**Title**

Linking functional traits to multiscale statistics of leaf venation networks

**Authors**

Benjamin Blonder [1, 2, *] - ORCID 0000-0002-5061-2385

Sabine Both [3,4] – ORCID 0000-0003-4437-5106

Miguel Jodra [1]

Hao Xu [5]

Mark Fricker [6] – ORCID 0000-0002-8942-6897

Ilaíne S. Matos [2] – ORCID 0000-0001-5557-5133

Noreen Majalap [7]

David F.R.P. Burslem [3] – ORCID 0000-0001-6033-0990

Yit Arn Teh [8] – ORCID 0000-0001-7976-6794

Yadvinder Malhi [1] – ORCID 0000-0002-3503-4783

**Addresses**

1: Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, United Kingdom

2: Department of Environmental Science, Policy, and Management, University of California, Berkeley, California, United States

3: School of Biological Sciences, University of Aberdeen, Aberdeen, United Kingdom

4: School of Environmental and Rural Science, University of New England, Armidale, New South Wales, Australia

5: Institute of Biomedical Engineering, Department of Engineering Science, Old Road Campus Research Building, University of Oxford, Oxford, United Kingdom

6: Department of Plant Sciences, University of Oxford, Oxford, United Kingdom

7: Forest Research Centre, Sabah Forestry Department, Sandakan, Sabah, Malaysia

8: School of Natural and Environmental Sciences, University of Newcastle, Newcastle, United Kingdom

*: Corresponding author, benjamin.blonder@berkeley.edu, +1 510 664 5339

**Social media information:**

Twitter: @benjaminblonder

**Word counts**

Summary: 195

Total: 6564

Introduction: 2156

Materials and Methods: 1457

Results: 1176

Discussion: 1775

**Figures**

Number of figures: 7

Color figures: 3,4,5,6,7

**Summary**

(1) Leaf venation networks evolved along several functional axes, including resource transport, damage resistance, mechanical strength, and construction cost. Because functions may depend on architectural features at different scales, network architecture may vary across spatial scales to satisfy functional tradeoffs.

(2) We develop a framework for quantifying network architecture with multiscale statistics describing elongation ratios, circularity ratios, vein density, and minimum spanning tree ratios. We quantify vein networks for leaves of 260 southeast Asian tree species in samples of up to 2 $cm^2$, pairing multiscale statistics with traits representing axes of resource transport, damage resistance, mechanical strength, and cost.

(3) We show that these multiscale statistics clearly differentiate species' architecture and delineate a phenotype space that shifts at larger scales; functional linkages vary with scale and are weak, with vein density, minimum spanning tree ratio, and circularity ratio linked to mechanical strength (measured by force to punch) and elongation ratio and circularity ratio linked to damage resistance (measured by tannins); and phylogenetic conservatism of network architecture is low but scale-dependent.

(4) This work provides tools to quantify the function and evolution of venation networks. Future studies including primary and secondary veins may uncover additional insights.

**Key words**

Venation network, network architecture, functional trait, resource transport, damage resistance, mechanical strength, construction cost, leaf

## Introduction

Leaves have venation networks with architecture that varies widely, from a single vascular strand (e.g. pines) to purely branching structures (e.g. *Ginkgo*) to open net patterns (e.g. many ferns) to mostly parallel structures in monocots, to highly reticulate patterns in many angiosperms. Networks vary over multiple spatial scales, with several levels of branching at length scales from $10^{-5}$ m (the radius of a single vein) – $10^{0}$ m (the length of some large leaves) (Roth-Nebelsick *et al.*, 2001; Sack & Scoffoni, 2013). We define *spatial scale* as a characteristic feature of the network with a certain extent, e.g. the area of a vein loop, or the radius of a vein.

Network architecture may be closely linked to multiple functions (Ronellenfitsch & Katifori, 2019). Selection could act to maximize efficiency of resource transport, resistance to damage, mechanical strength, or to minimize total construction cost (Blonder *et al.*, 2018). For resource transport, veins are implicated in photosynthesis and transpiration, as vein-mediated resource transport comprises a large portion of total leaf conductance in most species (Brodribb *et al.*, 2007), though variation in other tissues (e.g. bundle sheath extensions) may also be important (Ohtsuka *et al.*, 2018). For resistance to damage, the presence of loops in the network (Katifori *et al.*, 2010) could prevent the propagation of embolisms that reduce conductance under low water potential (Brodribb *et al.*, 2016), as well as the propagation of tears or cracks (Vincent, 1982; Niklas, 1999) caused by wind or herbivores. Additionally, defensive secondary compounds (e.g. latex) transported by the network could promote further resistance against herbivory (Agrawal & Konno, 2009), as damaged leaves with redundant flow pathways could still deliver latex or other signaling molecules to other portions of the leaf. For mechanical strength, the network could provide a skeleton that allows leaves to remain upright in wind, ultimately supporting photosynthesis (Givnish, 1979; Givnish *et al.*, 2005). For construction cost, lignified tissue comprising veins is costly to construct relative to other tissues (John *et al.*, 2017), and may also displace leaf volume that could be allocated to photosynthetic functioning. Thus, the realized network architecture may reflect trade-offs among these different functions.

These functional tradeoffs and network architecture may both contribute to function at a range of spatial scales within a leaf. Thus, network features of different spatial scales may make varying contributions to functioning. For example, minor veins contribute more to conductance and photosynthesis due to their close spacing and disproportionate impact on hydraulic resistance (Brodribb *et al.*, 2007; McKown *et al.*, 2010). Conversely, major veins contribute more to cost

(John *et al.*, 2017) and mechanical strength (Hua *et al.*, 2019) due to their disproportionate mass allocation.

From an evolutionary perspective, some aspects of network architecture may be more labile than others. Larger-scale veins could be phylogenetically conserved over deeper time (Ellis *et al.*, 2009). Smaller-scale vein patterns are thought to evolve more quickly and show developmental plasticity (Sack & Scoffoni, 2013).

At a single scale (e.g. of minor veins, typically ~10 μm radius), network architecture can be described using a range of geometrical or topological statistics (**Table 1**). Single-scale studies of network architecture and leaf function have sometimes produced clear functional linkages, e.g. between vein density and resource transport, or between vein branching and defense chemistry (Blonder *et al.*, 2011; Blonder *et al.*, 2016; Blonder *et al.*, 2017; Hua *et al.*, 2019), but in other contexts have not, e.g. between vein density and traits associated with cost, like leaf mass per area (Li *et al.*, 2015). Recent models linking network architecture to leaf function (Blonder *et al.*, 2011) have neglected this multiscale variation.

If network function varies across spatial scales, then network architecture should also be described across spatial scales, to better understand the diversity of evolved forms and the rules underlying this multiscale architecture. While such a perspective is sometimes implicit (Sack *et al.*, 2012; Sack & Scoffoni, 2013; Hua *et al.*, 2019; Kawai & Okada, 2019), prior studies have used categorical descriptors to characterize features at other spatial scales (Sack & Frole, 2006; Sack *et al.*, 2008; Ellis *et al.*, 2009).

There is a need for multiscale description of venation network architecture, as well as for comprehensive assessment of its implications for multiple leaf functions. Recent efforts have focused on the idea of hierarchical loop decomposition (HLD) (Katifori & Magnasco, 2012; Mileyko *et al.*, 2012; Ronellenfitsch *et al.*, 2015). In this approach, a leaf venation network is considered as comprising a set of loops that are nested within each other. The leaf can be partially summarized via a hierarchical tree describing how smaller loops are nested within larger loops, and the statistical properties of the tree (e.g. nesting ratio, topological length) can be used to characterize variation between species. This approach provides an advance in terms of explicitly considering architecture across spatial scales. However, empirical applications of this framework have been constrained by limited venation data and by the unclear linkage between

the hierarchical tree properties and hypotheses about leaf functioning (Ronellenfitsch *et al.*, 2015).

**Framework for multiscale venation statistics**

Here we present a framework, built on HLD concepts, to define multiscale venation statistics. The goal of the framework is to describe scale transitions in network architecture statistics (Table 1), so that a set of single-scale statistics $\{\omega_i\}$, can be transformed into a set of continuous functions $\{\omega_i(r)\}$ , where $r$ is a metric of spatial scale.

HLD analyses are based on 'spatial graph' extractions of leaf venation networks, i.e. with all vein segments described in terms of their position, length, radius, and connectivity with other vein segments, and all areoles described in terms of their area and shape. The network extraction step can be carried out using software available in Xu *et al.* (2020). This information is necessary to determine which areoles are nested within which other areoles. Loops are defined either in terms of regions wholly enclosed by veins, or regions enclosed by veins and leaf boundary; regions not wholly enclosed by veins (e.g. due to cropping of an input image) are excluded from the analysis. Vein segments that constitute boundaries between loops are identified and described in terms of their radius ($r$). Then, a pair of loops with the minimum radius boundary segment ($r_{min}$) are fused together, and the boundary segment is deleted from the network. This process begins with fusing the areoles linked by the vein with the smallest $r_{min}$, and is iterated until all loops are fused together into one large loop representing the entire leaf, with $r_{min}$ equal to the width of the largest vein segment (**Figure 1**). Only loops that are complete are considered (i.e. excluding those cut off by the image boundary). Large veins are split into multiple segments spanning branching points with other segments.

Single-scale statistics $\{\omega_i\}$ are then calculated at each of the network fusion steps, where $i$ indexes the various statistics. If a network has $n$ boundary vein segments, then the hierarchical loop decomposition will yield $n$ values of $\{\omega_i\}$, each at a different value of $r$. The outcome of this process is a scale-dependent description of the network that remains when considering only veins above a certain size class.

These $\{\omega_i\}$ values can then be converted into multiscale statistics, $\{\omega_i(r)\}$, which encapsulate the scale transitions that may occur in network architecture. Importantly, not all leaves will have veins present at all scales. Thus, to obtain scale-dependent functions, the first

step is to construct a hurdle function describing the presence or absence of a vein at a given scale, $\{\eta_i(r)\}$. Next, if veins are present, i.e. $\{\eta_i(r)\} \neq 0$, a 'multiscale statistic' $\{\sigma_i(r)\}$ is constructed describing the statistic of interest, e.g. VD(r), which is the vein density across scales. The final function is:

$$\omega_i(r) = \begin{cases} \eta_i(r) = 1, & \sigma_i(r) \\ \eta_i(r) = 0, & NA \end{cases}$$

Estimation proceeds by binning the data into uniform ranges of *r*, then calculating bin-median values of $\eta_i$ and $\sigma_i$. Using raw data leads to undesirable sample size – network scale correlations, while fitting functional forms to the curve via regression is not necessarily appropriate because the $\omega_i$ values are piecewise continuous rather than fully continuous, i.e. include gaps where no features of a given size class exist. Vein absences can arise either because a vein is truly absent at a scale, or because the sampled area was too small to be representative (i.e. undersampling). If the input network is complete (i.e. comprising a whole leaf) then the absences are real, because the input data are a statistical population rather than a statistical sample. In leaf subsections, however, undersampling biases may be important. In these cases, one can assume that network architecture falls on a scale continuum over the scale range of the data, i.e. $\eta_i(r) = 1$ for all *r*, and then $\sigma_i(r)$ can be gap-filled using interpolation.

Sampling limitations can also increase the frequency of gaps and/or bias estimates. The sampling uncertainty in $\omega_i$ necessarily increases with increases in *r*. This is because the number of network components must decrease as portions of the network are deleted. Therefore, analyses should be restricted to ranges of *r* where more than a certain fraction of samples have $\eta_i(r) = 1$.

Once the multiscale statistics are estimated, they can be used to describe variation among leaves and species, and to serve as potential predictors of leaf function. To proceed, one can use the values directly, or can reduce them to scalars like the linear slope estimate of $\omega_i(r)$ vs. *r*. When using the predicted values of $\{\omega_i(r)\}$ in regression analyses, it is necessary to account for the non-independence of values at similar spatial scales, e.g. via partial least squares (PLS) methods.

**Hypotheses about trade-offs between network architecture and function**

We focused on traits related to four multiscale statistics: vein density (VD), the mean elongation ratio of loops (ER), loop circularity (CR), and the minimum spanning tree ratio (MST) (**Table 1**).

We also focus on linking them to four leaf functions (resource transport, damage resistance, mechanical strength, and construction cost) (**Table 2**).

VD, ER, and MST have been shown to capture leading axes of variation in network architecture across many species at single scales (Blonder *et al.*, 2018), while CR(r) may reflect relative allocation of vein perimeter relative to loop area (Blonder *et al.*, 2011). Based on the availability of trait data in this study, we represented resource transport by maximum photosynthetic capacity ($A_{max}$ in common usage), though recognizing that this metric of carbon flux does not fully represent water fluxes, i.e. maximum hydraulic conductance (Brodribb *et al.*, 2007), for which we did not have data. Damage resistance was represented by tannin fraction, as this type of secondary chemistry is commonly used as an herbivory defense (Coley & Barone, 1996). To represent mechanical strength we used force to punch, as this is a direct metric of strength (Pérez-Harguindeguy *et al.*, 2013). Finally, we represented cost by leaf mass per area, as dry mass investment is strongly related to tissue construction cost (Poorter *et al.*, 2009).

Specifically, we asked 1) whether multiscale statistics enable useful measurement of species network architecture, 2) whether network architectures vary systematically across spatial scales, 3) whether expected architecture-function trade-offs are supported across different spatial scales, and 4) whether the phylogenetic conservatism of architecture varies across spatial scales. Then, we hypothesized that:

1. the four statistics of network architecture will represent statistically independent axes of variation, such that a wide range of network architectures are possible.
2. different portions of the architectural space will be occupied at different scales. That is, network architecture will shift systematically as scale increases.
3. distinct trade-offs exist between network architecture and functioning across spatial scales (**Table 2**). We predict that resource transport will be positively linked to VD(r) at small values of *r*, due to the importance of small veins in determining hydraulic conductance (McKown *et al.*, 2010), as well as, positively linked to MST(r) across scales, as more tree-like networks distribute resources better (Katifori *et al.*, 2010). Damage resistance will be negatively linked to MST(r) and CR(r) across scales, as high secondary chemistry investment may offset investment in redundant flow pathways. Mechanical strength will be positively linked to VD(r) at large scales, as large veins may contribute disproportionately to stiffness and because lignified vessel

sclerenchyma are associated with fracture resistance (Choong *et al.*, 1992), and negatively linked to ER(r) at small scales, as more parallel-veined laminas may be easier to tear. Cost will be positively linked to VD(r) at large scales, due to the importance of large veins (which are lignified and dense) in determining total leaf cost (John *et al.*, 2017).

4. phylogenetic niche conservatism in all network multiscale statistics will be lowest at small values of *r* due to their presumed development and evolutionary lability (Trivett & Pigg, 1996; Roth-Nebelsick *et al.*, 2001; Ellis *et al.*, 2009).

We tested these hypotheses using a phylogenetically diverse set of 260 tree species in a forest in Malaysian Borneo. Our analysis was based on leaf samples of ~ 2 $cm^2$, which contained veins of several orders but typically did not include primary or secondary veins.

## Materials and Methods

This study integrates venation network imagery (Blonder *et al.*, 2019) with functional trait data that were collected from the study sites (**Table 3**) (Both *et al.*, 2018).

### *Sites and sampling design*

Samples were collected in Malaysian Borneo during July-December 2015, within eight 1-ha permanent forest plots comprising mixed dipterocarp lowland forest (**Table 3**). We pooled data from all sites, as this study focused on functional and evolutionary rather than environmental questions. Vouchers are stored at the Danum Valley Field Centre, Sabah, Malaysia.

Within each plot, all trees ≥ 10 cm in diameter at breast height (DBH) were identified and measured for DBH and height. Branches of 2-4 cm diameter were then collected using rope climbing or pole pruning techniques. Branches were collected from species with highest contribution to biomass, i.e. species comprising the top 70% of plot basal area. Additional leaves of rarer taxa were also collected from all trees in three 20 × 20 m subplots within each plot. When possible, both sunlit and shaded branches were collected from each tree, resulting in sampling heights across the range of 2 – 53 m.

### *Venation networks*

A single mature undamaged leaf was collected from each branch, cleaned with a wet rag, then pressed flat and dried at 60°C for several days. A 1-4 $cm^2$ section was then cut from each leaf midway from the base to the apex and midway from a primary vein to either another primary vein (for palmate leaves) or to the margin (for pinnate leaves). These dried sections were rehydrated to reverse the effects of sample shrinkage (Blonder *et al.*, 2012) and chemically cleared and stained (Pérez-Harguindeguy *et al.*, 2013; Blonder *et al.*, 2018).

Samples were mounted on glass slides and imaged using a compound microscope (Olympus, BX43) with 2x apochromat objective and a color camera (3840 × 2748 pixel resolution; Olympus, SC100). Approximately 16 image fields were then stitched together to obtain a full image of each sample (resolution 595 px $mm^{-1}$). Image contrast was enhanced using a contrast-limited adaptive histogram equalization on the green channel of each image. Images were then imported into GIMP (GNU Project) image-editing software and the boundaries of the sample (vs. background) were manually delineated with a polygon. Next, an approximately 700×700 $px^2$ region in each sample was manually segmented using a tracing tablet. The hand-traced vein images were then segmented into binary representations, in which vein pixels were given one value, non-vein pixels another value, and image background a NA value.

These validated data (comprising 686,881,432 manually segmented pixels) were used to train a machine learning algorithm, as described in Xu *et al.* (2020). An ensemble of 6 convolutional neural networks (CNNs) was developed, with each network implemented via a U-Net architecture (Ronneberger *et al.*, 2015). To avoid overfitting, input image data were repeatedly rotated and scaled during training. Each CNN was trained on 5/6 of the input dataset and predicted using a sliding window approach on the remaining 1/6. Probability outputs were then averaged across all CNNs in the ensemble and converted to binary predictions using a threshold determined by Receiver Operating Characteristic (ROC) plots based on the prediction relative to the manually delineated validation region in each image. The algorithms sometimes under-segmented large veins that were not included in the validation dataset. Any veins with widths greater than approximately 500 μm were also manually segmented throughout the entire sample. These segmentations were then added to the CNN-based segmentations. Segmentations were then masked to the manually delineated boundaries. A total of 32,815,701,653 pixels were algorithmically segmented using this methodology.

The segmented images were then filtered, so that images with sample areas less than 90 $mm^2$ and Fbeta2 scores lower than 0.8 were discarded. Among the retained images, the segmentation quality was high (Fbeta2 mean=0.95 ± 0.03 s.d.). A small number of images were assigned new branch codes different than those in the original image dataset, based on recent revision of handwritten sample labels (**Table S1**).

Spatial graph representations of the networks were extracted from each segmented image. Next, all unique loops and their boundary vein segments were identified. Unclosed loops (on the boundary of the sample) were removed from the analysis. The four architectural network statistics (**Table 1**) then were obtained following Xu *et al.* (2020).

After calculating these statistics, the network was iteratively pruned, and loops fused, by sequential removal of the smallest remaining boundary vein segments. After calculating these statistics, the network was iteratively pruned, and loops fused, by sequential removal of the smallest intervening vein segment. After each fusion event, vein and areole metrics were re-calculated, along with the minimum vein radius ($r_{min}$, μm). Multiscale statistics were truncated to the 0 – 0.2 mm range of $r_{min}$ to minimize undersampling biases (**Supporting Information Figure S1**), though veins of up to 0.58 mm were present in the data.

*Trait measurements*

Traits were measured from each branch using mature undamaged and cleaned leaves. Different leaves were used for each trait, then data were pooled at branch level. Measurements are described in **Table 2** and in Both *et al.* (2018).

*Statistical analysis*

We first constructed multiscale statistics for each leaf by binning data into 50 $r_{min}$ bins spanning 0 – 0.2 mm. As we worked with leaf subsections, the absence of veins at a certain scale could be caused either by true absences or by undersampling. In cases where $\eta_i(r) = 0$ occurred, we used two approaches. In the first approach, we assumed undersampling, then set $\eta_i(r) = 1$ and filled the missing values of $\sigma_i(r)$ using linear interpolation. Extrapolation was handled by using the value at the closest data extreme. In the second approach, we assumed the data were accurate as observed, then separately modeled the $\eta$ and $\sigma$ components of $\omega$. For all the analyses described below, except the test of Hypothesis 3, we only used the first approach, as the large number of

incomplete cases prevented application of the necessary statistics. Analyses were conducted treating branches as independent and identically distributed random replicates.

To determine the independence of architectural axes across scales (Hypothesis 1), we carried out a principal component analysis across the four axes of VD, ER, CR, and MST, using values for each leaf at each scale as replicates. We then assessed variance allocated to each principal component to determine the relative independence of axes.

To evaluate whether different portions of the architectural space are occupied at different scales (Hypothesis 2), we also visualized the principal component scores using convex hulls and 95% confidence ellipses at each scale.

To assess evidence for architecture-function trade-offs across scales (Hypothesis 3), we used two complementary approaches. In the first approach, we carried out partial least squares (PLS) regressions for each of the functional traits. Functional traits were treated as response variables and the multiscale statistics in each bin as predictors. PLS allows for incorporation of non-independence of the statistics across scales. All four multiscale statistics were included as predictors. Each multiscale statistic was *z*-transformed before analysis to improve comparability of parameter estimates across multiscale statistics. For all models, most of the variation was explained by the first component, so we restricted analysis to this component. We reported the variance explained by the first component, and also assessed the scale-dependent effect of each venation trait at each scale as the value of the respective loading coefficient. Each hypothesis was then tested by examining plots of loading coefficients vs. $r_{min}$.

In the second approach, we used ordinary least squares regression models, one for each functional trait, multiscale statistic, and $r_{min}$ binned value combination. At each combination, we pooled data for all species with $\eta_i(r) = 1$ (i.e. only those species with observed veins at that scale). We then conducted the regression using the functional trait values as the response variable and the $\sigma_i(r)$ values as the predictor variable. We reported the overall variance explained by the regression and the slope estimates. This approach does not account for non-independence of data across spatial scales, or non-independence of predictors, nor does it allow us to assess statistical significance (due to the high false-discovery rate). However, it does allow for use of more data in each analysis, as missing cases do not need to be dropped or interpolated.

To determine levels of phylogenetic niche conservatism across scales (Hypothesis 4), we used the binned values of each multiscale statistic. For each spatial scale and each multiscale

statistic, we computed the K statistic (Blomberg *et al.*, 2003). Values of K<1 indicate faster trait evolution than expected under a Brownian model (trait lability), while values of K>1 indicate slower trait evolution (phylogenetic niche conservatism). The tree was estimated based on two recently published mega-trees (Jin & Qian, 2019).

All statistical analyses were conducted in R software (version 3.5.1), using the *V.PhyloMaker*, *ape*, and *pls* packages.

## Results

The final merged vein and trait dataset included 639 branches from 260 species in 47 angiosperm families. Overall, $r_{min}$ ranged from 13 – 68 μm (1% – 99% quantiles), and network extents ranged from 94 – 180 $mm^2$ (1% – 99% quantiles). For visual clarity, we discuss the results in the context of four species selected to illustrate contrasting network architectures (**Figure 2**) – *Dryobalanops lanceolata* (Dipterocarpaceae)*, Pometia pinnata* (Sapindaceae)*, Mallotus wrayi* (Euphorbiaceae)*,* and *Horsfieldia crassifolia* (Myristicaceae).

*Quantification of species network architecture across scales*

All four multiscale statistics demonstrated high variation among species, as well as complex patterns across spatial scales **(Figure 3)**, giving support to our Hypothesis 2. In general, relationships for each multiscale statistic were consistent among leaves within species, i.e. intraspecific plasticity for network architecture traits was low, but variation between species was high. Data for focal species are shown in **Figure 3**, with data for all species in **Supporting Information Figure S2-S5**.

Some species maintained low values of ER(r) ($ER \approx 1$) across scales (**Figure 3a, Supporting Information Figure S2**), reflecting circular loops nested within larger circular loops. Other species showed sharp increases ($ER \approx 5 - 10$) at $r_{min} \approx 0.1$ mm, reflecting nesting of circular loops in longer loops (e.g. *Dryobalanops lanceolata*). Most species then reduced ER at $r_{min} \approx 0.2$ mm, while a small number maintained high or increasing values (e.g. *Mallotus wrayi*).

Many species had a high value of CR ($CR \approx 0.75$) at small $r_{min}$ , then reached a lower value ($CR \approx 0.25$) by $r_{min}$=0.2 mm, reflecting a transition from loops with more complex boundaries to loops with simpler boundaries (**Figure 3b, Supporting Information Figure S3**).

However, the rate of decrease varied strongly among species, with some showing leveling off at intermediate scales (*Pometia pinnata*) and others increases at larger scales (*Dryobalanops lanceolata*).

VD(r) decreased monotonically with $r_{min}$ in all species, reflecting the greater prevalence of small veins relative to large veins in these species at the studied scales (**Figure 3c, Supporting Information Figure S4**). However, the rates of decrease varied across species. Some species showed rapid linear decreases (*Horsfieldia crassifolia*), consistent with vein tapering and a scale continuum, while other species showed abrupt decreases, consistent with discrete transitions between veins of different orders (*Mallotus wrayi* and *Dryobalanops lanceolata*).

Contrarily, MST(r) increased monotonically with $r_{min}$ in all species (**Figure 3d, Supporting Information Figure S5**), reflecting a tendency for more branching and less looping at larger spatial scales. At small $r_{min}$, MST values ranged from 0.5 – 0.8, indicating that a large fraction of all vein segments contributed to loops rather than branches, while at larger $r_{min}$, MST values approached 1 (i.e. no loops). Some species (*Horsfieldia crassifolia*) had much more tree-liked networks at all scales. Although MST values must reach 1 ultimately (i.e. when only a single vein remains in the network), it is mathematically possible for MST to both increase and decrease with $r_{min}$ at intermediate scales. However, no decreases occurred in this dataset. Abrupt increases at some scales were present in some species (*Pometia pinnata*), indicating discrete transitions in branching architecture.

The principal components analysis showed that the four multiscale statistics can be distinguished in at least three axes of variation (**Figure 4**), partially supporting Hypothesis 1. The first axis represented high VD and low MST, and explained 67% of the variation in the data. The second axis represented high ER, and explained 23% of the variation in the data. A third axis represented high CR and explained 8% of the variation in the data. As $r_{min}$ increased, species tended to increase in PC1, decrease in PC2, and increase then decrease in PC3 (**Supporting Information Figure S6**), indicating shifts at larger spatial scales from high to low VD, high to low MST, high to low ER, and low to high to low CR. Additionally, there was a pronounced shift from variation occurring primarily along PC1 and PC2 at small scales ($r_{min} < 0.05$ mm), to variation occurring primarily along PC1 at intermediate scales ($r_{min} > 0.05$ mm), and then primarily along PC2 at larger scales ($r_{min} > 0.1$ mm).

*Trade-offs in network architecture – function across spatial scales*

PLS analysis showed that network architecture was more associated with mechanical (in the first PLS component force to punch explained 10% of model variation) and damage resistance traits (tannins fraction: 8.9%) than to resource transport (photosynthetic capacity: -2%) and cost (leaf mass per area: 3.5%) (**Figure 5**). Inclusion of additional PLS components did not qualitatively improve explained variation or reveal other network architecture linkages (data not shown).

The importance of each multiscale statistic varied across spatial scales. For force to punch (mechanical strength proxy), VD and MST made large contributions at $r_{min} < 0.05$ mm, while ER and CR made large contributions at $r_{min} > 0.05$ mm (**Figure 5a**). These results are largely not consistent with the hypothesized positive linkage to VD at large scales and negative link to ER at small scales. For tannins (damage resistance proxy), only VD contributed at $r_{min} < 0.04$ mm, while CR, ER, and MST contributed at $r_{min} = 0.05$ mm, and only ER contributed at $r_{min} > 0.15$ mm (**Figure 5d**). These results are not consistent with the hypothesized negative linkage to MST and CR at all scales. Results for photosynthetic capacity (resource transport proxy) (**Figure 5c**) had slightly negative explained variation (i.e. the model was worse than one not including the predictors), also not consistent with the hypothesized positive linkage to VD at small scales and positive link to MST at all scales. Results for leaf mass per area (cost proxy) (**Figure 5b**) also had low explained variation, not consistent with the hypothesized positive linkage to VD at large scales. Therefore, we did not find strong evidence for the predictions of Hypothesis 3.

We also repeated these analyses using single-variable regressions at each scale. Results were qualitatively similar (**Supporting Information Figure S7**), indicating that the choice of methodology did not strongly influence our findings.

*Phylogenetic conservatism of network architecture*

All four multiscale statistics varied extensively across the phylogeny with some visual evidence of clustering at certain spatial scales (**Figure 6**). Notably, clustering was apparent at small $r_{min}$ for VD and MST for the family Euphorbiaceae and Dipterocarpaceae, which comprise many species that are ecologically dominant in southeast Asia. In contrast, CR and ER varied extensively across the phylogeny at small $r_{min}$. At larger $r_{min}$, more variation was apparent, with

low values of CR present in some clades (e.g. the Clusiaceae), and more rapid increases in MST in the Fagaceae and Dipterocarpaceae. ER showed extensive variation but no clear association with phylogeny at any scale. That is, patterns of phylogenetic clustering were highly scale dependent.

The K analysis was consistent with this perspective. All four multiscale statistics varied at least as rapidly (K<1) as under a Brownian trait evolution model (**Figure 7**). In the case of VD and CR, Brownian evolution (K=1) could be rejected at almost all scales; for MST, K<1 was supported most $r_{min} < 0.12$ mm, and for ER, at some scales of $r_{min} < 0.08$ mm. Contrary to Hypothesis 4, there was a trend for more trait lability (smaller K) at larger $r_{min}$.

## Discussion

Broadly, we found that 1) that network architecture varied extensively among species and across scales, 2) that the phenotype space shifted and decreased at larger scales, 3) that relationships between venation multiscale statistics and leaf function in this dataset were weak, and 4) phylogenetic niche conservatism in multiscale statistics was low and variable across scales.

### *Independent axes of variation*

Our framework provided a vocabulary for making quantitative and specific statements about network architecture variation. The four axes of VD, CR, ER, and MST provide complementary information about network architecture, with likely three independent axes present (**Figure 4**). Other metrics of network architecture not considered here (e.g. vein branching angles; (Hickey, 1973)), may also be useful, so these four statistics should not be considered a complete set of descriptors.

### *Variation in architecture across scales*

The architectural phenotype space varied across scales (**Figure 3**). In particular, we found a reduction in the size of phenotypic space at larger scales, and a shift to more tree-like and less elongated network geometry. The overall shape of this phenotypic space is potentially consistent with the existence of a number of phenotypic extremes, in which intermediate phenotypes occur as tradeoffs among functions represented by the extreme functional archetypes (a 'Pareto front'; Shoval *et al.* (2012)). Consistent with this perspective, theoretical work has identified extreme

network phenotypes (e.g. all branching, all looping) corresponding to optimization of resource transport and resistance functions, respectively (Ronellenfitsch & Katifori, 2019). Our work now extends these models to a wider range of phenotypic axes and empirical datasets. However, the weak empirical results for Hypothesis 3 (see below) leave unclear whether the extreme phenotypes do indeed represent functional archetypes.

The phenotype space appeared to vary along certain axes differentially across scales, suggestive of multiple drivers affecting patterning at each scale. That is, variation (as seen via confidence ellipses and convex hulls) appeared primarily along PC2 at small and large scales, but along PC1 at intermediate scales. A simple null explanation is that reduced sampling at large scales reduces the available variation (e.g. MST must reach a value of 1 for all leaves once $r_{min}$ reaches the maximum vein size). However, this idea seems insufficient to explain the differential variation along other axes, as well as the contraction of confidence ellipses at intermediate scales.

These results highlight the complexity of the phenotype space, and provide new targets for models meant to capture their evolution (Ronellenfitsch & Katifori, 2017; Ronellenfitsch & Katifori, 2019). Broadly, we found that network statistics at one scale are not predictive of statistics at another scale, and that the limits to phenotype space are scale-dependent. This high heterogeneity also invites questions about the reasons why species may occupy different subsets of the architectural phenotype space, what factors drive shifts, and whether such variation leads to divergent or convergent functioning.

*Functional linkages between network architecture and function*

Each model explained less than 10% of variation in each functional trait (**Figure 5**). Stronger linkages were obtained for tannin fraction (damage resistance) and force to punch (mechanical strength), rather than LMA (cost) and photosynthetic capacity (resource transport). This finding is consistent with prior studies finding vein linkages to structural and defensive traits (Li *et al.*, 2015; Blonder *et al.*, 2018; Hua *et al.*, 2019; Kawai & Okada, 2019) and inconsistent with those finding linkages to hydraulic or resource flux traits (Sack *et al.*, 2008; Brodribb *et al.*, 2016).

Although we found a linkage between force to punch and venation traits, the relationships observed were in the opposite directions of our predictions (Hypothesis 3). For example, we expected mechanical strength to exhibit a positive link to VD at larger scales and to ER at small

scales, but we found the opposite pattern (**Figure 5a**). The most likely explanation is that spatial scale variation in $r_{min}$ in these data was not enough to capture some of the 'large'-scale functional linkages. The largest veins studied here were 0.2 mm, when midveins can be at least 100x larger; similarly, the largest sample was approximately 4 $cm^2$ in area, while leaves can be nearly 1 $m^2$ in area for some species (Kattge *et al.*, 2011). These larger features could be important for explaining mechanical strength, and other axes of functioning. Therefore, describing network architecture across all orders of magnitude is a priority for future investigations. However, one cannot easily quantify large image extents at high resolution, at least with brightfield microscopy. One solution is to take a hybrid approach, using higher resolution microscopy for measurements at smaller scales and lower resolution cameras or scanners at larger scales. Because both analyses yield multiscale statistics with the same scale variable ($r_{min}$), it should be possible to fuse multiscale statistics from multiple ranges of scales, as shown in **Supporting Information Figure S8**.

For damage resistance (tannin), we predicted a negative link to MST and CR at all scales, but found a significant contribution only at $r_{min}$ = 0.05 mm (**Figure 5d**). This result suggests that a high investment in secondary chemistry may not be enough to offset investment in redundant flow pathways. However, we recognize that we examined damage resistance primarily through the lens of herbivory defense, but there are other aspects that may also be linked to venation. For example, the ability to resist damage from drought may also be linked to the venation network, as the probability of cavitation varies with vein size (Sack *et al.*, 2008; Brodribb *et al.*, 2016). Another important functional axis not explored in this study was resilience to damage. While resistance is related to the ability to prevent damages, resilience is the capacity to maintain function after damage has occurred. The presence of vein loops may provide redundant pathways that enable continued flow after damage, caused either by abiotic (e.g. drought, frost) or biotic factors (e.g. herbivory) (Sack et al. 2008; Katifori et al. 2010). However, too much redundancy may be detrimental and might actually decrease resilience, by facilitating the spread of embolisms or diseases.

This study assessed functional linkages of network architecture with a limited set of functional traits and functional axes. The underlying trait dataset we used here actually includes a larger number of traits, including secondary chemistry and isotope and elemental stoichiometry data. These other traits may map onto the functional axes we have already discussed. However,

analyses using these additional traits as proxies for each functional axis did not qualitatively shift results, with maximum PLS1 $R^2$ for any trait of 11% (**Supporting Information Figure S9**). However, other proxies might have relevant connections. In particular, the resource transport axis could be refined in future studies, separating out functions related to carbon uptake and water transport. For example, hydraulic conductance (unmeasured) might be more closely linked to water transport functions, whereas photosynthetic capacity would be more linked to carbon uptake functions (measured). The role of venation in predicting hydraulic conductance appears to be complex, even at single scales. Studies examining the linkage between minor veins and hydraulic conductance have obtained divergent results, suggesting that selection for this relationship may be clade- or biogeographically- dependent (Brodribb *et al.*, 2007; Walls, 2011; Sack *et al.*, 2013; Gleason *et al.*, 2016). More broadly, various aspects of leaf functioning may be decoupled, if there is independent selection on multiple axes of functioning beyond the 'fast-slow' continuum (Wright *et al.*, 2004; Reich, 2014), for example, as has been found in the decoupling between hydraulic and economic functioning in Chinese (Li *et al.*, 2015) and Australian (Gleason *et al.*, 2016) species. Such patterns might also occur if functioning is achieved through other traits not directly related to venation. For example, leaves may modify traits like their mesophyll density or their thickness in order to meet different structural constraints (John *et al.*, 2017), and similarly, outside-xylem pathways (e.g. bundle sheath extensions) may be important factors in determining hydraulic conductivity (Buckley *et al.*, 2015; Ohtsuka *et al.*, 2018).

An alternate explanation for the weak architectural-functional relationships found in this study is that not all aspects of network architecture have immediate functional linkages. That is, much of the diversity of form seen in networks may reflect evolution in the absence of strong selection. For example, the weak relationships between leaf functional traits and ER (**Figure 5**), may indicate that elongation of areoles may occur, but that it is a consequence rather than a cause of other evolutionary forces. For example, leaf aspect ratio and areole elongation are sometimes linked in non-monocots (Blonder *et al.*, 2016), and are potentially coupled via developmental mechanisms related to leaf elongation (Kang & Dengler, 2004). As such, selection on leaf aspect ratio could indirectly drive variation in ER.

*Phylogenetic conservatism of network architecture*

The results also showed that all venation network multiscale statistics varied more rapidly than expected under a Brownian evolution model (Blomberg's $K < 1$). Moreover, there was no increase in K at large scales for any statistic (**Figure 7**). That is, the prediction of more niche conservatism at larger scales (Hypothesis 4) was apparently rejected. These results are consistent with prior analyses of fine-scale network architecture variation (Boyce *et al.*, 2009; Sack & Scoffoni, 2013; Blonder *et al.*, 2018), and extend them to a wider range of scales. This finding runs contrary to ideas that large scale network architecture should be highly conserved, as is expected in plants systematics (Hickey, 1973), but is consistent with rapid rates of evolution seen in some adaptive radiations (Dunbar-Co *et al.*, 2009; Blonder *et al.*, 2016). However, these results should be interpreted with some caution as the sampling design was focused on dominant species; thus, patterns may change if rare species were included.

As in the functional analysis (Hypothesis 3), there was likely insufficient scale variation present in the image data to fully address this research question. Thus, our results do not yet call into question the utility of primary and secondary venation in systematics. It seems likely that larger-scale analyses would likely shift the conclusions of this analysis, especially given the extensive prior use of primary and secondary venation characters as aids to species identification and systematics. Our framework provides an approach that, with additional data, will be able to delineate the scales and contexts in which these applications are well-founded.

## Conclusion

We advanced a framework to quantify network architecture across species and spatial scales. By building on ideas for hierarchical loop decomposition (Katifori & Magnasco, 2012), and leveraging recent machine learning algorithms (Xu *et al.*, 2020), we were able to move past topological (Ronellenfitsch *et al.*, 2015) or single scale (Blonder *et al.*, 2018) descriptions of leaves toward multiscale statistics. Datasets with wider scale variation than ours will further clarify understandings of network form-function linkages.

## Data Availability Statement

Venation data, subsetted trait data, the phylogeny are available at https://doi.org/10.5287/bodleian:QR11d1PD2. R code to replicate all analyses using these files is available at https://doi.org/10.5287/bodleian:E9JP2gjyP. Venation network image data are available in Blonder *et al.* (2019). Segmented venation network images (like those shown in **Figure 2**) are available in Xu *et al.* (2020). Algorithms to analyze images and to calculate multiscale statistics are available in Xu *et al.* (2020). The full trait dataset (including numerous other variables) is available https://doi.org/10.5281/zenodo.3247631.


## Acknowledgments

Fieldwork was supported by Unding Jami, Matheus Henrique Nuñes, Rudi Saul Cruz Chino, Milenka Ximena Montoya, and South East Asia Rainforest Research Program (SEARRP) staff. Research was facilitated by Rob Ewers, Laura Kruitbos, Reuben Nilus, Glen Reynolds, and Charles Vairappan. Species identifications were made by Bernadus Bala Ola, Bill McDonald, Alexander Karolus, and MinSheng Khoo. This work also was supported by the UK Natural Environment Research Council (NERC) (#NE/M019160/1, to BB) and the US National Science Foundation (#DEB- 2025282, to BB). This publication is a contribution from the NERC Human-modified Tropical Forest Programme (#NE/M017508/1, to YAT) and Biodiversity And Land-use Impacts on Tropical Ecosystem Function (BALI) consortium (#NE/K016253/1, to YM and #NE/K016253/1, to YAT). The SAFE Project was funded by the Sime Darby Foundation and the UK NERC. The study areas are part of the Global Ecosystems Monitoring Network (GEM) via an ERC Advanced Investigator Award to YM (#321131). The project also was supported by the Stability of Altered Forest Ecosystems (SAFE) Project, the Sabah Biodiversity Council (SaBC, permits JKM/MBS.1000-2/2 JLD.3-126 and -154), the Institute for Tropical Biology and Conservation (ITBC) at the University of Malaysia, Sabah (UMS), the Sabah Forest Research Centre (FRC) at Sepilok, the Sabah Forestry Department, the SEARRP, Yayasan Sabah (Maliau Basin Conservation Area), and the Maliau Basin and Danum Valley Management Committees. Sean Gleason and several anonymous reviewers provided constructive feedback on the manuscript.


## Author contributions

BB designed the project. SB led the field work. DB, YAT, YM and NM organized the wider research initiative. BB and ISM carried out the statistical analyses and drafted the manuscript. MJ led the venation sample chemical processing. HX and MF developed the algorithms to analyze images of the venation samples. All authors contributed to revising the manuscript.

## References


**Agrawal AA, Konno K. 2009.** Latex: a model for understanding mechanisms, ecology, and evolution of plant defense against herbivory. *Annual Review of Ecology Evolution and Systematics* **40**: 311-331.

**Blomberg SP, Garland T, Ives AR, Crespi B. 2003.** Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* **57**(4): 717-745.

**Blonder B, Baldwin BG, Enquist BJ, Robichaux RH. 2016.** Variation and macroevolution in leaf functional traits in the Hawaiian silversword alliance (Asteraceae). *Journal of Ecology* **104**: 219-228.

**Blonder B, Both S, Jodra M, Majalap N, Burslem D, Teh YA, Malhi Y. 2019.** Leaf venation networks of Bornean trees: images and hand-traced segmentations. *Ecology* **100**(11): e02844.

**Blonder B, Buzzard V, Simova I, Sloat L, Boyle B, Lipson R, Aguilar-Beaucage B, Andrade A, Barber B, Barnes C, Bushey D, Cartagena P, Chaney M, Contreras K, Cox M, Cueto M, Curtis C, Fisher M, Furst L, Gallegos J, Hall R, Hauschild A, Jerez A, Jones N, Klucas A, Kono A, Lamb M, Matthai JDR, McIntyre C, McKenna J, Mosier N, Navabi M, Ochoa A, Pace L, Plassmann R, Richter R, Russakoff B, Aubyn HS, Stagg R, Sterner M, Stewart E, Thompson TT, Thornton J, Trujillo PJ, Volpe TJ, Enquist BJ. 2012.** The leaf-area shrinkage effect can bias paleoclimate and ecology research. *American Journal of Botany* **99**(11): 1756-1763.

**Blonder B, Salinas N, Patrick Bentley L, Shenkin A, Chambi Porroa P, Valdez Tejeira Y, Boza Espinoza T, Goldsmith GR, Enrico L, Martin R, Asner GP, Díaz S, Enquist B, Malhi Y. 2018.** Structural and defensive roles of angiosperm leaf venation network reticulation across an Andes-Amazon elevation gradient. *Journal of Ecology* **106**(4): 1683-1699.

**Blonder B, Salinas N, Patrick Bentley L, Shenkin A, Orland Chambi Porroa P, Valdez Tejeira Y, Violle C, Goldsmith GR, Martin RE, Asner G, Díaz S, Enquist B, Malhi Y. 2017.** Predicting trait-environment relationships for venation networks along an Andes-Amazon elevation gradient. *Ecology* **98**(5): 1239-1255.

**Blonder B, Violle C, Bentley LP, Enquist BJ. 2011.** Venation networks and the origin of the leaf economics spectrum. *Ecology Letters* **14**(2): 91-100.

**Both S, Riutta T, Paine CET, Elias DMO, Cruz RS, Jain A, Johnson D, Kritzler UH, Kuntz M, Majalap-Lee N, Mielke N, Montoya Pillco MX, Ostle NJ, Arn Teh Y, Malhi Y, Burslem DFRP. 2018.** Logging and soil nutrients independently explain plant trait expression in tropical forests. *New Phytologist* **221**(4): 1853-1865.

**Boyce CK, Brodribb TJ, Feild TS, Zwieniecki MA. 2009.** Angiosperm leaf vein evolution was physiologically and environmentally transformative. *Proceedings of the Royal Society B: Biological Sciences* **276**(1663): 1771-1776.

**Brodribb T, Feild T, Jordan G. 2007.** Leaf maximum photosynthetic rate and venation are linked by hydraulics. *Plant physiology* **144**(4): 1890.

**Brodribb TJ, Bienaimé D, Marmottant P. 2016.** Revealing catastrophic failure of leaf networks under stress. *Proceedings of the National Academies of Sciences* **113**(17): 4865–4869.

**Buckley TN, John GP, Scoffoni C, Sack L. 2015.** How Does Leaf Anatomy Influence Water Transport outside the Xylem? *Plant Physiology* **168**(4): 1616-1635.

669 **Choong MF, Lucas PW, Ong JSY, Pereira B, Tan HTW, Turner IM. 1992.** Leaf fracture
670 toughness and sclerophylly: their correlations and ecological implications. *New*
671 *Phytologist* **121**(4): 597-610.
672 **Coley PD, Barone J. 1996.** Herbivory and plant defenses in tropical forests. *Annual Review of*
673 *Ecology and Systematics* **27**(1): 305-335.
674 **Dunbar-Co S, Sporck MJ, Sack L. 2009.** Leaf trait diversification and design in seven rare taxa
675 of the Hawaiian Plantago radiation. *International Journal of Plant Sciences* **170**(1): 61-
676 75.
677 **Ellis B, Daly D, Hickey L. 2009.** *Manual of Leaf Architecture*. New York, NY: New York
678 Botanical Garden.
679 **Givnish T 1979.** On the adaptive significance of leaf form. In: Solbrig O, Jain S, Johnson G,
680 Raven P eds. *Topics in plant population biology*: Columbia University Press, New York,
681 375–407.
682 **Givnish TJ, Pires JC, Graham SW, McPherson MA, Prince LM, Patterson TB, Rai HS,**
683 **Roalson EH, Evans TM, Hahn WJ, Millam KC, Meerow AW, Molvray M, Kores**
684 **PJ, O'Brien HE, Hall JC, Kress WJ, Sytsma KJ. 2005.** Repeated evolution of net
685 venation and fleshy fruits among monocots in shaded habitats confirms a priori
686 predictions: evidence from an ndhF phylogeny. *Proceedings of the Royal Society B*
687 **272**(1571): 1481-1490.
688 **Gleason SM, Blackman CJ, Chang Y, Cook AM, Laws CA, Westoby M. 2016.** Weak
689 coordination among petiole, leaf, vein, and gas-exchange traits across Australian
690 angiosperm species and its possible implications. *Ecology and Evolution* **6**(1): 267-278.
691 **Hickey LJ. 1973.** Classification of the architecture of dicotyledonous leaves. *American Journal*
692 *of Botany* **60**(1): 17-33.
693 **Hua L, He P, Goldstein G, Liu H, Yin D, Zhu S, Ye Q. 2019.** Linking vein properties to leaf
694 biomechanics across 58 woody species from a subtropical forest. *Plant Biology* **22**(2):
695 212-220.
696 **Jin Y, Qian H. 2019.** V. PhyloMaker: an R package that can generate very large phylogenies for
697 vascular plants. *Ecography* **42**(8): 1353-1359.
698 **John GP, Scoffoni C, Buckley TN, Villar R, Poorter H, Sack L. 2017.** The anatomical and
699 compositional basis of leaf mass per area. *Ecology Letters* **20**: 412-425.
700 **Kang J, Dengler N. 2004.** Vein pattern development in adult leaves of Arabidopsis thaliana.
701 *International Journal of Plant Sciences* **165**(2): 231-242.
702 **Katifori E, Magnasco MO. 2012.** Quantifying loopy network architectures. *PLoS ONE* **7**(6):
703 e37994.
704 **Katifori E, Szöllősi GJ, Magnasco MO. 2010.** Damage and fluctuations induce loops in
705 optimal transport networks. *Physical Review Letters* **104**(4): 048704.
706 **Kattge J, Díaz S, Lavorel S, Prentice IC, Leadley P, BöNisch G, Garnier E, Westoby M,**
707 **Reich PB, Wright IJ, Cornelissen JHC, Violle C, Harrison SP, Van Bodegom PM,**
708 **Reichstein M, Enquist BJ, Soudzilovskaia NA, Ackerly DD, Anand M, Atkin O,**
709 **Bahn M, Baker TR, Baldocchi D, Bekker R, Blanco CC, Blonder B, Bond WJ,**
710 **Bradstock R, Bunker DE, Casanoves F, Cavender-Bares J, Chambers JQ, Chapin**
711 **Iii FS, Chave J, Coomes D, Cornwell WK, Craine JM, Dobrin BH, Duarte L, Durka**
712 **W, Elser J, Esser G, Estiarte M, Fagan WF, Fang J, Fernández-Méndez F, Fidelis**
713 **A, Finegan B, Flores O, Ford H, Frank D, Freschet GT, Fyllas NM, Gallagher RV,**
714 **Green WA, Gutierrez AG, Hickler T, Higgins SI, Hodgson JG, Jalili A, Jansen S,**

**Joly CA, Kerkhoff AJ, Kirkup D, Kitajima K, Kleyer M, Klotz S, Knops JMH, Kramer K, Kühn I, Kurokawa H, Laughlin D, Lee TD, Leishman M, Lens F, Lenz T, Lewis SL, Lloyd J, Llusià J, Louault F, Ma S, Mahecha MD, Manning P, Massad T, Medlyn BE, Messier J, Moles AT, Müller SC, Nadrowski K, Naeem S, Niinemets Ü, Nöllert S, Nüske A, Ogaya R, Oleksyn J, Onipchenko VG, Onoda Y, Ordoñez J, Overbeck G, Ozinga WA, Patiño S, Paula S, Pausas JG, Peñuelas J, Phillips OL, Pillar V, Poorter H, Poorter L, Poschlod P, Prinzing A, Proulx R, Rammig A, Reinsch S, Reu B, Sack L, Salgado-Negret B, Sardans J, Shiodera S, Shipley B, Siefert A, Sosinski E, Soussana JF, Swaine E, Swenson N, Thompson K, Thornton P, Waldram M, Weiher E, White M, White S, Wright SJ, Yguel B, Zaehle S, Zanne AE, Wirth C. 2011.** TRY – a global database of plant traits. *Global Change Biology* **17**(9): 2905-2935.

**Kawai K, Okada N. 2019.** Leaf vascular architecture in temperate dicotyledons: correlations and link to functional traits. *Planta* **251**(1): 17.

**Li L, McCormack ML, Ma C, Kong D, Zhang Q, Chen X, Zeng H, Niinemets Ü, Guo D. 2015.** Leaf economics and hydraulic traits are decoupled in five species-rich tropical-subtropical forests. *Ecology Letters* **18**(9): 899-906.

**McKown AD, Cochard H, Sack L. 2010.** Decoding leaf hydraulics with a spatially explicit model: principles of venation architecture and implications for its evolution. *American Naturalist* **175**(4): 447-460.

**Mileyko Y, Edelsbrunner H, Price CA, Weitz JS. 2012.** Hierarchical Ordering of Reticular Networks. *PLoS ONE* **7**(6): e36715.

**Niklas KJ. 1999.** A mechanical perspective on foliage leaf form and function. *New Phytologist* **143**(1): 19-31.

**Ohtsuka A, Sack L, Taneda H. 2018.** Bundle sheath lignification mediates the linkage of leaf hydraulics and venation. *Plant, Cell & Environment* **41**(2): 342-353.

**Pérez-Harguindeguy N, Díaz S, Garnier E, Lavorel S, Poorter H, Jaureguiberry P, Bret-Harte M, Cornwell W, Craine J, Gurvich D. 2013.** New handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of botany* **61**(3): 167-234.

**Poorter H, Niinemets Ü, Poorter L, Wright IJ, Villar R. 2009.** Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. *New Phytologist* **182**(3): 565-588.

**Reich PB. 2014.** The world-wide 'fast–slow' plant economics spectrum: a traits manifesto. *Journal of Ecology* **102**(2): 275-301.

**Ronellenfitsch H, Katifori E. 2017.** The Phenotypes of Fluctuating Flow: Development of Distribution Networks in Biology and the Trade-off between Efficiency, Cost, and Resilience. *arXiv:1707.03074*.

**Ronellenfitsch H, Katifori E. 2019.** Phenotypes of Vascular Flow Networks. *Physical Review Letters* **123**(24): 248101.

**Ronellenfitsch H, Lasser J, Daly DC, Katifori E. 2015.** Topological Phenotypes Constitute a New Dimension in the Phenotypic Space of Leaf Venation Networks. *PLoS Computational Biology* **11**(12): e1004680.

**Ronneberger O, Fischer P, Brox T 2015**. U-net: Convolutional networks for biomedical image segmentation.In Navab N, Hornegger J, Wells WM, Frangi AF. *International conference on medical image computing and computer-assisted intervention*. Munich: Springer. 234-241.

**Roth-Nebelsick A, Uhl D, Mosbrugger V, Kerp H. 2001.** Evolution and function of leaf venation architecture: a review. *Annals of Botany* **87**(5): 553-566.

**Sack L, Dietrich EM, Streeter CM, Sanchez-Gomez D, Holbrook NM. 2008.** Leaf palmate venation and vascular redundancy confer tolerance of hydraulic disruption. *Proceedings of the National Academies of Sciences* **105**(5): 1567-1572.

**Sack L, Frole K. 2006.** Leaf structural diversity is related to hydraulic capacity in tropical rain forest trees. *Ecology* **87**(2): 483-491.

**Sack L, Scoffoni C. 2013.** Leaf venation: structure, function, development, evolution, ecology and applications in the past, present and future. *New Phyt* **198**(4): 983-1000.

**Sack L, Scoffoni C, John GP, Poorter H, Mason CM, Mendez-Alonzo R, Donovan LA. 2013.** How do leaf veins influence the worldwide leaf economic spectrum? Review and synthesis. *Journal of Experimental Botany* **64**(13): 4053-4080.

**Sack L, Scoffoni C, Mckown AD, Frole K, Rawls M, Havran JC, Tran H, Tran T. 2012.** Developmentally based scaling of leaf venation architecture explains global ecological patterns. *Nature Communications* **3**: 837.

**Shoval O, Sheftel H, Shinar G, Hart Y, Ramote O, Mayo A, Dekel E, Kavanagh K, Alon U. 2012.** Evolutionary Trade-Offs, Pareto Optimality, and the Geometry of Phenotype Space. *Science* **336**(6085): 1157.

**Trivett ML, Pigg KB 1996.** A survey of reticulate venation among fossil and living land plants. In: Taylor DW, Hickey L eds. *Flowering plant origin, evolution & phylogeny*. New York, NY: Chapman and Hall, 8-31.

**Vincent JF. 1982.** The mechanical design of grass. *Journal of Materials Science* **17**(3): 856-860.

**Walls RL. 2011.** Angiosperm leaf vein patterns are linked to leaf functions in a global-scale data set. *American Journal of Botany* **98**(2): 244-253.

**Wright I, Reich P, Westoby M, Ackerly D, Baruch Z, Bongers F, Cavender-Bares J, Chapin T, Cornelissen J, Diemer M. 2004.** The worldwide leaf economics spectrum. *Nature* **428**(6985): 821-827.

**Xu H, Blonder B, Jodra M, Malhi Y, Fricker M. 2020.** Automated and accurate segmentation of leaf venation networks via deep learning. *bioRxiv. doi:10.1101/2020.07.19.206631*.

**Supporting Information Legends**

**Table S1.** Revised branch codes for samples.

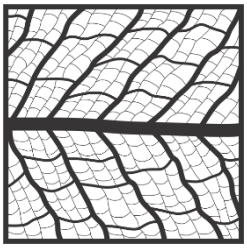
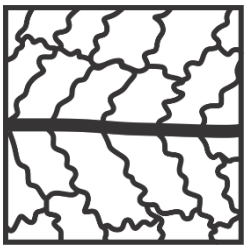
**Fig. S1.** Number of samples with a vein present at each spatial scale ($r_{min}$).
**Fig. S2.** Multiscale statistic for elongation ratio (ER).
**Fig. S3.** Multiscale statistic for circularity ratio (CR).
**Fig. S4.** Multiscale statistic for vein density (VD).
**Fig. S5.** Multiscale statistic for minimum spanning tree ratio (MST).
**Fig. S6.** Mean scores for the principal component analysis in **Fig. 3**.
**Fig. S7.** Alternate linear regression analysis of the effect of multiscale statistics on functional traits.
**Fig. S8.** Illustration of data fusion method for multiscale statistics.
**Fig. S9.** Additional partial least squares analyses for seventeen other traits reported in Both *et al.* (2018).

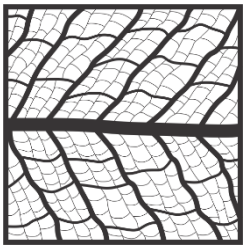
808 **Tables**

809

810 **Table 1**. Venation network multiscale statistics measured in this study.

811

| Statistic | Units | Notes | Low value | High value |
|---|---|---|---|---|
| Vein density (VD) | mm $mm^{-2}$ | calculated as the length of all vein segments divided by the area of the leaf sample. Higher VD values indicate more veins. |  |  |
| Loop elongation Ratio (ER) | dimensionless, ranging from 0 (short loop) to ∞ (long loop) | calculated by fitting an ellipse to each loop, and dividing the major axis length by the minor axis length, then taking the median ratio across all loops and subtracting 1 (for later convenience of fits). Higher ER values indicate longer, more "stretched" loops. |  |  |
| Loop circularity (CR) | dimensionless, ranging from 0 (more infolding) to 1 (less infolding) | calculated by dividing the loop area by the squared loop perimeter for each loop, then multiplying by 4π, then taking the median value across all loops. Higher CR values indicate loops |  |  |

| | | | | |
|---|---|---|---|---|
| | | with less infolding. | | |
| Minimum spanning tree ratio (MST) | dimensionless, ranging from 0 (more loops) to 1 (fewer loops) | calculated by computing the length of the minimum spanning tree connecting all vein junctions divided by the length of all veins. Higher MST values indicate more branching (tree-like) pathways with fewer loops. |  |  |

812

813

814 **Table 2.** Functional traits used as proxies for different functional axes.

| *Functional axis* | *Trait proxy* | *Units* |
|---|---|---|
| Resource transport | Photosynthetic capacity ($A_{max}$) | (%) |
| Damage resistance | Total tannins fraction | (mg $g^{-1}$) |
| Mechanical strength | Force to punch (Fp) | N $mm^{-1}$ |
| Construction cost | Leaf mass per area (LMA) | g $cm^{-2}$ |

815

816 **Table 3.** Site information. Site code (SI) in this database refers to the filename structure used for
817 branch identification in **File S1**. Site code (FP) refers to the plot name in the ForestPlots
818 database.
819

| *Site code (SI)* | *Site code (FP)* | *Site name* | *Number of branches with vein and trait data* | *Number of species with vein and trait data* | *Number of vein segments in all branches* |
|---|---|---|---|---|---|
| DAS1 | DAN-04 | Danum GEM Carbon Plot 1 | 81 | 48 | 1186605 |
| DAF2 | DAN-05 | Danum GEM Carbon Plot 2 | 52 | 38 | 849519 |
| BEL | MLA-01 | Maliau SAFE GEM Carbon Plot: Belian | 78 | 50 | 1077431 |
| SER | MLA-02 | Maliau SAFE GEM Carbon Plot: Seraya | 86 | 52 | 1587871 |
| BSO | SAF-01 | SAFE GEM Carbon Plot B South | 85 | 56 | 1474711 |
| BNT | SAF-02 | SAFE GEM Carbon Plot B North | 103 | 55 | 2031577 |
| ESA | SAF-03 | SAFE GEM Carbon Plot E | 120 | 81 | 1997597 |
| SLF | SAF-04 | SAFE GEM Carbon Plot LF | 34 | 19 | 265037 |

820
821

**Figure Legends**

**Figure 1.** Cartoon of the hierarchical loop decomposition process. **A)** Areoles are connected by vein segments of varying radii. The lowest-radius vein segment is identified and the adjacent areoles are fused into a larger region. **B-D)** This process is iterated until a single region remains. At each iteration, multiscale statistics of the fused network are calculated.

**Figure 2.** Example venation networks for **(a)** *Dryobalanops lanceolata* (code BNT-T212-BSH), **(b)** *Pometia pinnata* (DAS1-T010227-BSH), **(c)** *Mallotus wrayi* (SER-T462-BSH), and **(d)** *Horsfieldia crassifolia* (SLF-T53-BSH). Images are cropped sections of machine learning algorithms applied to raw imagery. Each panel shows 8.4 × 8.4 mm of lamina. These images correspond to the highlighted species in **Figure 3**.

**Figure 3.** Multiscale statistics for variation in **(a)** elongation ratio (ER), **(b)** circularity ratio (CR), **(c)** vein density (VD), and **(d)** minimum spanning tree ratio (MST), at each value of minimum vein size ($r_{min}$). Black lines indicate data for all leaves in the full dataset. Colored lines indicate data for four focal leaves shown in **Figure 1**. Gaps in lines indicate scales at which no veins were observed.

**Figure 4.** Principal components analysis of the four-dimensional architectural space comprising elongation ratio (ER), circularity ratio (CR), vein density (VD), and minimum spanning tree ratio (MST) across scales, at each value of minimum vein size ($r_{min}$). 95% confidence ellipses enclose the data at $r_{min}$ value to show the central tendency in the data; convex hulls are also shown to highlight species with extreme phenotypes. Parenthetical values indicate variance explained by each axis.

**Figure 5.** Summaries of partial least squares (PLS) models predicting leaf functional traits from elongation ratio (ER), circularity ratio (CR), vein density (VD), and minimum spanning tree ratio (MST) across scales, at each value of minimum vein size ($r_{min}$). Each facet indicates a model with a different leaf functional trait: **(a)** force to punch, **(b)** leaf mass per area, **(c)** photosynthetic capacity, and **(d)** tannins fraction. Colored lines indicate standardized effects of each multiscale

853 statistic at each spatial scale. Thicker colored lines indicate models with higher $R^2$. Loading
854 coefficients reflect the first PLS component only.
855
856 **Figure 6.** Phylogenetic distribution of multiscale statistics for all species in the study. Heatmaps
857 indicate values of **(a)** elongation ratio (ER), **(b)** circularity ratio (CR), **(c)** vein density (VD), and
858 **(d)** minimum spanning tree ratio (MST) at each value of minimum vein size ($r_{min}$). Shaded clades
859 are the Dipterocarpaceae (blue), Clusiaceae *sensu latu* (inc. Hyperiaceae) (green), Euphorbiaceae
860 (orange), and Fagaceae (red).
861
862 **Figure 7.** Blomberg's K statistic of phylogenetic conservatism for elongation ratio (ER),
863 circularity ratio (CR), vein density (VD), and minimum spanning tree ratio (MST) at each value
864 of minimum vein size ($r_{min}$) (colored dots). Values of 1 (dashed horizontal black line) indicate
865 Brownian trait evolution; values <1 indicate more rapid evolution than under Brownian trait
866 evolution. Points are solid if significantly different from 1 ($p<0.05$) and transparent otherwise.
867
868