. \tag{Eq. 4}$$

This transformation stabilizes the variance and allows plotting the reference L_{pois} as a straight line with a value of 0 (i.e., positive values indicate aggregation, negative ones indicate uniformity). Edge corrections were implemented (Ripley, 1988; Baddeley, 1998). Finally, note that the K and the L functions are explicit functions of the distance r from the typical point of the pattern (R_{nn} is not), which implies that aggregated, random and/or uniform distributions can all co-occur at different spatial scales for the same point pattern (cf. Results section for one example).

Conduit-to-conduit interaction models—Pairwise interaction models have probability densities of the form

$$f(x) = \alpha\beta(x_i)^n\gamma(x_i, x_j) \tag{Eq. 5a}$$

for the homogeneous case and

$$f(x) = \alpha\left\{\prod_{i=1}^{n} b(x_i)\right\}\prod_{i<j}\gamma(x_i, x_j) \tag{Eq. 5b}$$

for the heterogeneous case (Baddeley et al., 2005). Again, α is the normalizing constant, $\beta(x_i)$ is the "first order} term (i.e., related to conduit density), n is the number of points in the pattern and $\gamma(x_i, x_j)$ (with $x_i, x_j \in W$) is the "second order" or "pairwise interaction" term between points x_i and x_j of the window W. The pairwise interaction term γ introduces dependency (either "repulsion" or "attraction") between conduits. In the hard core (HC) model, for example (e.g., Diggle, 2003), the term γ is chosen in such a way that:

$$\gamma(u, v) = 1, \text{ if the distance } \|u - v\| > l \text{ or}$$

$$\gamma(u, v) = 0, \text{ if } \|u - v\| \le l, \tag{Eq. 6}$$

where l is the HC radius and u, v are the coordinates of points x_i, x_j. In this way, it is impossible for points to be closer to each other than a distance equal to the HC radius. Beyond the HC radius, this model assumes no conduit-to-conduit interaction, and it reduces to a Poisson model (Møller and Waagepetersen, 2003).

The model that we selected for modeling xylem cross sections belongs to the class of Geyer models (Geyer, 1999). The saturation point process with interaction radius m, saturation threshold s, and parameters β and γ, is the point process in which each point x_i in the pattern X contributes a factor $\beta\gamma^{\min[s,t(x_i)]}$ to the intensity and the probability density function is:

$$f(x) = \alpha\beta(x_i)^n\gamma^{\min[s,t(x_i)]}, \tag{Eq. 7}$$

where $t(x_i)$ denotes the number of "close neighbors" of x_i in the pattern X. A "close neighbor" of x_i is a point x_j, such that the distance between x_i and x_j is less than or equal to m. Beyond the interaction distance m, no further conduit-to-conduit interactions are allowed and the model reduces to a Poisson model (i.e., $\gamma = 1$ for $\|u - v\| > m$) (Møller and Waagepetersen, 2003). Within distance m, if the saturation threshold s is set to $s = 0$, the model reduces to the Poisson point process. If s is a finite positive number, the interaction parameter γ may take any positive value, with $\gamma > 1$ describing a "clustered" or "attractive" pattern and $\gamma < 1$ describing an "ordered", "repulsive", or "inhibitive" pattern. Also, if $\gamma = 1$ the model reduces to a Poisson process, while if $\gamma = 0$ the model is a HC process (i.e., no points allowed within distance m of point x_i).

A HC radius was present by definition in all point patterns, because finite-sized conduits were represented by points. Therefore, we fitted a piecewise Geyer model (i.e., with two interaction distances), to all species. This approach

produced two values of γ (γ1 and γ2), with γ1 indicating the HC effect and γ2 representing the most important dimension of conduit-to-conduit interaction. For some species, up to six values of γ, for six different spatial scales, were necessary to minimize the AIC.

The best values of the parameters β and γ for the Geyer model were estimated by maximizing a pseudolikelihood function for a range of values of *m* and *s* (i.e., profile pseudolikelihood; Baddeley and Turner, 2005). To fit the point patterns, we used a Metropolis–Hastings algorithm, based on a Markov chain run to near-equilibrium using 10^6 iterations. The 95% confidence intervals of the γ values were obtained by simulating each point pattern stochastically 10 times using the same Metropolis-Hastings algorithm for the same set of parameters values and by refitting a Geyer model to obtain a distribution of γ values.

Goodness-of-fit tests—For the Poisson models, goodness-of-fit tests for the first-order processes were obtained with diagnostic plots of model residuals plotted against the two dimensions (*x*and *y*) of the point pattern (Baddeley et al., 2005). Goodness-of-fit of second-order processes was assessed instead using simultaneous critical 95% envelopes (Ripley 1981) obtained by Monte Carlo simulations of the point pattern and by using diagnostic plots of data quantiles plotted against mean quantiles of simulation residuals (with critical 95% bands; Baddeley et al., 2005).

Other analyses—The spatial tests based on the point patterns were run using the program Spatstat 1.15-4 for R 2.9.0 (Baddeley and Turner, 2005). To test the assumption that vessels can be represented as points, we applied Poisson models directly to all binary images using the Programita software (Wiegand and Moloney, 2004; Wiegand et al., 2006), which extends PPA to objects of finite size and variable shape (however, Geyer models cannot be fitted). As explained in the Results section, the assumption of the point representation of conduits was valid for all species except some of the ring-porous species, for which the very large size of the earlywood vessels violated the assumption of the point representation of conduits. Strictly, it is also inappropriate to treat earlywood and latewood vessels as two subpopulations of the same population, because they are spatially segregated into different regions of the cross section. Hence for the ring-porous species, we conducted marked point pattern analysis in Spatstat, by marking earlywood and latewood vessels separately. If Programita and Spatstat coincided on the spatial pattern, the ring porous species was then classed under that spatial pattern. However, if the spatial patterns differed between the two types of analyses, the species could not be classified and it was discarded. Programita and Spatstat always agreed for conifers, diffuse porous, and semi-ring porous species.

Phylogenetic analyses—The central phylogenetic tree was obtained using the maximally resolved seed plant tree from Phylomatic (http://www.phylodiversity.net/phylomatic) (Davies et al., 2004; Webb and Donoghue, 2005). Additional information regarding specific families, genera, and species within genus was added using the user-supplied data repository in Phylomatic for the Pinaceae (Rydin, 2002), Rosaceae (Potter, 2007) and Betulaceae (Chen, 1999). Most of the remaining polytomies were resolved by conducting a reference search for the following clades: the Pyrinae (Campbell et al., 2007), the genera *Acer* (Li et al., 2006), *Alnus* (Navarro et al., 2003), *Betula* (Järvinen et al., 2004), *Pinus* (Wang et al., 2000; Gernandt et al., 2005), *Populus* (Hamzeh and Dayanandan, 2004), *Prunus* (Wen et al., 2008), and *Rosa* (Koopman et al., 2008). The genera *Salix*, *Ribes*, *Lonicera*, and *Sorbus* were left as unresolved polytomies.

The phylogenetic tree was originally composed of the 105 species of the data set. This needed to be pruned to a final number of 99 species because of missing estimates for some of the spatial parameters. Because our final phylogeny was obtained by combining trees from various sources, we could not incorporate information on branch length, and we assumed equal branch lengths throughout. The 99 species were distributed among the major clades of the core eudicots phylogeny and the Pinales. Among the eudicots, the order Saxifragales ($N = 3$ species), euasterids I ($N = 1$), euasterids II ($N = 11$), eurosids I ($N = 60$), and eurosid II ($N = 6$) were represented. In addition, in the asterids, the Cornales, the Vitaceae, the Buxaceae, the Proteales and the Ranunculales were also represented. Among the Pinales, the Pinaceae ($N = 9$), Cupressaceae ($N = 2$), Taxaceae ($N = 1$) were represented.

The variance–covariance matrix of the phylogenetic distances was then calculated (Freckleton et al., 2002; Paradis, 2006), together with the coefficient λ (Pagel, 1999; Freckleton et al., 2002), which gives a quantitative measure of phylogenetic dependency of a variable, based on a Brownian motion evolutionary model. A value of λ of zero indicates evolution of traits independent of phylogeny, whereas a value of 1 indicated that traits on the given phylogeny have evolved according to a Brownian motion model. The departures of λ from the theoretical values of 0 and 1 were assessed using a χ^2 test (Freckleton et al., 2002). This approach is considered more informative than the one based on phylogenetic independent contrasts (Freckleton et al., 2002). All phylogenetic tests were run using APE, pglm3.2 and ade4 for R 2.9.0 (Freckleton et al., 2002; Paradis, 2006).

RESULTS

One species in detail—We first illustrate the procedure adopted in determining the statistical properties of the point pattern, i.e., the application of the Poisson and the piecewise Geyer function, for one species taken as an example. Then, we move on to illustrate the results obtained for the other species.

The cross section employed to quantify the properties of the xylem of one growth ring of *Acer pseudoplatanus* L. is plotted in Fig. 1A. The spatial kernel detected a significant spatial trend in the *Y* direction for this species, with conduit densities varying from around 70 conduits·mm^{-2} in the bottom part of the image to around 40 conduits·mm^{-2} in the top part (not shown), for a mean conduit density of around 54 conduits·mm^{-2}. As a consequence, a heterogeneous model fitted the data better than a homogeneous one (AIC dropped from 5502 to 5484 from the homogeneous to the heterogeneous Poisson model) and a spatial kernel was retained for all subsequent analyses. This occurred in virtually all cross sections examined; hence, the use of a broad spatial kernel was set as a default for all species. Further spatial trends are evident in Fig. 1A. For example, radial parenchyma rays constrain the position of the conduits in more or less regular files. Additionally, vessels often (but not always) appear forming multiples of two or three, although in one case (center of the picture) as many as eight are detected in a group. Figure 1B shows the representation of the xylem network of Fig. 1A in terms of points centered in the geometrical center of each vessel. The size of the point is arbitrary. Figure 1C shows the application of an *L* function to this cross section as a function of the mean radius *r* of a circle drawn around a random point of the pattern. In the case of a homogeneous Poisson, the theoretical reference line for a random distribution would be a horizontal line at $Y = 0$. However, because of the spatial variability in conduit density accounted for by the kernel and the application of edge corrections, the theoretical line estimated using 38 Monte Carlo realizations (represented with a dashed line in the figure) deviates somewhat from the horizontal zero line, particularly at larger distances, i.e., greater than 200 μm. The deviations in the empirical Poisson curve (shown as a bold line) at distances shorter than about 200 μm are instead caused by point-to-point interactions. The first minimum at $r = 20$–30 μm (which significantly departs from the critical downward 95% envelope (given as a dotted line) corresponds to the HC radius (mean conduit radius for this species was 24.7 μm), the second upward peak (which is not significant) at about $r = 50$ μm corresponds to a distance of about twice the mean conduit radius and reflects the aggregation of some of the points in groups of predominantly two. The third downward peak at $r =$ 103 μm (which is barely not significant) corresponds to the mean distance between parenchyma rays, which tends to cause a more regular distribution of the points along the *y*-axis. Application of the piecewise Geyer model to detect departures from the random distribution described by the Poisson model (Fig. 1D) detected in sequence the three levels of spatial

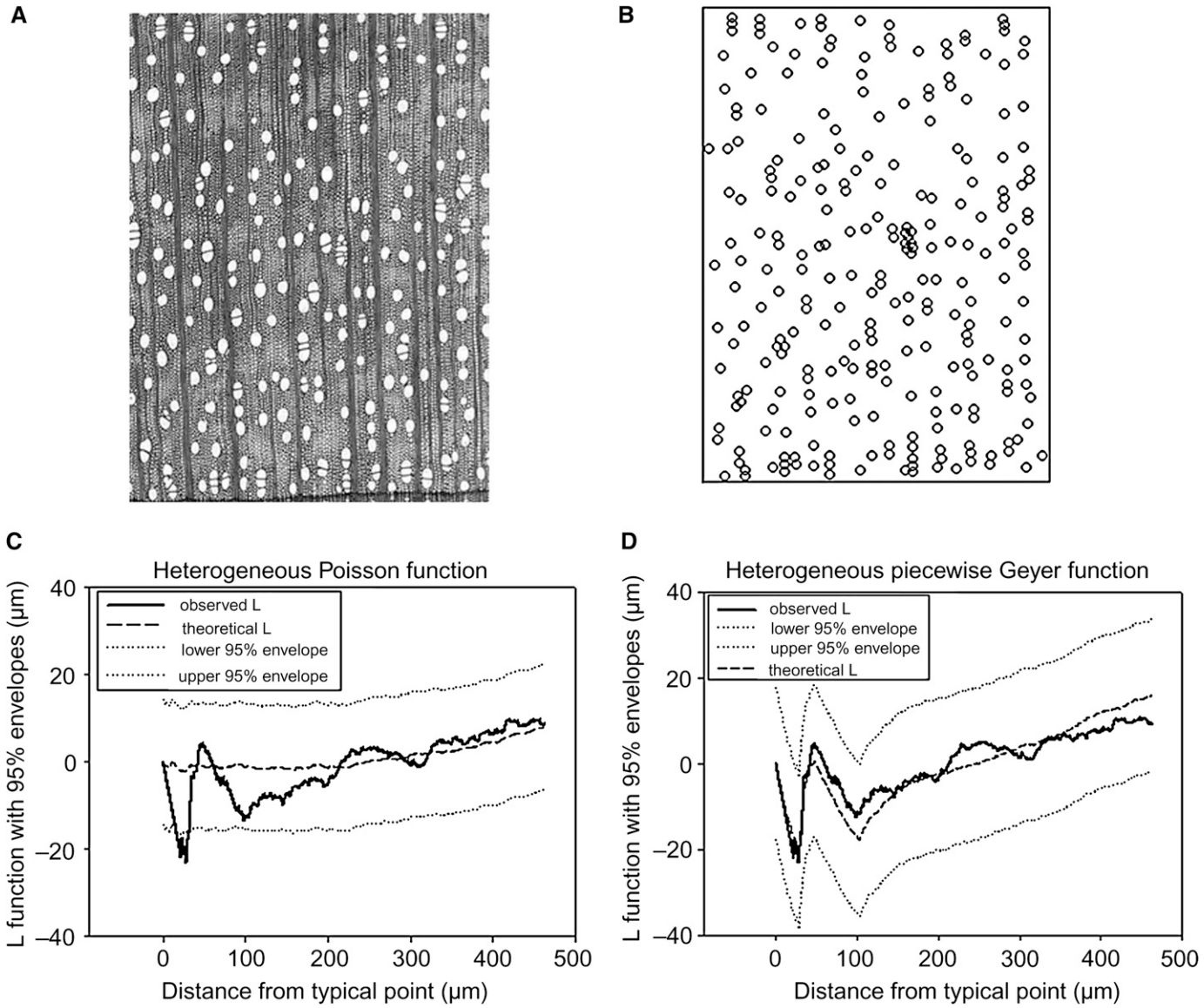

Fig. 1. Example of the steps applied for the spatial analysis of a cross section of *Acer presudoplatanus*. (A) Original cross-sectional image used for binary acquisition and point determination. (B) The resulting point pattern. (C) The application of a random Poisson model in the form of the empirical (bold solid line) and theoretical (dashed line) *L* function (with simultaneous critical 95% upward and downward envelopes, shown as dotted lines), estimated using Monte Carlo simulations. (D) The application of a piecewise Geyer model, with three interaction distances over which conduit-to-conduit interaction takes place. Line scheme as in (C).

organization, corresponding to the three processes previously described and also visible in Fig. 1A. Hence the measured *L* curve now lies well within the 95% envelopes and sits very close to the estimated theoretical line. In addition, the piecewise Geyer model progressively lowered the AIC to 5441, 5423 and 5247, all three of which are well below the value obtained with the heterogeneous Poisson model (5484). The three conduit-to-conduit interaction parameters ($\gamma 1$, $\gamma 2$, and $\gamma 3$) for the three spatial scales had values of 0.24, 1.53 and 0.66 (with 95% confidence intervals of ±0.30, ±0.35 and ±0.05), respectively.

Spatial analyses for ring porous species—Due to the limitations inherent in current available software for spatial PPA, a special decision-making process had to be adopted for the ring porous species. An example of the marked-point analysis conducted on a ring porous species, *Ulmus glabra* Huds., is given in Fig. 2. Figure 2A shows the corresponding binary image employed in Programita. The species shows clear clustering in the latewood vessels. The few earlywood vessels also suggest clustering, but they occur in very small numbers in the cross section. To conduct the spatial analysis, we marked large earlywood and small latewood vessels separately, using the size of the conduit of mean conductance (i.e., $[\Sigma r(i)^4]^{1/4}$ to separate the two groups. A Poisson model was then applied to the two populations independently, accounting for their different spatial locations in the cross-section. Figure 2B (top left panel) shows that the small vessels taken in isolation tend to cluster at spatial scales larger than about 40 μm, after the occurrence of a significant HC radius up to r = 30 μm. Unsurprisingly, given their locations, small vessels tend to cluster away from the large

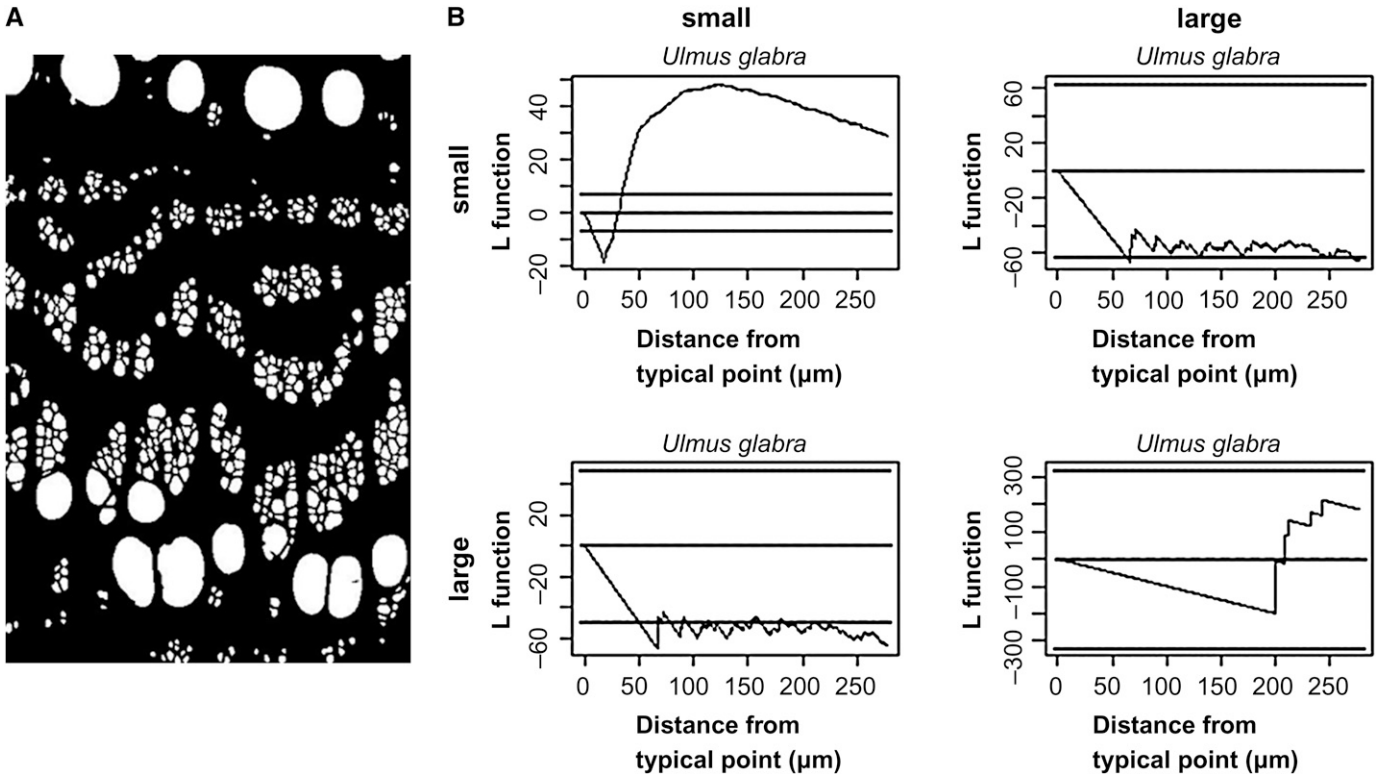

Fig. 2. Example of the approach used to quantify conduit spatial distribution in the case of ring porous species. (A) Binary image of *Ulmus glabra* used directly in Programita. (B) Marked point pattern function obtained in Spatstat for the interaction between the two subpopulations of large earlywood and small latewood vessels. Top left panel shows the interaction among small latewood vessels, the bottom right the interaction among the large earlywood vessels. Distribution of the small latewood vessels in relation to the large earlywood ones is shown in the bottom left panel; the opposite distribution of the large earlywood vessels in relation to the small latewood ones is shown in the top right panel. In each chart, the central horizontal line is the theoretical Poisson line drawn at $L = 0$; the outer two horizontal lines are the 95% envelopes.

vessels at almost all spatial scales (bottom left panel), whereas large vessels only cluster away from the small vessels at the HC radius of about $r = 70$ μm (top right panel; compare with binary image). Nothing can be said of the behavior of the large vessels taken in isolation due to their very large 95% envelopes (bottom right panel). A Geyer model applied to the two separate populations of vessels gave values of γ of 2.0 (small vessels, significant at 95%) and 28.0 (large vessels, not significant at 95%). The analysis with the Programita software (which randomizes objects with their real size using the binary image of Fig. 2A) revealed a significant (95%) overall aggregation. This species was therefore classed as aggregated with an overall $\gamma 2 = 1.37$ (from the piecewise Geyer model; confidence intervals of ±0.06). A summary of the results obtained with this method for the ring porous species is given in Table 1. In many cases, the results of the Programita software coincided with those obtained with Spatstat, with the exception of *Clematis vitalba* L. and *Prunus armeniaca* L. These species (plus *Fraxinus excelsior* for which the binary image was not available) were therefore excluded from subsequent calculations. Generally speaking, aggregation prevailed in latewood and randomness prevailed in earlywood for all these species.

Spatial analyses for all species—The results of the application of the heterogenous Poisson model to 105 species are reported in Appendix S2 (see online Supplemental Data), including the classifications into uniform, random and aggregated distribution based on the *L* plots and on the calculation of the mean uniformity index, R_{nn}. The figures with the *L* functions for all the 105 species with the fit to the Poisson model are given in online Appendix S3.

A summary of the results of the heterogeneous piecewise Geyer model is given in Table 2, by grouping the 105 species according to the spatial distribution pattern obtained (uniform, random, aggregate, Table 2A) or according to their xylem type (tracheid-bearing, ring porous, semi-ring porous, diffuse porous, Table 2B). The results for all individual species are given in online Appendix S4, while the figures with the *L* functions and the fit to the piecewise Geyer model to individual species are given in online Appendix S5. The two interaction parameters (γ1 and γ2) correspond to the HC radius and the main Interaction distance.

Overall, the results were improved considerably compared to the Poisson model (Fig. 1D; figures reported in the Supplemental Data). The HC radius appeared to match or slightly exceed the radius of the conduit with mean cross section, with the exception of the ring-porous species. Correspondingly, γ1 (which indicates the strength of the interaction for distances up to the HC radius) gave extremely low values, indicating the very small probability of encountering another conduit within the (mean) radius of the typical conduit. Among the wood type groups, the tracheid-bearing conifers showed the shortest interaction distance, a reflection of their small average conduit size and the extreme uniformity in their behavior, whereas the largest distances were found for the ring-porous species (whose conduits were larger) and in the diffuse-porous species

(in which aggregation prevailed, cf. their $\gamma 2$ values). As a group, the semi-ring porous showed a tendency toward uniformity (53 of 59 species) and the ring-porous a tendency toward aggregation (six of 11 species). The saturation parameter varied little across groups.

For 26 species, a third fitting step was required to bring the observed L function within the 95% envelopes and minimize the AIC, as previously shown for *Acer pseudoplatanus* L. They are summarized in online Appendix S6. *Acer pseudoplatanus* L. and *Picea abies* (L.) H. Karst. were the only cases in which this resulted in the simultaneous presence of both uniform and aggregated behavior at different spatial scales. Classification of these two species was based on the significance of the primary departure from Poisson.

Relationships among variables—The two spatial variables, $\gamma 2$ and R_{nn}, were significantly and inversely related to each other (Fig. 3, $R^2 = 0.84$, $P < 0.001$), suggesting that a large fraction of the spatial structure that was measured by the Geyer model occurred at the small spatial scale of the nearest neighbor where the uniformity index is measured. Vice versa, the discrepancy among the two indices for some species (top right quadrant) suggests that some of the spatial structuring occurred at larger spatial scales. Note that the nature of this relationship (i.e., the exponential fit) is the one predicted by the exponential family of the distribution functions (cf. Eqs.1, 5, and 7). To facilitate the understanding of this relationship, we have reported one typical binary image in each of the three quadrants where points were found (Fig. 4), with each image corresponding to the numbered points of Fig. 3. *Carpinus betulus* (point 1 and top left quadrant) is a typical example of aggregative behavior both at the scale of the nearest-neighbor (as estimated by R_{nn}) and at larger scales (with separate groups of conduits themselves aggregated along vertical bands). *Pyrus communis* (point 2 and bottom right quadrant) is instead an example of uniform distribution of individual conduits and of the section as a whole. *Hedera helix* (point 3 and top right quadrant) shows an example of species in which individual conduits are uniformly distributed, but clear structures are visible at larger scales, with groups of conduits aggregated along vertical bands. Interestingly, no species was found in the fourth quadrant (i.e., with small scale aggregation and large scale uniformity). This combination could be realized, for example, by having conduits arranged in groups of two but with the groups uniformly dispersed in the cross section. Some of the *Acer* species come close to this distribution without however reaching it.

Although they explained a relatively low proportions of variance, we found highly significant correlations between conduit density and both $\gamma 2$ (Fig. 5A, $R^2 = 0.30$, $P < 0.01$) and R_{nn} (Fig. 5B, $R^2 = 0.46$, $P < 0.01$). Cross sections with low values of $\gamma 2$ (i.e., where uniformity was present) tended to have significantly more conduits/mm^2 than sections with large values of $\gamma 2$ (i.e., those characterized by aggregation). Although the regressions given in Fig. 5 are largely driven by the conifers, systematic differences in conduit density were evident also across the major angiosperms groups (data not shown).

Phylogenetic analyses—Significant phylogenetic effects were detected for the coefficient λ (Table 3, χ^2 test for the null hypothesis that $\lambda = 0$) for the two spatial variables considered, i.e., $\gamma 2$ and mean R_{nn}. The result was robust to conducting the analysis for all species combined and separately for the angiosperms, whereas a significant phylogenetic effect for the 12 tracheid-bearing conifers was not detected, possibly because of the very small sample size. Interestingly, for all species combined as well as for the angiosperms (but not for R_{nn}), values of λ were also significantly different from 1, indicating that trait evolution was not entirely explained by a Brownian evolutionary model along the considered phylogeny. For all the species combined, and separately for the angiosperms, values of λ were larger for R_{nn} compared to $\gamma 2$.

The final phylogenetic tree is shown in Fig. 6, with the tree being color-coded to give the parsimony reconstruction according to R_{nn}, the uniformity index. The Fagales (in genera such as *Alnus*, *Betula*, *Ostrya*, *Corylus*, and *Carpinus*), the Fabales (genera *Robinia* and *Laburnum*) and some of the Rosales (genera *Frangula*, *Rhamnus*, some of the *Ulmus*), all of which are in the eurosids I, emerged as the orders characterized by abundance of aggregative structures ($R_{nn} < 1$). The genera *Ilex* (in euasterids II), *Berberis* (in Ranunculales), and *Vitis* (in rosids) also showed R_{nn} smaller than 1, indicating aggregation. Vice versa, the Rosaceae were notably characterized by a clear uniformity in conduit distribution, with particularly high values of R_{nn} recorded for *Amelanchier*, *Pyrus*, and some *Prunus*.

The parsimony reconstruction suggested several independent origins for conduit uniformity and aggregation. In particular,

TABLE 1. Results of the combined Spatstat and Programita spatial analyses for the ring-porous species.

Species	Geyer γ for earlywood vessels and level of significance	Geyer γ for latewood vessels and level of significance	Classification according to Programita (early- and latewood vessels combined)	Final $\gamma 2$ value from piecewise Geyer model
Castanea sativa	(34.82) NS	**1.13** *	Random	1.13
Clematis vitalba	(7.62) NS	**1.43** *	Uniform	Discarded
Fraxinus excelsior	(2021.4) NS	**1.61** *	NA	NA
Laburnum alpinum	(1.67) NS	**1.61** *	Aggregate	1.33
Laburnum anagyroides	(**1.60**) *	**2.33** *	Aggregate	1.42
Prunus armeniaca	(68.14) NS	13.00 NS	Random	Discarded
Quercus petraea	(28.25) NS	**4.55** *	Random	0.96
Quercus robur	(4.79) NS	**1.16** *	Random	0.99
Robinia pseudoacacia	(535.34) NS	**1.46** *	Aggregate	1.82
Rosa canina	(1.64) NS	1.60 NS	Random	0.62
Ulmus campestris	(24.95) NS	**1.42** *	Aggregate	1.42
Ulmus laevis	(614.17) NS	**1.35** *	Aggregate	1.99
Ulmus scabra	(28.05) NS	**2.05** *	Aggregate	1.47
Vitis vinifera	(366.61) NS	**1.50** *	Aggregate	1.45

Notes: Level of significance for Geyer γ obtained by multitype analysis, using the simultaneous 95% envelopes around the Poisson distribution. Values in boldface and *, $P < 0.05$; values not in bold and NS, not significant; NA, not available.

TABLE 2. Summary of statistics (±SEs in brackets) for the Geyer piecewise saturating models on the 105 species grouped according to (A) the spatial distribution of their conduits or (B) their wood type.

	Mean radius	Hard core radius	Interaction distance	Saturation parameter	AIC	β	γ1	γ2
A) Conduit spatial distribution group								
Uniform distribution	21.59 (0.12)	24.85a (0.14)	41.98 a (0.33)	2.75 a (0.02)	7224 (38)	2.65 a (0.025)	0.18 (0.003)	0.52 (0.002)
Uniform distribution (without conifers)	23.21 (0.13)	26.68 (0.16)	46.12 (0.40)	2.87 (0.03)	7025 (48)	2.02 (0.01)	0.18 (0.004)	0.55 (0.002)
Random distribution	19.60 (1.99)	16.85 (0.70)	40.00 (3.82)	1.11 (0.04)	5095 (295)	0.90 (0.05)	0.033 (0.01)	1.54 (0.20)
Aggregate distribution	22.79 (0.56)	22.53 a (0.30)	67.50 b (1.01)	2.62 a (0.07)	4456 (130)	0.44 b (0.007)	0.048 (0.006)	2.12 (0.04)
B) Conduit type group								
Semi-ring-porous angiosperms	21.47 (0.15)	23.88 a (0.15)	43.77 b (0.38)	2.51 a (0.02)	7035 (51)	1.69 b (0.02)	0.13 (0.00)	0.86 (0.02)
Ring-porous angiosperms	33.48 (0.82)	25.30 a (0.47)	60.71 a (2.15)	3.21 a (0.15)	4286 (207)	0.70 c (0.03)	0.05 (0.01)	1.23 (0.04)
Diffuse-porous angiosperms	20.75 (0.33)	26.49 a (0.62)	65.75 a (1.42)	2.60 a (0.12)	4963 (133)	1.50 b (0.06)	0.20 (0.01)	1.94 (0.09)
Conifers	11.46 (0.36)	15.72 b (0.51)	21.25 c (0.52)	2.17 a (0.13)	8220 (144)	5.81 a (0.18)	0.14 (0.01)	0.35 (0.01)

Notes: Mean radius, radius (μm) of the conduit of mean cross-sectional area; hard core radius, radius (μm) of circle within which no other conduits are allowed; interaction distance, radius of circle (μm) within which pairwise conduit-to-conduit interactions take place; saturation parameter, number of neighbors allowed within the Interaction distance for which the Geyer function saturates; AIC, Akaike information criterion; β, intensity parameter; γ1, pairwise interaction parameter for the circle of hard core radius; γ2, pairwise interaction parameter for the circle of radius equal to the interaction distance. Letters after the means refer to planned comparisons in one-way ANOVA using Fisher's least significance difference (LSD).

high values of conduit uniformity were found in all the major clades (i.e., eurosids I, eurosids II, euasterids I and II). In addition, the three nodes deepest in the tree, marking the departures of the Ranunculales, the Proteales, and the Buxaceae, also showed widely ranging values of both R_{nn} and γ2, i.e., suggesting strong aggregation for *Berberis* but strong uniformity for both *Platanus* and *Buxus*.

DISCUSSION

General considerations—We applied PPA to the investigation of the spatial distribution of xylem conduits in cross-sectional images. We applied methods using classical PPA, which assumes that objects can be represented by points located at their geometrical center, as well as methods that expand on classical PPA, by analyzing objects with real size directly from binary images. In general, both approaches worked well, although current analytical achievements are limited by the development of the statistical theory behind PPA and by the availability of appropriate software that can overcome various limitations. For example, the Geyer model that we deployed is suitable to detect only moderate levels of aggregation, whereas it tends to underestimate the degree of grouping that is detectable with, for example, Cox process models (Møller and Waagepetersen, 2003; Baddeley and Turner, 2005). Our choice of applying a Gibbs approach was due to the flexibility of the piecewise Geyer model, which allows one to determine a set of interaction coefficients whose interpretation is relatively straightforward and can accommodate a range of situations.

Progress in software availability is also required. Spatstat is a very flexible package, but it currently lacks the capability of handling binary images to study spatial patterns of finite-sized objects. On the other hand, Programita does not have the suite of functions that are available in Spatstat, and model fitting and simulation are significantly more limited.

Xylem spatial patterns and their significance—On the basis of our Poisson analysis, between 87 and 93 species of 105 (depending on whether we classified their distribution on the basis of the nearest neighbor test or the 95% envelopes of the Poisson *L* function) appeared to violate the assumption of random vessel spatial distribution. In addition, based on our Geyer analysis, significantly better models (as judged by AIC) were obtained in all cases when a conduit-to-conduit interaction parameter was added compared to a simpler random distribution model; finally, in only nine species of 105, was γ2 not significantly different from 1.00 (as estimated from the calculated 95% confidence

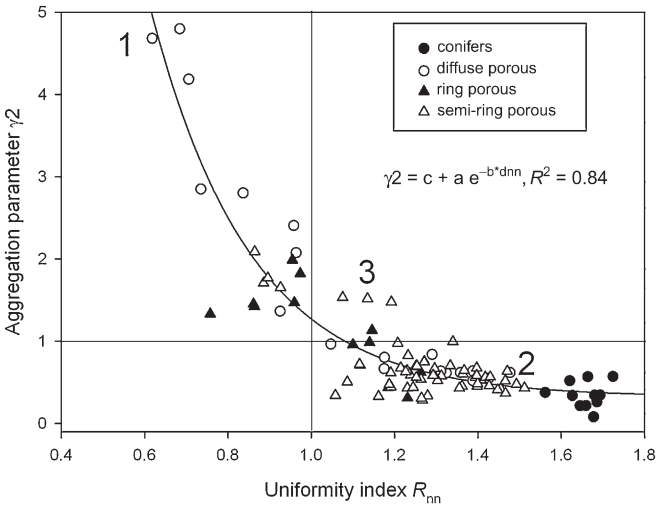

Fig. 3. Exponential relationship between the conduit-to-conduit interaction parameter γ2 and the uniformity index R_{nn}. The vertical line at R_{nn} = 1.0 and the horizontal line at γ2 = 1 separate the figure into four quadrants. The top left one identifies the species with aggregated conduit distribution ($R_{nn} < 1.0$, γ2 > 1). The bottom right one identifies the images with uniform conduit distribution ($R_{nn} > 1.0$, γ2 < 1). The remaining two quadrants identify regions of the space where the two indices indicate uniformity according to R_{nn} and aggregation according to γ2 (top right) or vice versa (bottom left). Only the first one of these two is occupied by a few points. Numbers 1, 2, and 3 refer to the binary images used in Fig. 4. Points belonging to the major wood types are plotted using different symbols.

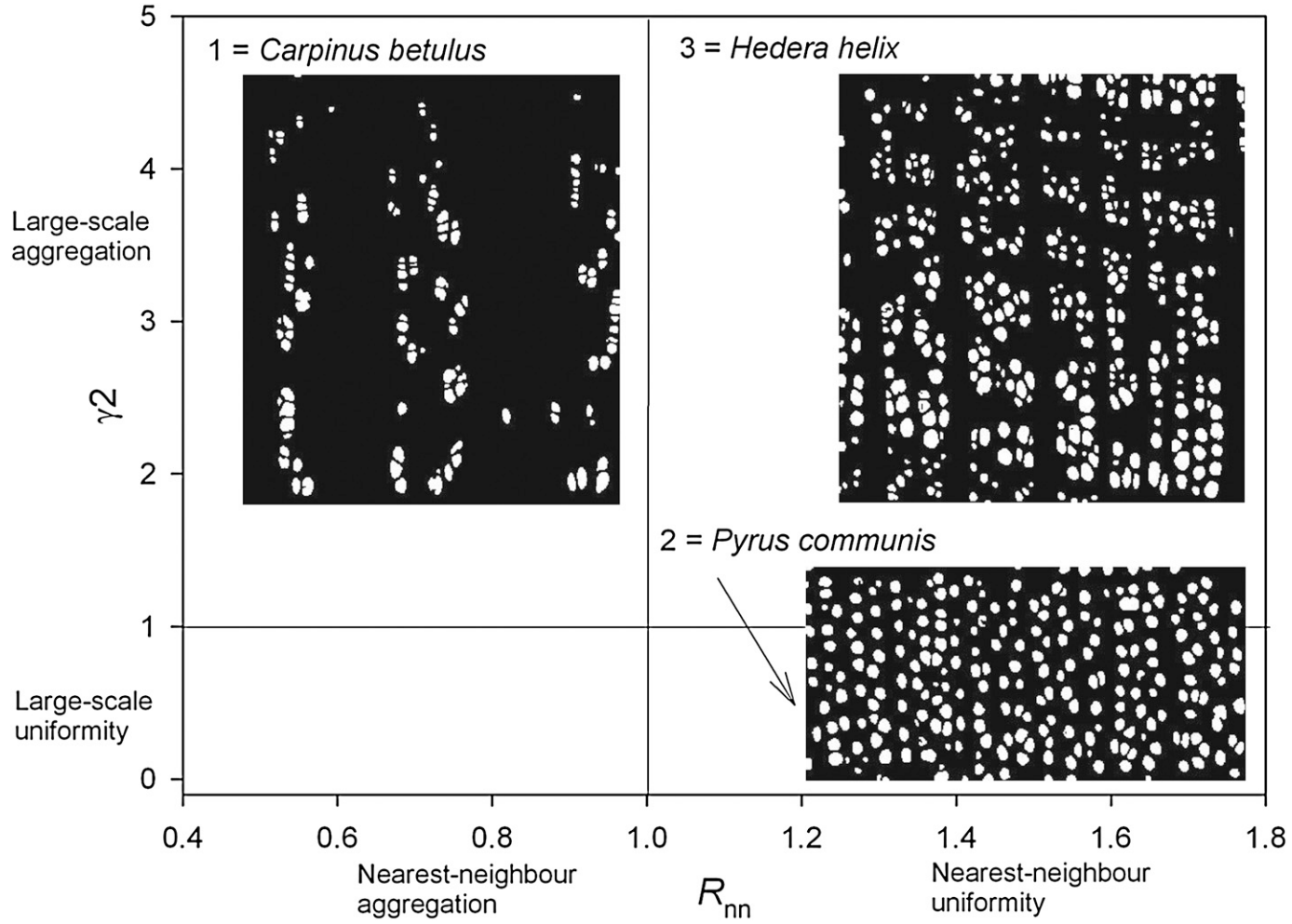

Fig. 4. Plot similar to Fig. 3 for the points numbered 1 to 3. The top left quadrant shows the example of *C. betulus*, a species with aggregated conduit distribution ($R_{nn} < 1.0, \gamma 2 > 1$). The bottom right quadrant reports the example of *P. communis*, a species with uniform conduit distribution ($R_{nn} > 1.0, \gamma 2 < 1$). The top right quadrant reports the example of *H. helix*, a species with conduits uniformly distributed at small spatial scale, but overall aggregated in groups and vertical bands at larger scales.

intervals). Taken together, these results raise the question of why so many species appear to have a spatial distribution which follows non–random patterns (cf., Carlquist and Hoekman, 1985; Wheeler et al., 2005; Rosell et al., 2007).

Although preliminary, our results suggest that spatial conduit distribution is likely to be an important anatomical character, for two main reasons. First, highly significant relationships were found between both R_{nn}, $\gamma 2$, and the density of conduits, suggesting that species characterized by conduit aggregation tended to have fewer conduits than species with conduit uniformity. Second, for the two spatial parameters examined ($\gamma 2$ and R_{nn}), values of λ highly significantly different from 0 suggested that phylogeny exerted a significant constraint in the evolution of these characters (Pagel, 1999; Freckleton et al., 2002).

With regard to the first aspect, it is interesting to note that significant relationships were found for both R_{nn} and $\gamma 2$ and conduit density. As previously explained, R_{nn} quantifies only one aspect of the spatial patterns, i.e., that related to physical closeness to the nearest neighbor, hence the close agreement between the two suggests that much of the conduit-to-conduit interactions take place at small spatial scales. In four species (*Castanea sativa* Mill., *Juglans regia* L., *Rhamnus cathartica* L., and *Hedera helix* L.) a clear discrepancy in the final classification was found, with R_{nn} implying uniformity and $\gamma 2$ implying aggregation. This demonstrated that, in those four species, uniformity occurred at small scale and aggregation at larger scales, in dendritic patterns. Multiple processes at different spatial scales probably occurred in many more species (e.g., Fig. 1D). In a few cases (e.g., including *Rhamnus*), this was demonstrated by the use of a piecewise Geyer model (see summary figures in the Supplemental Data), but in other cases, this may have gone undetected because of our use of a HC radius.

Despite the frequent occurrence of multiple processes at different spatial scales, in many cases this simply reinforced the typical pattern, i.e., multiple occurrences of aggregation or multiple occurrences of uniformity. This explains why, despite the presence of multiple spatial scales of interaction, R_{nn} and $\gamma 2$ largely agreed with one another. Application of the Geyer model directly on binary images should detect the presence of multiple processes at different spatial scales more effectively than it was possible in this work.

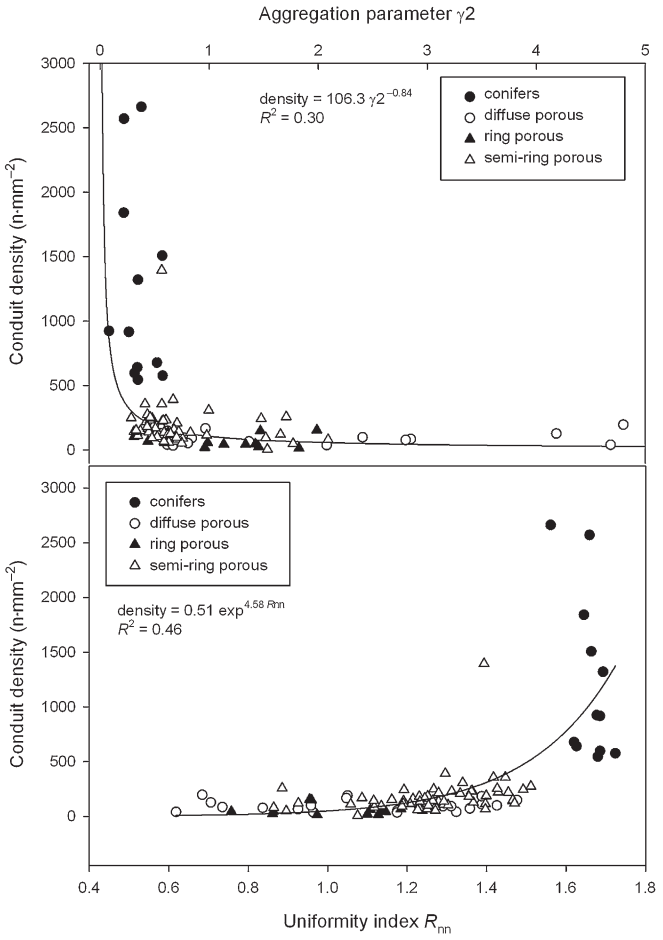

Fig. 5. Univariate regressions between the interaction parameter γ2 (panel A), the uniformity index R_{nn} (panel B) and conduit density. Although the percentages of variance explained are relatively low, both slopes are significantly different from zero ($P < 0.0001$). Wood types plotted by different symbols.

Because it is defined in relation to the distance to the nearest neighbor, R_{nn} may be regarded as being closer to the definition of the traditional anatomical grouping indices (e.g., Carlquist, 2001), with their emphasis on physical contact among conduits. In other words, the fact that both γ2 and R_{nn} were similarly related to conduit density, suggests that the relevant physiological dimension of conduit spatial distribution had largely to do with physical contact or proximity to the nearest neighbors. This was confirmed by our estimates of the Saturation parameter (Table 3), which typically had values of between 2 and 3 conduits.

The preliminary conclusion that conduit density was lower in species with conduit aggregation contrasts with the conclusions gained using different indices of conduit grouping (i.e., conduit grouping index sensu Carlquist), which instead generally suggested that more vessels tended to be present in species with clustered conduits (e.g., Carlquist and Hoekman, 1985). This discrepancy can be due to various reasons, i.e.,(1) the lack of comparability between our definition of uniformity and aggregation vs. the definition of vessel clustering sensu Carlquist, (2) the difficulty of detecting the small vessels tightly packed in the cross sections we examined, i.e., a systematic underestimation of conduit density for the species with aggregation, or (3) the occurrence of different syndromes among woody plants from temperate vs. Mediterranean and desert regions.

In any case, the finding that conduit spatial distribution can affect conduit density is important and needs to be explored further in future studies, to clarify better both the mechanisms likely to be involved as well as the possible physiological and ecological tradeoffs.

With regard to the second aspect (the phylogenetic signal), we first note that our sample was limited to one cross section per species, selected primarily for its image quality. Hence, we do not make any claim regarding the repeatability of the spatial distribution at the scale of the individual species studied. However, broad comparisons within families clearly showed that our spatial indices did capture an evolutionary axis that appears to be well conserved within particular clades, as reflected by values of λ values relatively close to 1 for both indices. This is not surprising, because spatial features constitute important species-specific identification characters (IAWA Committee, 1989). Broadly speaking, the prevalence of aggregated distributions in the Fagales and the Fabales and the prevalence of uniformity in the Rosaceae corresponds to known spatial trends (e.g., Carlquist and Hoekman, 1985; Herendeen et al., 1999).

Of interest was also the fact that, for both spatial parameters, the values of λ were often significantly lower than 1.00, suggesting that, although important, phylogenetic constraints did not impede evolution of similar spatial features across unrelated clades or vice versa evolution of fairly dissimilar patterns within the same clade. This observation is important and suggests that spatial characters could partly be the result of adaptive evolutionary processes, which is in line with previous observations that some spatial features tend to be found in species living under particular environmental conditions (e.g., Carlquist, 1984, Zanne et al., 2006). Values of λ for R_{nn} were much closer to 1 than those for γ2, and in the case of the angiosperms, the χ^2 tests showed that only λ for γ2 significantly differed from 1.00. This suggests that small scale spatial structures (as estimated by R_{nn}) are more phylogenetically conserved than large scale patterns (as estimated by γ2). This may be understood

TABLE 3. Values of the parameter λ (and associated log-likelihood ratio in brackets), quantifying the extent by which phylogenetic effects are present for the two spatial variables γ2 and R_{nn}. γ2, conduit-to-conduit interaction coefficient from the Geyer piecewise saturating model; R_{nn}, uniformity index. χ^2, χ^2 test to determine whether the observed λ is significantly different from 0 and from 1, respectively (where λ=0 indicates character evolution independent of phylogeny, and λ=1 indicates character evolution following a pure Brownian motion model). Symbols for significance of the χ^2 test are given in bold.

	All species			Angiosperms only			Conifers only		
Spatial indices	λ (ln lik)	χ^2 (λ = 0)	χ^2 (λ = 1)	λ (ln lik)	χ^2 (λ = 0)	χ^2 (λ = 1)	λ (ln lik)	χ^2 (λ = 0)	χ^2 (λ = 1)
γ2	0.62 (–93.0)	26.1 ***	40.5 ***	0.66 (–90.3)	26.8 ***	29.0 ***	1.00 (5.35)	0.30 **ns**	0.00 **ns**
R_{nn}	0.92 (40.79)	41.4 ***	6.19 *	0.96 (29.6)	46.10 ***	2.83 **ns**	0.00 (22.2)	0.00 **ns**	5.89 *

Notes: ns, $P > 0.05$; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

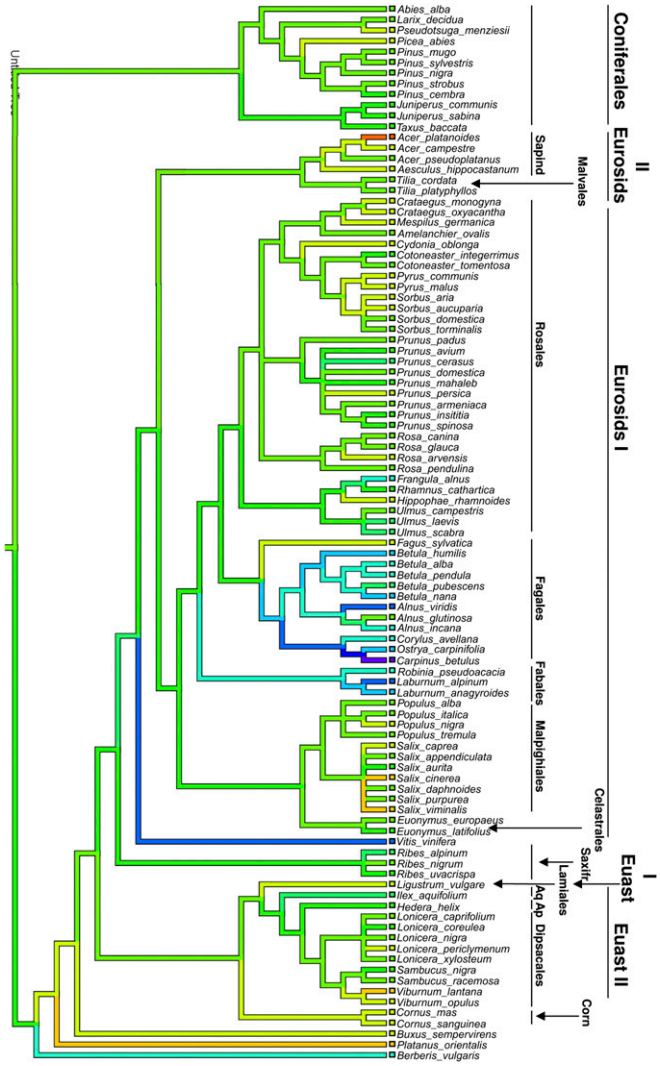

Fig. 6. Phylogenetic tree employed to account for the phylogenetic dependence among the 105 species employed in these analyses. Character reconstruction was obtained using a parsimony method on the measured values of R_{nn}. A similar tree constructed using $\gamma 2$ values is given in Appendix S3 of the online Supplemental Data. Color scheme: red to orange: strong uniformity; orange to light green, mild uniformity; intense green to pale blue: mild aggregation; intense blue to violet: strong aggregation. Ap = Apiales, Aq = Aquifoliales, Corn = Cornales, Euast = euasterid, Sapind = Sapindales

by assuming, e.g., that large-scale dendritic patterns show a lower degree of phylogenetic conservatism or that the frequency of other anatomical elements affecting large scale patterns in conduit distribution (such as parenchyma rays) show a lower phylogenetic dependency.

We note, however, that the values of λ presented here refer only to the effects of the phylogeny on spatial variables analyzed individually. As previously mentioned, spatial variables need to be interpreted in the context of simultaneous changes in both vessel density (cf., Fig. 5) and vessel size. Hence, λ also needs to be calculated in the context of the relationship between all spatial variables involved. This will be carried out separately.

The observation that the three deepest nodes showed very divergent spatial indices (from highly aggregated to highly uniform) suggests that these two syndromes were already present early on in the evolution of the core eudicot clade. In addition, our parsimony reconstruction clearly showed that accentuated conduit uniformity and accentuated aggregation must have evolved independently several times, with highly uniform distributions being now present in all the major core eudicot clades and with aggregation also present in several of them.

Physiological importance of xylem spatial patterns—Does the species-to-species variability in xylem spatial conduit distribution affect the physiological properties of their water transport system? For example, should one expect that a xylem structure characterized by a few aggregated conduits has a higher or a lower transport efficiency (as measured by hydraulic conductivity), or a higher or a lower transport safety (as measured by vulnerability to water-stressed induced embolism), compared to a xylem structure with many uniformly distributed conduits? These questions have not been examined empirically, by direct comparison of species selected along a gradient of conduit aggregation, nor have they been examined theoretically, using spatially explicit models of water transport. The available anatomical and physiological evidence appears to suggest that a syndrome of many small aggregated conduits is typically seen in drought-prone regions, which would imply an ecological advantage of conduit aggregation in resisting drought stress. Carlquist (1984) proposed that this advantage likely consists of providing parallel pathways of water transport, which may be required during periods of partial embolization of the transport pathway. This view was reinforced by the abundant presence of "true" tracheids (with fully functional pits) in many species with isolated conduits, and the absence of "true" tracheids in species with grouped conduits. However, it is not clear that a large local connectivity within the single patch of aggregated conduits necessarily provides redundancy at the level of the whole cross section, given that the various patches are necessarily more isolated. Additionally, the argument that water is more easily rerouted in a group of conduits if a component of the group cavitates does not consider that water and air are rerouted following the same routes and, by the same token, emboli are also more easily rerouted. Finally, as previously remarked, the Carlquist index is hard to interpret for across-species comparisons, because the number of contacts occurring among conduits purely as a result of their random placement in a cross section will depend on conduit size as well as density.

The spatially explicit model presented by Loepfe et al. (2007) provides an opportunity to test these ideas, by linking it with an explicit three-dimensional representation of the conduit distribution. The results presented by Loepfe et al. (2007) would suggest that a trend opposite to the one proposed by Carlquist (1984) is more likely, i.e., that the increased local connectivity caused by the aggregation of conduits would likely enhance the local conductivity of the patch while increasing of spreading of run-away embolism under drought conditions. An explicit test of these hypotheses seems necessary.

Finally, it can be asked why we focused on the determination of a relatively abstract set of statistical parameters and not instead on parameters such as average wall area of contact between two conduits or fraction of vessel length in contact with another vessel. These are all important parameters that need to be measured to identify empirical patterns in the physiological determinants of the degree of xylem connectivity and to parameterize models of water transport. However, we suspect that

those patterns will reflect the large-scale spatial distribution of conduits in a cross section and hope that this statistical approach will prove useful in understanding the broader scale patterns.

Further opportunities for spatial analysis applied to functional xylem anatomy—Beyond the analyses presented in this paper, several additional avenues are available to expand our knowledge of xylem structure and functioning using more advanced statistical tools. Firstly, PPA theory is already available for images in 3D as opposed to 2D. This approach should be particularly promising if 3D binary images could be directly fed from, e.g., high power x-ray ultrascan analyses (e.g., Steppe et al., 2004). In addition, space–time PPA is also a promising avenue, e.g., to analyze the spatial and temporal patterns of xylem embolism in relation to the pre-existing presence of spatial structure in the population of conduits (e.g., Gatrell et al., 1996; Baddeley et al., 2005). This should conclusively demonstrate whether, and under what conditions, the spread of embolism is definitely a neighbor-to-neighbor process. In more general terms, beside the identification and the modeling of point patterns, additional techniques are also available to detect the presence of linear or curvilinear structures in a matrix of noisy points (Stanford and Raftery, 2000). This should be particularly useful to determine objectively whether radial, tangential or diagonal vessel multiples can be statistically identified in cross sections.

Conclusions—We demonstrated that xylem conduit spatial distribution can be determined using point pattern analysis, a statistical technique that evaluates conduit spatial location against the null hypothesis of complete spatial randomness. We also modeled the spatial properties of xylem cross sections belonging to 105 temperate tree species, using a piecewise saturated Geyer model. We used these analyses to show that two indices of conduit spatial distribution exhibited a strong phylogenetic signal. In addition, we showed that these indices of spatial distribution were significantly related to conduit density, with aggregated conduits being associated with lower conduit densities. This suggests an important physiological and ecological role for conduit aggregation and, its opposite, uniformity.

LITERATURE CITED

BAAS, P., AND S. CARLQUIST. 1985. A comparison of the ecological wood anatomy of the floras of southern California and Israel. *International Association of Wood Anatomists Bulletin* 6: 349–353.

BADDELEY, A. 1998. Spatial sampling and censoring. *In* O. Barndorff-Nielsen, W. Kendall, and M. van Lieshout [eds.], Stochastic geometry: Likelihood and computation, Monographs on statistics and applied probability, no. 80, 37–78. Chapman and Hall, London, UK.

BADDELEY, A., AND T. R. TURNER. 2005. Spatstat: An R package for analysing spatial point patterns. *Journal of Statistical Software* 12: 1–42.

BADDELEY, A., T. R. TURNER, J. MØLLER, AND M. HAZELTON. 2005. Residual analysis for spatial point processes. *Journal of the Royal Statistical Society, B. Methodological* 67: 617–666.

BERMAN, M., AND T. R. TURNER. 1992. Approximating point process likelihoods with GLIM. *Applied Statistics* 41: 31–38.

CAMPBELL, C. S., R. C. EVANS, D. R. MORGAN, T. A. DICKINSON, AND M. P. ARSENAULT. 2007. Phylogeny of subtribe *Pyrinae* (formerly the Maloideae, Rosaceae): Limited resolution of a complex evolutionary history. *Plant Systematics and Evolution* 266: 119–145.

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

CARLQUIST, S. 2001. Comparative wood anatomy. Systematic, ecological and evolutionary aspects of dicotyledon wood, 2nd ed. Springer, Berlin, Germany.

CARLQUIST, S., AND D. A. HOEKMAN. 1985. Ecological wood anatomy of the woody southern Californian flora. *International Association of Wood Anatomists Bulletin* 6: 319–347.

CHEN, Z. 1999. Phylogeny and evolution of the Betulaceae as inferred from DNA sequences, morphology, and paleobotany. *Plant Systematics and Evolution* 86: 1168–1181.

CHOAT, B., T. W. BRODIE, A. R. COBB, M. A. ZWIENIECKI, AND N. M. HOLBROOK. 2006. Direct measurements of intervessel pit membrane hydraulic resistance in two angiosperm tree species. *American Journal of Botany* 93: 993–1000.

CHOAT, B., A. R. COBB, AND S. JANSEN. 2008. Structure and function of bordered pits: New discoveries and impacts on whole-plant hydraulic function. *New Phytologist* 177: 608–626.

CHRISTMAN, M. A., J. S. SPERRY, AND F. R. ALDER. 2009. Testing the 'rare pit' hypothesis for xylem cavitation resistance in three species of *Acer*. *New Phytologist* 182: 664–674.

DAVIES, T. J., T. G. BARRACLOUGH, M. W. CHASE, P. S. SOLTIS, D. E. SOLTIS, AND V. SAVOLAINEN. 2004. Darwin's abominable mystery: Insights from a supertree of the angiosperms. *Proceedings of the National Academy of Sciences, USA* 101: 1904–1909.

DIGGLE, P. 2003. Statistical analysis of spatial point patterns, 2nd ed. Arnold, London, UK.

FRECKLETON, R. P., P. H. HARVEY, AND M. PAGE. 2002. Phylogenetic analysis and comparative data: A test and review of evidence. *American Naturalist* 160: 712–726.

GATRELL, A. C., T. C. BAILEY, P. J. DIGGLE, AND B. S. ROWLINGSON. 1996. Spatial point pattern analysis and its application in geographical epidemiology. *Transactions of the Institute of British Geographers* 21: 256–274.

GERNANDT, D. S., G. G. LÓPEZ, S. O. GARCÍA, AND A. LISTON. 2005. Phylogeny and classification of *Pinus*. *Taxon* 54: 29–42.

GEYER, C. 1999. Likelihood inference for spatial point processes. *In* O. Barndorff-Nielsen, W. Kendall, and M. van Lieshout [eds.], Stochastic geometry: Likelihood and computation, Monographs on statistics and applied probability, no. 80, 79–140. Chapman and Hall, London, UK.

GIL, D., A. HERNANDEZ, O. RODRÍGUEZ, F. MAURI, AND P. RADEVA. 2006. Statistical strategy for anisotropic adventitia modelling in IVUS. *IEEE Transactions on Medical Imaging* 27: 1022–1030.

GIL, D., A. HERNANDEZ-SABATE, M. BURNAT, S. JANSEN, AND J. MARTINEZ-VILLALTA. 2009. Structure-preserving smoothing of biomedical images. Proceedings 13th International Conference on Analysis and Image Processing, Lecture Notes in Computer Science 5702, 427–434, 2009, Münster, Germany. Springer, Berlin, Germany, ISBN 978-3-642-03766-5.

HAMZEH, M., AND S. DAYANANDAN. 2004. Phylogeny of *Populus* (Salicaceae) based on nucleotide sequences of chloroplast *trnT-trnF* region and nuclear rDNA. *American Journal of Botany* 91: 1398–1408.

HERENDEEN, P. S., E. A. WHEELER, AND P. BAAS. 1999. Angiosperm wood evolution and the potential contribution of paleontological data. *Botanical Review* 65: 278–300.

IAWA [INTERNATIONAL ASSOCIATION OF WOOD ANATOMISTS] COMMITTEE. 1989. IAWA list of microscopic features for hardwood identification. *International Association of Wood Anatomists Bulletin* 10: 219–332.

ILLIAN, J., A. PENTTINEN, H. STOYAN, AND D. STOYAN. 2008. Statistical analysis and modelling of spatial point patterns. John Wiley, Chichester, UK.

JANSEN, S., B. CHOAT, AND A. PLETSERS. 2009. Morphological variation of intervessel pit membranes and implications to xylem function in angiosperms. *American Journal of Botany* 96: 409–419.

JÄRVINEN, P., A. PALMÉ, L. O. MORALES, M. LÄNNENPÄÄ, M. KEINÄNEN, T. SOPANEN, AND M. LASCOUX. 2004. Phylogenetic relationships of *Betula* species (Betulaceae) based on nuclear *ADH* and chloroplast *matK* sequences. *American Journal of Botany* 91: 1834–1845.

KOOPMAN, W. J. M., V. WISSEMANN, K. DE COCK, J. VAN HUYLENBROECK, J. DE RIEK, G. J. H. SABATINO, D. VISSER, ET AL. 2008. AFLP markers as a tool to reconstruct complex relationships: a case study in *Rosa* (Rosaceae). *American Journal of Botany* 95: 353–366.

KREBS, C. J. 1999. Ecological methodology. Benjamin Cummings/ Prentice Hall, Menlo Park, California, USA.

LI, J., J. YUE, AND S. SHOUP. 2006. Phylogenetics of *Acer* (Aceroideae, Sapindaceae) based on nucleotide sequences of two chloroplast noncoding regions. *Harvard Papers in Botany* 11: 101–115.

LOEPFE, L., J. MARTINEZ-VILALTA, J. PIÑOL, AND M. MENCUCCINI. 2007. The relevance of xylem network structure for plant hydraulic efficiency and safety. *Journal of Theoretical Biology* 247: 788–803.

LYNCH, H. J., AND P. R. MOORCROFT. 2008. A spatiotemporal Ripley's K-function to analyze interactions between spruce budworm and fire in British Columbia, Canada. *Canadian Journal of Forest Research* 38: 3112–3119.

MØLLER, J., AND R. WAAGEPETERSEN. 2003. Statistical inference and simulation for spatial point processes. Chapman and Hall/CRC, Boca Raton, Louisiana, USA.

NAVARRO, E., J. BOUSQUET, A. MOIROUD, A. MUNIVE, D. PIOU, AND P. NORMAND. 2003. Molecular phylogeny of *Alnus* (Betulaceae), inferred from nuclear ribosomal DNA ITS sequences. *Plant and Soil* 254: 207–217.

PAGEL, M. 1999. Inferring the historical patterns of biological evolution. *Nature* 401: 877–884.

PARADIS, E. 2006. Analysis of phylogenetics and evolution with R. Springer, Berlin, Germany.

POTTER, D. 2007. Phylogeny and classification of Rosaceae. *Plant Systematics and Evolution* 266: 5–43.

RIPLEY, B. 1981. Spatial statistics. John Wiley, New York, New York, USA.

RIPLEY, B. 1988. Statistical inference for spatial processes. Cambridge University Press, Cambridge, UK.

ROSELL, J. A., M. E. OLSON, R. AGUIRRE-HERNANDEZ, AND S. CARLQUIST. 2007. Logistic regression in comparative wood anatomy: Tracheid types, wood anatomical terminology, and new inferences from the Carlquist and Hoekman southern Californian data set. *Botanical Journal of the Linnean Society* 154: 331–351.

RYDIN, C. 2002. Seed plant relationships and the systematic potion of gnetales based on nuclear and chloroplast DNA: Conflicting data, rooting problems, and the monophyly of conifers. *International Journal of Plant Sciences* 163: 197–214.

SCHENK, H. J., S. ESPINO, C. M. GOEDHART, M. NORDENSTAHL, H. I. MARTINEZ CABRERA, AND C. S. JONES. 2008. Hydraulic integration and shrub growth form linked across continental aridity gradients. *Proceedings of the National Academy of Sciences, USA* 105: 11248–11253.

SCHOCH, W., I. HELLER, F. H. SCHWEINGRUBER, AND F. KIENAST. 2004. Wood anatomy of central european species. Online version: http://www.woodanatomy.ch [accessed 1 September 2009].

SCHWEINGRUBER, F. H. 1990. Microscopic wood anatomy: Structural variability of stems and twigs ion recent and subfossil woods from Central Europe, 3rd ed. Eidgen-ssische Forschungsanstalt [Swiss Federal Research Institute] WSL, Birmensdorf, Switzerland.

STANFORD, D. C., AND A. E. RAFTERY. 2000. Finding curvilinear features in spatial point patterns: Principal curve clustering with noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22: 601–609.

STEPPE, K., V. CNUDDE, C. GIRARD, R. LEMEUR, J.-P. CNUDDE, AND P. JACOBS. 2004. Use of X-ray computed microtomography for noninvasive determination of wood anatomical characteristics. *Journal of Structural Biology* 148: 11–21.

TYREE, M. T., S. D. DAVIS, AND H. COCHARD. 1994. Biophysical perspectives of xylem evolution: is there a tradeoff of hydraulic efficiency for vulnerability to dysfunction? *International Association of Wood Anatomists Bullettin* 15: 335–360.

WANG, X.-Q., D. C. TANK, AND T. SANG. 2000. Phylogeny and divergence times in Pinaceae: Evidence from three genomes. *Molecular Biology and Evolution* 17: 773–781.

WEBB, C. O., AND M. J. DONOGHUE. 2005. Phylomatic: Tree assembly for applied phylogenetics. *Molecular Ecology Notes* 5: 181–183.

WEN, J., S. T. BERGGREN, C.-H. LEE, S. ICKERT-BOND, T.-S. YI, K.-O. YOO, L. XIE, ET AL. 2008. Phylogenetic inferences in *Prunus* (Rosaceae) using chloroplast *ndhF* and nuclear ribosomal ITS sequences. *Journal of Systematics and Evolution* 46: 322–332.

WHEELER, E. A., P. BAAS, AND S. RODGERS. 2007. Variations in dicot wood anatomy: A global analysis based on the InsideWood database. *International Association of Wood Anatomists Journal* 28: 229–258.

WHEELER, J. K., J. S. SPERRY, U. G. HACKE, AND N. HOANG. 2005. Intervessel pitting and cavitation in woody Rosaceae and other vesselled plants: A basis for a safety versus efficiency trade-off in xylem transport. *Plant, Cell & Environment* 28: 800–812.

WIEGAND, T., W. D. KISSLING, P. A. CIPRIOTTI, AND M. A. AGUIAR. 2006. Extending point pattern analysis for objects of finite size and irregular shape. *Journal of Ecology* 94: 825–837.

WIEGAND, T., AND K. A. MOLONEY. 2004. Rings, circles, and null-models for point pattern analysis in ecology. *Oikos* 104: 209–229.

ZANNE, A. E., K. SWEENEY, M. SHARMA, AND C. M. ORIANS. 2006. Patterns and consequences of differential vascular sectoriality in 18 temperate tree and shrub species. *Functional Ecology* 20: 200–206.

ZWIENIECKI, M. A., P. J. MELCHER, AND M. N. M. HOLBROOK. 2001. Hydrogel control of xylem hydraulic resistance in plants. *Science* 291: 1059–1062.