www.sciencemag.org/cgi/content/full/science.1231574/DC1

Supplementary Materials for

Global Leaf Trait Relationships: Mass, Area, and the Leaf Economics Spectrum

Jeanne L. D. Osnas,* Jeremy W. Lichstein, Peter B. Reich, Stephen W. Pacala

*Corresponding author. E-mail: jldosnas@gmail.com

Published 28 March 2013 on *Science* Express
DOI: 10.1126/science.1231574

This PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S6
Tables S3 and S4
Full Reference List

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/science.1231574/DC1)

Tables S1 and S2 (as Excel files)

Methods

Data set description and modifications

We modified the original 2004 Wright et al. (*2*) GLOPNET leaf trait data set by deleting the 178 (out of 2548 total) observations that lacked LMA data and two observations not identified to genus ("Unidentified compositae" and "Leguminosae sp"), leaving us with 2368 observations. We also converted units of mass-normalized N and P from percent to g g^{-1} (by dividing by 100). Units for area-normalized N and P were g m^{-2}. Finally, we converted area-normalized Amax and Rdark to units of nmol m^{-2} s^{-1} so that they were more easily compared to the mass-normalized units of nmol g^{-1} s^{-1}. Units of LMA and LL were g m^{-2} and months, respectively. This left us with 2368 observations for LMA, 1957 for N, 763 for Amax, 739 for P, 274 for Rdark and 677 for LL.

After these modifications the 2368 observations remaining in the dataset were comprised of 1852 unique species. While there are multiple observations of some species in the GLOPNET dataset, no trait values are the same within a species, and no species has more than 8 observations. This version of the dataset that includes intraspecific variation is the appropriate dataset to investigate the contribution of both normalization-independent and normalization-induced correlation to the LES presented by Wright et al. (*2*). A version of the dataset free of intraspecific variation, however, is more appropriate for characterizing interspecific normalization-independent trait relationships and estimating the area- and mass-proportionality of traits. Toward this end, we created

another version of the GLOPNET dataset purged of intraspecific variation by replacing multiple observations per species with the mean of those values. We repeated analyses with this dataset. Doing so only negligibly altered the results and did not change our conclusions, so we only present analyses using the complete dataset. So, we will henceforth here and in the main text refer to any observation in GLOPNET as a "species" observation.

Models

We used a suite of statistical models to quantify the degree to which traits are distributed across species proportional to mass vs. area (see explanation of terms below). In this section, we describe the models. In subsequent sections, we described the methods used to (i) obtain maximum likelihood estimates and confidence intervals of model parameters, (ii) simulate random data from select models, and (iii) partition traits into area- and mass-associated fractions.

Although LMA does vary within species, we treat it here as a species-specific constant because global variation in LMA is much larger between species than within (*11, 19*). Below we report analyses of mass- and area-normalized values reported in GLOPNET. We repeated these analyses except that we used each species' mean LMA times its mass-normalized trait value in place of the area-normalized value reported in GLOPNET; and similarly, we used each species' area-normalized trait value divided by its mean LMA instead of the mass-normalized value in GLOPNET. These values differ slightly from the area- and mass-normalized values reported in GLOPNET because area- and mass-normalized are not inter-converted exactly with species mean LMA values in GLOPNET. Nevertheless, these analyses returned nearly identical results to those described below (based on the mass- and area- normalized values reported in GLOPNET), which further justifies our treatment of LMA as a species-specific constant.

In addition to analyzing both mass- and area-normalized data, we also obtained unpublished values of total leaf mass and total leaf area for a subset of the leaves in GLOPNET. These mass and area values allowed us to fit the models described below using total (non-normalized) trait data. These analyses (which we do not report here) confirmed that our models yield nearly identical results whether fit to non-normalized, mass-normalized, or area-normalized trait data.

Formally, a trait is distributed proportional to mass if normalizing by mass makes the trait statistically independent of all measures of leaf size, including mass, area and LMA. Area-normalization similarly removes all effects of leaf size from an area-proportional trait. A definition of mass-proportionality that is easier to visualize is the per-unit-mass increase in the total (non-normalized) trait amount as mass increases across leaves *while controlling for leaf area*. We refer to the mass-proportional trait fraction as being 'distributed *across* species proportional to mass.' Similarly, we define area-proportionality as the per-unit-area increase in the total trait amount as area increases across leaves *while controlling for leaf mass*. We refer to the area-proportional trait fraction as being 'distributed *across* species proportional to area.' Leaf traits may be purely mass-proportional, purely area-proportional, or both mass- and area-proportional. To illustrate our use of these terms, consider a collection of leaves with the *same area but variable mass*. If at least some fraction of trait-*i* were *distributed proportional to mass*, then the total (non-normalized) amount of trait-*i* would tend to increase with increasing

mass across these equal-area leaves. Conversely, consider a collection of leaves with the *same mass but variable area*. If at least some fraction of trait-*i* were *distributed proportional to area*, then the total amount of trait-*i* would tend to increase with increasing area across these equal-mass leaves.

It is important to note that, according to our definitions, the terms mass- and area-proportionality do not imply anything about the physical location of traits within individual leaves. For example, a trait could (in principle) be distributed *within leaves* uniformly throughout the entire leaf mass (or thickness), yet be distributed *across species* proportional to area (and independent of mass) if the amount of trait per unit mass were inversely proportional to LMA. Conversely, a trait could (in principle) be concentrated *within leaves* in a thin layer (e.g., the palisade layer), yet be distributed *across species* proportional to mass (and independent of area) if the amount of trait per unit area were proportional to LMA. Thus, as defined here, the terms mass-proportional and area-proportional simply refer to how the total amount of trait tends to change across leaves as mass and area change (i.e., how traits are distributed *across species*), but do not imply anything about how traits are distributed *within leaves* (e.g., distributed uniformly throughout the thickness of the leaf vs. concentrated in a thin layer).

We now describe five statistical models that quantify how the total amount of trait in a leaf depends on leaf mass and/or area. The first model, which we refer to as 'Model-M&A,' allows for both mass- and area-proportional trait. Let $\bar{X}_{Mi}$ and $\bar{X}_{Ai}$ be constants that quantify mass- and area-proportionality, respectively, for trait-*i*. The units of $\bar{X}_{Mi}$ and $\bar{X}_{Ai}$ are the amount of trait-*i* per-unit-mass and per-unit-area, respectively. The expected value of the total amount of trait-*i* in a species-*k* leaf is $\text{Mass}_k \bar{X}_{Mi} + \text{Area}_k \bar{X}_{Ai}$. The model has three mathematically identical forms, depending on how data are normalized:

(Model-M&A) $$X_{ik} = (Mass_k \bar{X}_{Mi} + Area_k \bar{X}_{Ai})\varepsilon_{ik}$$

(Model-M&A) $$X_{Mik} = (\bar{X}_{Mi} + \bar{X}_{Ai} LMA_k^{-1})\varepsilon_{ik}$$

(Model-M&A) $$X_{Aik} = (LMA_k \bar{X}_{Mi} + \bar{X}_{Ai})\varepsilon_{ik}$$

where X_{ik}, X_{Mik}, and X_{Aik} are, respectively, the total, mass-normalized, and area-normalized amounts of trait *i* in a species-*k* leaf; and ε_{ik} is a species-specific, unit-mean, unitless random variable that quantifies interspecific variation (we assume that measurement error in GLOPNET is small relative to interspecific variation, and we henceforth ignore it). The three forms of Model-M&A are easily inter-converted via mass, area, or LMA. Note that X_{Mik} and X_{Aik} are simply normalized trait values for species-*k*; these values imply nothing about mass- or area-proportionality as defined here. In contrast, $\bar{X}_{Mi}$ and $\bar{X}_{Ai}$ quantify mass- and area-proportionality across species. No matter how large or small the mass- or area-proportional amounts might be for trait-*i* of species-*k* ($\text{Mass}_k \bar{X}_{Mi}$ or $\text{Area}_k \bar{X}_{Ai}$, respectively), we can still express the trait on a mass or area basis (X_{Mik} or X_{Aik}) by dividing the total trait amount (X_{ik}) by mass or area. This highlights the difference between mass- or area-*normalization* (i.e., expressing a trait on a

mass- or area-basis, which has nothing to do with how the traits are distributed across species) and mass- or area-*proportionality* (which reflects how traits are distributed across species, regardless of how they are normalized).

The second model is similar to Model-M&A above, but allows for species-specific variation separately in the mass- and area-associated traits amounts. We call this second model 'Model-C,' where the letter "C" indicates that the likelihood includes a convolution integral (see next section below). Model-C is used in the main text and in Fig. 2 to quantify the species distribution of trait fractions associated with mass vs. area; similar mean fractions are obtained from Model-M&A described above (see details below in the section titled "Partitioning Traits into Area- and Mass-Fractions"). In Model-C, the amount of mass-associated trait-*i* in a species-*k* leaf is the product of three terms: leaf mass ($Mass_k$); the mass-proportional constant ($\bar{X}_{Mi}$, with per-unit-mass units); and a species-specific, unit-mean, unitless value (ε_{Mik}). Similarly, the amount of area-associated trait-*i* in a species-*k* leaf is the product of leaf area ($Area_k$); the area-proportional constant ($\bar{X}_{Ai}$, with per-unit-area units); and a species-specific, unit-mean, unitless value (ε_{Aik}). As in Model-M&A, $\bar{X}_{Mi}$ and $\bar{X}_{Ai}$ quantify, respectively, mass- and area-proportionality *across species*. In contrast, ε_{Mik} and ε_{Aik} quantify species-specific deviations from the expected mass- and area-proportional trait amounts. The three mathematically identical forms of Model-C are:

(Model-C) $$X_{ik} = Mass_k \bar{X}_{Mi} \varepsilon_{Mik} + Area_k \bar{X}_{Ai} \varepsilon_{Aik}$$

(Model-C) $$X_{Mik} = \bar{X}_{Mi} \varepsilon_{Mik} + LMA_k^{-1} \bar{X}_{Ai} \varepsilon_{Aik}$$

(Model-C) $$X_{Aik} = LMA_k \bar{X}_{Mi} \varepsilon_{Mik} + \bar{X}_{Ai} \varepsilon_{Aik}$$

The third and fourth models (Model-M and Model-A) assume that traits are either purely mass-proportional or purely area-proportional and are used in the main text to evaluate the two competing hypotheses that traits are either mass- or area-proportional. We define a trait as being 'purely mass-proportional' if the total amount of trait increases isometrically with mass (controlling for area) but is unrelated to area (controlling for mass). In other words, a trait is purely mass-proportional if the mass-proportional constant ($\bar{X}_{Mi}$) is greater than zero, but the area-proportional constant ($\bar{X}_{Ai}$) is zero. In this case, the total trait amount may increase across leaves as area increases, but only due to the positive correlation between area and mass. If a trait is purely mass-proportional, then mass, not area, causes the total trait amount to increase from small leaves to large leaves. This extreme case, which we label as Model-M, has three mathematically identical forms:

(Model-M) $$X_{ik} = Mass_k \bar{X}_{Mi} \varepsilon_{Mik}$$

(Model-M) $$X_{Mik} = \bar{X}_{Mi} \varepsilon_{Mik}$$

(Model-M) $$X_{Aik} = \bar{X}_{Mi} LMA_k \varepsilon_{Mik}$$

where terms are defined as above for Model-M&A. The non-normalized form of Model-M (the first expression above) states that if a trait is purely mass-proportional, then

interspecific variation in the amount of trait in a leaf is due to two factors alone: interspecific variation in leaf mass ($Mass_k$) and interspecific variation in the amount of trait per unit leaf mass (ε_{Mik}). This seemingly trivial statement has important implications. Firstly, if we divide both sides of the expression by leaf mass to obtain the mass-normalized form of Model-M (second expression above), we see that mass-normalized values of purely mass-proportional traits are uncorrelated across species with mass, area, or LMA. Secondly, if we instead divide through by leaf area to obtain the area-normalized form of Model-M (third expression above), we see that area-normalized values of purely mass-proportional traits increase across species with LMA. *Thus, if traits are purely mass-proportional, then mass-normalized trait values should be uncorrelated with LMA, whereas area-proportional trait values should increase with LMA*.

We define a trait as being 'purely area-proportional' if the total amount of trait increases isometrically with area (controlling for mass) but is unrelated to mass (controlling for area). In other words, a trait is purely area-proportional if the area-proportional constant ($\overline{X}_{Ai}$) is greater than zero, but the mass-proportional constant ($\overline{X}_{Mi}$) is zero. This extreme case, which we label as Model-A, has three mathematically identical forms:

(Model-A) $X_{ik} = Area_k \overline{X}_{Ai} \varepsilon_{Aik}$

(Model-A) $X_{Aik} = \overline{X}_{Ai} \varepsilon_{Aik}$

(Model-A) $X_{Mik} = \overline{X}_{Ai} LMA_k^{-1} \varepsilon_{Aik}$

The interpretation of Model-A is analogous to Model-M, described above. Briefly, the non-normalized form of Model-A (the first expression above) states that if a trait is purely area-proportional, then interspecific variation in the amount of trait in a leaf is due to two factors alone: interspecific variation in leaf area ($Area_k$) and interspecific variation in the amount of trait per unit leaf area (ε_{Aik}). The area-normalized form (second expression) states that area-normalized values of purely area-proportional traits are uncorrelated with mass, area, or LMA. Finally, the mass- normalized form (third expression) states that mass-normalized values of purely area-proportional traits are inversely proportional to LMA.

Finally, the fifth model (Model-LN) is equivalent to the normalization-independent LES defined by Equations 1 and 2 in the main text. Like Model-M&A and Model-C above, Model-LN allows for both mass- and area-associated trait amounts. This model's three mathematically identical forms are:

(Model-LN) $X_{ik} = K_i (Mass_k^{S_{Ai}} Area_k^{1-S_{Ai}}) \varepsilon_{ik}$

(Model-LN) $X_{Aik} = K_i (LMA_k^{S_{Ai}}) \varepsilon_{ik}$

(Model-LN) $X_{Mik} = K_i (LMA_k^{S_{Ai}-1}) \varepsilon_{ik}$

where K_i and S_{Ai} are constants and ε_{ik} is a unitless, unit-mean value representing interspecific variation in the amount of trait. The area- and mass-normalized forms of Model-LN (second and third expressions, respectively) are obtained from the non-normalized form (first expression) by dividing through by area or mass, respectively.

Note that the residual variation (ε_{ik}) is identical across the three forms of Model-LN. Although Models-LN, -M&A, and -C have different mathematical forms, all three lead to a similar quantitative partitioning of mass and area trait fractions (see the section below titled "Comparisons of Models-LN and -C").

Taking natural logs of the area- and mass-normalized forms of Model-LN, and assuming that ε_{ik} is lognormally distributed, leads directly to Equations 1 and 2 in the main text:

$$\text{(Model-LN)} \quad \ln X_{Aik} = I'_i + S'_{Ai} \ln LMA_k + n'_{ik}$$

$$\text{(Model-LN)} \quad \ln X_{Mik} = I'_i + \left(S'_{Ai} - 1\right) \ln LMA_k + n'_{ik}$$

where $I'_i = \ln(K_i) - \sigma^2_i/2$, and n'_{ik} is a zero-mean normal random variable with variance σ^2_i. The $\sigma^2_i/2$ term in the intercept I'_i is due to the fact that the mean of $\exp(n'_{ik})$, a lognormal random variable with a mean of zero on the log-scale, is $\exp(\sigma^2_i/2)$. The n'_{ik} terms in the above equations are formally the random variables for the natural log of trait-i conditional on ln(LMA). Because the logarithms of the traits are normal, the constants are estimated by ordinary least squares (OLS) regression, and the variance-covariance matrix of the random variables (n'_{ik}) is estimated from the variances and covariances of the OLS residuals. The slope of the regression line for mass-normalized data, S'_{Mi}, is equal to $S'_{Ai} - 1$. To see this, subtract $\ln(\mathrm{LMA}_k)$ from both sides of the first equation above to convert it into the second. By dividing both sides of the above equations by ln(10), and thus converting to base ten logarithms, we arrive at Equations 1 and 2 in the main text, where $I = I'/\ln(10)$, $S = S'/\ln(10)$ and $n = n'/\ln(10)$. The mathematical equivalence of the mass- and area-normalized forms of Model-LN is the reason that $\log_{10}$LMA-OLS regression residuals from mass- and area-normalized data (Equations 1 and 2 in main text) produce the same normalization-independent LES (Fig. 4, Fig. S4 and Fig. S5).

Maximum Likelihood Estimation and Confidence Limits

With the exception of Model-C, all of the models described above can be fit using standard linear and nonlinear least-squares statistical software by (i) assuming that the random variables, ε, are lognormally distributed; and (ii) log-transforming the models so that the random variables become additive, normally distributed errors. However, Model-C includes two random variables and, therefore, cannot be fit with standard software. For consistency, we fit all models by maximum likelihood without log-transformations (i.e., we back-transformed the $\log_{10}$ values reported in GLOPNET). To perform equivalent analyses on a log-scale with least-squares software (for all models except for Model-C), one must adjust for the fact that the expected value of a lognormal random variable is $\exp(\mu_{\log} + \sigma^2/2)$, where $\mu_{\log}$ and σ^2 are, respectively, the mean and variance on the log-scale. Ignoring this fact would still yield results that are qualitatively similar to the results we report here.

We fit the models to the GLOPNET data for Amax, Rdark, N and P by assuming that each species' set of six traits represents an independent draw from a multivariate lognormal distribution. For example, in Model-M&A, the ε_{ik} for $k = 1, 2, \ldots, Q$ (where Q is the number of species) are assumed to be independent unit-mean lognormal random

variables for trait-i. Traits are allowed to be correlated within a species: ε_{ik} may be correlated with ε_{jk}. In Model-C, these same assumptions also apply to ε_{Aik} and ε_{Ajk}, and to ε_{Mik} and ε_{Mjk}. Additionally, ε_{Aik} and ε_{Mik} are assumed to be statistically independent in Model-C.

We first treat the four models with a single random variable (ε): Models-M&A, -M, -A, and -LN. Let D_{ik} be the measurement for leaf trait-i and species-k in the GLOPNET data; and let $P(D_{ik}|\mu_i(LMA_k),\sigma^2_i)$ be the lognormal probability density for leaf trait-i, where the function $\mu_i(LMA_k)$ is the mean for species-k, and σ^2_i is the variance of the logarithm of trait-i. In our notation, $\mu(LMA)$ is the expectation on the linear (non-log) scale; i.e., $\mu(LMA) = \exp(\mu_{\log} + \sigma^2/2)$. If data are area-normalized then $\mu_i(LMA_k)$ is $\bar{X}_{Ai} + \bar{X}_{Mi} LMA_k$ for Model-M&A, $\bar{X}_{Ai}$ for Model-A, $\bar{X}_{Mi} LMA_k$ for Model-M, and $K_i(LMA_k^{S_{Ai}})$ for Model-LN. If data are mass-normalized then $\mu_i(LMA_k)$ is $\bar{X}_{Ai}/LMA_k + \bar{X}_{Mi}$ for Model-M&A, $\bar{X}_{Ai}/LMA_k$ for Model-A, $\bar{X}_{Mi}$ for Model-M, and $K_i(LMA_k^{S_{Ai}-1})$ for Model-LN. The log-likelihood function for trait-i in the above four models is: $\ell = \sum_{k=1}^{Q} \ln P(D_{ik} \mid \mu_i(LMA_k),\sigma_i^2)$.

The log-likelihood function for Model-C is given by the convolution of two lognormal random variables:

$$\ell = \sum_{k=1}^{Q} \ln\left[\int_0^{D_{ik}} P_A\left(y \mid \mu_{Ai}\left(LMA_k\right),\sigma_{Ai}^2\right) P_M\left(D_{ik} - y \mid \mu_{Mi}\left(LMA_k\right),\sigma_{Mi}^2\right) dy\right]$$

where $P_A(\cdot)$ is the lognormal density for the amount of trait-i that is associated with area, and $P_M(\cdot)$ is the lognormal density for the amount associated with mass. If data are area-normalized then $\mu_{Ai}(LMA_k)$ is $\bar{X}_{Ai}$ and $\mu_{Mi}(LMA_k)$ is $\bar{X}_{Mi} LMA_k$. If data are mass-normalized then $\mu_{Ai}(LMA_k)$ is $\bar{X}_{Ai}/LMA_k$ and $\mu_{Mi}(LMA_k)$ is $\bar{X}_{Mi}$.

All of the above models were fit using R software (*20*). Maximum likelihood estimates were obtained with the "mle2" function in the "bbmle" package (*21*), and confidence intervals were estimated from the second derivatives of the likelihood surface by specifying the "quadratic" method in the "confint" function.

Random Data Sets

To illustrate how normalization induces correlations among traits, we generated simulated data using Models-C, -M, and -A. For each model, we produced a null data set (no functional correlations) containing area- and mass-normalized values of Amax, Rdark, N and P. For each null data set, values of X_{Mi} and X_{Ai} were produced by combining (i) randomly generated values of ε_{Mik} and/or ε_{Aik} (depending on the model) using the maximum likelihood estimates for their variances, (ii) maximum likelihood estimates for $\bar{X}_{Mi}$ and/or $\bar{X}_{Ai}$ (depending on the model), and (iii) the actual LMA and LL values reported in GLOPNET. The null data sets lacked functional correlations among Amax, Rdark, N and P (or between these four traits and LMA or LL) because the values of ε_{Mik} and ε_{Aik} for any two traits within the same species were statistically independent. Thus, leaves in the null model with the same (or similar) LMA had statistically independent values of Amax, Rdark, N and P.

Partitioning Traits into Area- and Mass-Fractions

For Models-M&A and -C, we estimated the species-distribution of the fraction of each leaf trait-i that is associated with leaf area (p_{Ai}); i.e., the "area-proportional" fraction, as defined in this paper. These species distributions quantify, for each trait-i, the fraction of species whose area-associated fraction lies in a given interval $a \le p_{Ai} < b$, where $0 \le a < b \le 1$. For Model-M&A, the area-associated fraction for trait-i and species-k is

$$p_{Aik} = Area_k \bar{X}_{Ai} / (Area_k \bar{X}_{Ai} + Mass_k \bar{X}_{Mi}) = \bar{X}_{Ai} / (\bar{X}_{Ai} + LMA_k \bar{X}_{Mi}),$$

and for Model-C it is

$$p_{Aik} = Area_k \bar{X}_{Ai} \varepsilon_{Aik} / (Area_k \bar{X}_{Ai} \varepsilon_{Aik} + Mass_k \bar{X}_{Mi} \varepsilon_{Mik}) = \bar{X}_{Ai} \varepsilon_{Aik} / (\bar{X}_{Ai} \varepsilon_{Aik} + LMA_k \bar{X}_{Mi} \varepsilon_{Mik}).$$

To account for parameter value uncertainty in the calculation of p_{Aik}, we approximated the sampling distribution of each model's parameters as multivariate normal with mean given by the MLEs and variance-covariance matrix estimated from the second derivatives of the likelihood function. The parameter vectors needed to calculate p_{Aik} are $(\bar{X}_{Ai}, \bar{X}_{Mi})$ and $(\bar{X}_{Ai}, \bar{X}_{Mi}, \sigma^2_{Ai}, \sigma^2_{Mi})$ for Models-M&A and -C, respectively. For Model-M&A, we took 1000 random draws from the multivariate normal parameter distribution. For each of these 1000 parameter sets, we generated a p_{Aik} value for each GLOPNET species-k using the species' reported LMA value. We then estimated the distribution of p_{Ai} (mean and percentiles) by pooling all p_{Ai} values. For example, for N, there were 1.957×10^6 p_{Ai} values (1000 values for each of 1,957 N values reported in GLOPNET). For Model-C, we followed a similar procedure, but also accounted for interspecific variation in the mass- and area-associated trait fractions. As with Model-M&A, we took 1000 random draws from the multivariate normal parameter distribution of Model-C to create 1000 sets of parameters. For each of these 1000 parameter sets, we generated a p_{Aik} value for each GLOPNET species-k by combining each species' LMA value reported in GLOPNET with a random draw of ε_{Aik} and ε_{Mik} (one value of each random variable for each LMA value). We then calculated the distribution of p_{Ai} across species for trait-i by combining all random draws for all species and all 1000 parameter sets.

In addition to the above procedure described for Model-C, we also applied the Model-M&A procedure to Model-C (by ignoring interspecific variation, ε), which yielded qualitatively similar results. The species distributions from Model-M&A are summarized in Table S1. The species distributions from Model-C are summarized in the main text and in Fig. 2 (full analysis, including ε), and also in Table S1 (ignoring ε). The species distributions from both models are formally equivalent to posterior distributions assuming non-informative priors.

AIC Tests

We evaluated the models against one another using the AIC criterion and test. Comparisons of Model-C with Model-A or Model-M allowed us to determine if a trait's mass- or area-proportional fraction was statistically significant (see below).

Posterior probability of Models-M and -A

If we assume that Models-A and -M are the only possible models, and that each has prior probability equal to 0.5, then the posterior probability of Model-M is $1/[1+\exp(\Delta_L)]$, where Δ_L is the log-likelihood of Model-A evaluated at its maximum likelihood parameter estimates minus the corresponding quantity for Model-M. The above expression can alternatively be expressed as $L_M/(L_M + L_A)$, where L_M and L_A, are, respectively, the maximum likelihoods of Model-M and Model-A.

SOM Text

Effects of normalization on covariances

In this SOM section, we show how mass-normalization affects the covariances of area-proportional traits. [In the SOM section below titled "Understanding the normalization-independent LES," we analyze how both mass- and area-normalization affect the correlations and slopes between trait pairs in general (not necessarily area-proportional proportional traits).] Consider two traits, i and j, that are purely proportional to area (see explanation above in the "Methods / Models" section). By definition, if traits i and j are area-proportional, then the area-normalized traits, X_{Aik} and X_{Ajk}, are statistically independent of Mass_k and Area_k; i.e., the covariances between the traits and mass or area are zero. We do not assume that the covariance between X_{Aik} and X_{Ajk}, $\text{Cov}(X_{Aik}, X_{Ajk})$, is zero. It is convenient to take logarithms (of any base), which preserves the independence assumptions: $\text{Cov}(\log X_{Aik}, \log\text{Area}_k) = 0$, $\text{Cov}(\log X_{Aik}, \log\text{Mass}_k) = 0$, and $\text{Cov}(\log X_{Aik}, \log X_{Ajk})$ not necessarily equal to zero. We require three identities to quantify the effect of normalization on these covariances. Let z1, z2 and z3 be three random variables. Then:

(i) $\text{Cov}(z1 - z3, z2 - z3) = \text{Cov}(z1, z2) - \text{Cov}(z2, z3) - \text{Cov}(z1, z3) + \text{Var}(z3)$

(ii) $\text{Cov}(z1 - z3, z3) = \text{Cov}(z1, z3) - \text{Var}(z3)$

(iii) $\text{Cov}(z1, z2 - z3) = \text{Cov}(z1, z3) - \text{Cov}(z2, z3)$

where $\text{Cov}(\cdot)$ and $\text{Var}(\cdot)$ are, respectively, the covariance and variance between the variables in parentheses. An area-normalized trait is converted to a mass-normalized trait as $X_{Mik} = X_{Aik}/\text{LMA}_k$; or, in log-space: $\log X_{Mik} = \log X_{Aik} - \log\text{LMA}_k$. So, by identiy (i):

$\mathrm{Cov}(\log X_{Mik}, \log X_{Mjk}) = \mathrm{Cov}(\log X_{Aik} - \mathrm{logLMA}_k, \log X_{Ajk} - \mathrm{logLMA}_k) =$
$\mathrm{Cov}(\log X_{Aik}, \log X_{Ajk}) - \mathrm{Cov}(\log X_{Aik}, \mathrm{logLMA}_k) - \mathrm{Cov}(\log X_{Ajk}, \mathrm{logLMA}_k) +$
$\mathrm{Var}(\mathrm{logLMA}_k)$.

Because $\mathrm{logLMA}_k = \mathrm{logMass}_k - \mathrm{logArea}_k$, by identity (iii):

$\mathrm{Cov}(\log X_{Aik}, \mathrm{logLMA}_k) = \mathrm{Cov}(\log X_{Aik}, \mathrm{logMass}_k) - \mathrm{Cov}(\log X_{Aik}, \mathrm{logArea}_k) = 0$

where the two covariances on the right-hand-side of the expression above are zero, by definition, because traits-*i* and-*j* are assumed to be purely area-proportional. Combining the previous two expressions gives the covariance between two area-proportional traits that are expressed as mass-normalized values:

$\mathrm{Cov}(\log X_{Mik}, \log X_{Mjk}) = \mathrm{Cov}(\log X_{Aik}, \log X_{Ajk}) + \mathrm{Var}(\mathrm{logLMA}_k)$.

Also, by identity (ii):

$\mathrm{Cov}(\log X_{Mik}, \mathrm{logLMA}_k) = \mathrm{Cov}(\log X_{Aik} - \mathrm{logLMA}_k, \mathrm{logLMA}_k) =$
$\mathrm{Cov}(\log X_{Aik}, \mathrm{logLMA}_k) - \mathrm{Var}(\mathrm{logLMA}_k) = -\mathrm{Var}(\mathrm{logLMA}_k)$.

Finally, by identity (iii):

$\mathrm{Cov}(\mathrm{logLL}_k, \log X_{Mjk}) = \mathrm{Cov}(\mathrm{logLL}_k, \log X_{Ajk} - \mathrm{logLMA}_k) =$
$\mathrm{Cov}(\mathrm{logLL}_k, \log X_{Ajk}) - \mathrm{Cov}(\mathrm{logLL}_k, \mathrm{logLMA}_k)$.

$\mathrm{Cov}(\mathrm{logLL}_k, \mathrm{logLMA}_k)$ is approximately proportional to $\mathrm{Var}(\mathrm{LMA}_k)$, because the slope of an OLS linear regression of logLL_k on logLMA_k is close to one (1.1 ± 0.1; see Table S2).

So, to summarize, what happens to the covariance between area-proportional traits if they are normalized by mass instead of area? The covariance between the logarithms of mass-normalized values of two area-proportional traits, *i* and *j*, across all species is: $\mathrm{COV}(\log X_{\mathrm{M}ik}, \log X_{\mathrm{M}jk}) = \mathrm{COV}(\log X_{\mathrm{A}ik}, \log X_{\mathrm{A}jk}) + \mathrm{VAR}(\mathrm{logLMA}_k)$, where COV(·) is covariance and VAR(·) is variance. Thus, mass normalization of area-proportional traits induces positive covariance equal to the variance of logLMA_k. The induced covariance in (*2*) is large because LMA_k ranges from 14 to 1509 g m^{-2} in GLOPNET. Additionally, mass-normalization of an area-proportional trait induces covariance between it and logLMA_k equal to $-\mathrm{VAR}(\mathrm{logLMA}_k)$, and between it and logLL_k equal to $-\mathrm{COV}(\mathrm{logLL}_k, \mathrm{logLMA}_k)$, which is approximately proportional to $-\mathrm{VAR}(\mathrm{logLMA}_k)$.

<u>Estimates of model parameters obtained from mass- vs. area-normalized data</u>
Because GLOPNET reports values that are normalized by area (X_{Aik}) or mass (X_{Mik}), but not the values of area and mass themselves (X_{ik}), we fit the X_{Aik} and X_{Mik} forms of Models-M&A, -C, -M, -A, and -LN. Although the X_{ik}, X_{Aik}, and X_{Mik} model forms are all mathematically identical, they are not expected to yield identical results when applied to GLOPNET because X_{Aik} and X_{Mik} are not inter-converted exactly with LMA_k in this data

set. Nevertheless, fits to area- and mass-normalized data produced very similar results and so we only report the former in the main text.

Posterior probabilities of mass- vs. area-proportionality

The difference between the maximum log-likelihoods, Δ_L, for Models-A and -M (using results from fitting models to area-normalized data) was 435.5 for N, 613.4 for P, 58.9 for Rdark, and 270.2 for Amax (log-likelihoods in Table S1). Thus, assuming equal prior probabilities of Models-M and -A, the posterior probability of the hypothesis that the data are proportional to mass, $1/[1+\exp(\Delta_L)]$, is always less than 10^{-25}.

Estimates from Models-M&A, -C, -M, -A, and -LN

Table S1 contains maximum likelihood estimates and 95% confidence limits for model parameters using area-normalized and mass-normalized data. As noted above, analyses using mass-normalized data and area-normalized yield nearly identical values.

Fraction of each trait that is associated with mass vs. area

As explained in Methods, we accounted for parameter uncertainty and the distribution of LMA values in GLOPNET when estimating the fraction of each trait associated with (or "proportional to") area (p_A). We obtained similar results whether we calculated p_A from Model-M&A (Table S1), Model-C ignoring interspecific variation (Table S1), or Model-C accounting for interspecific variation (Fig. 2). Here, we focus on results from Model-C, unless otherwise noted. Estimated distributions of p_A show that almost all Amax and P is associated with area whether we ignore interspecific variation (Table S1) or account for it (Fig. 2). Although N and Rdark have greater mass-proportional fractions than Amax and P, the majority of species have more than 50% of N and Rdark associated with area (Fig. 2). Similarly, if we ignore interspecific variation, then the majority of N and Rdark are associated with area (Table S1). AIC values (Table S1) show that Model-C is much better than Model-A (which is much better than Model-M) for Amax, Rdark, and N (AIC difference > 25), indicating a mass-associated component that, while small, is clearly statistically significant ($p<0.001$). In contrast, Models-A and -C have similar AICs for P (AIC difference < 5). In summary, Amax, Rdark, N and P are all primarily associated with area, but all but P have a clearly statistically significant mass component.

Model-LN parameter estimates

Table S1 contains maximum likelihood parameter estimates and confidence limits for Model-LN. As explained in Methods, the residuals shown in Fig. 4C, Fig. S4, and Fig. S5 may be obtained from either the area- or mass-normalized forms of Model-LN, or from an OLS regression of either Equation 1 or 2 in the main text.

Predictions from the random data sets generated for Models-A, -M, and -C, assuming no functional correlations among Amax, Rdark, N and P

Figure S1 is the area-normalized counterpart of Figure 1 in the main text. The left three columns of Fig. S1 show the area-normalized LES from GLOPNET. The right three columns show area-normalized predictions from null Model-A. Recall that area-normalized traits are statistically independent from each other in null Model-A. The correlations in the area-normalized GLOPNET data (which are weak compared to the correlations in the mass-normalized GLOPNET data) are due to normalization-

independent correlations (see next section) and the minor proportion of traits that is distributed proportional to mass, as determined by our analysis described above. Area-normalization of the minor mass-proportional trait fractions induces weak correlation for the same reason that mass-normalization of the dominant area-proportional trait fractions induces strong correlation.

Figure S2 shows predictions from null Model-C: mass-normalized in the left three panels and area-normalized in the right three panels. These are similar to the predictions of null Model-A, because traits are distributed primarily proportional to area (i.e., for each trait, null Model-C preserves the higher degree of area- than mass-proportionality estimated from the GLOPNET data, while eliminating functional correlations among traits). Thus, the mass-normalized predictions from null Model-C are similar to the observed mass-normalized LES (left side of main-text Fig. 1) and to the mass-normalized predictions from null Model-A (right side of main-text Fig. 1). Null Model-C also predicts some of the correlation structure in the observed area-normalized LES (Fig. S1, left side; e.g., the relationships between N and LL, and between N and LMA) because area-normalization of the minor mass-proportional trait fractions induces weak correlation (again, for the same reason that mass-normalization of the dominant area-proportional trait fractions induces strong correlation).

Figure S3 shows predictions from null Model-M: mass-normalized in the left three panels and area-normalized in the right three panels. Model-M predicts no correlations among mass-normalized values and strong correlations among area-normalized values, which is unlike the pattern in the data (Fig. 1 and Fig. S1).

Understanding the normalization-independent LES

Model-LN (or, equivalently, Equations 1-2 in the main text) predicts that the variance-covariance matrix of the logarithms of Amax, Rdark, N, and P would be the same within any group of species with the same LMA, no matter what the value of the LMA. That is, within any group of species with the same LMA, a scatter plot of the logarithms of Amax, Rdark, N, and P would reveal a normalization-independent LES that is statistically identical to the one shown by the scatter in figs. S4-S5 (the residuals from the log-scale version of Model-LN; or, equivalently, from main-text Equations 1-2). However, Model-LN also predicts that a constant-LMA group's centroid (mean of the logarithms of Amax, Rdark, N, and P) will move as a function of logLMA along the black lines (arrows) in figs. S4-S5. Consider a pair of area-normalized traits X_{Ai} and X_{Aj}. For a constant-LMA group of species, the centroid of an ellipse representing the normalization-independent scatter in traits i and j would be located at the point ($I_i + S_{Ai} \log_{10}\text{LMA}$, $I_j + S_{Aj} \log_{10}\text{LMA}$) (see log-scale version of Model-LN, defined above; or, equivalently, Equation 1 in the main text, with $S_{Ai} = S_i$). If $\log_{10}$LMA is mean-centered (as in Fig. S4 and Fig. S5), then the centroid would be located at ($S_{Ai}\, \Delta\log_{10}\text{LMA}$, $S_{Aj}\, \Delta\log_{10}\text{LMA}$), where $\Delta\log_{10}$LMA is $\log_{10}$LMA minus the mean $\log_{10}$LMA across all species. Similarly, the corresponding centroid for a pair of mass-normalized traits (Equation 2 in the main text) would be located at the point ($I_i + S_{Mi} \log_{10}\text{LMA}$, $I_j + S_{Mj} \log_{10}\text{LMA}$), with $S_{Mi} = (S_{Ai} - 1)$; or, for mean-centered $\log_{10}$LMA, at the point ($S_{Mi}\, \Delta\log_{10}\text{LMA}$, $S_{Mj}\, \Delta\log_{10}\text{LMA}$).

Because the maximum likelihood parameter estimate for S_{Ai} for Model-LN are relatively close to zero for Amax, Rdark, N and P, the centroid will not move much along the axes for these traits as $\log_{10}$LMA changes if data are area-normalized (estimates in

Table S1). However, the movement will be larger if the data are mass-normalized because S_{Mi}, which equals $S_{Ai} - 1$, is relatively close to −1. This explains why the area-normalized LES (Fig. S1) looks more like the normalization-independent LES (Fig. S4 and Fig. S5) than does the mass-normalized LES (Fig. 1).

Figures S4-S5 show how mass-normalization strongly alters the normalization-independent LES, whereas area-normalization mildly changes it. These figures generalize the information in Fig. 4 of the main text to other combinations of traits. In the figures, the blue ellipses are isodensity ellipses obtained from the variance-covariance matrix of the residuals from main-text Equations 1 or 2 for each pair of traits. These ellipses are estimated to contain 95% of species in GLOPNET. The red ellipses are the corresponding quantities for the raw GLOPNET data expressed as $\log_{10}$-trait minus $\log_{10}$-trait mean. We express the raw data as mean-centered values in $\log_{10}$-space so that they are centered on the same mean as the normalization-independent residuals (0, 0). Thus, each blue ellipse shows the pattern of scatter predicted for a group of species with the same $\log_{10}$LMA. The blue ellipse in the center of each graph in Fig. S5 corresponds to the mean $\log_{10}$LMA. The two flanking blue ellipses in Fig. S5 correspond to the mean plus/minus two standard deviations of $\log_{10}$LMA. Graphs on the left hand side are for area-normalized data. Graphs on the right hand side are for mass-normalized data. The three blue ellipses in each graph in Fig. S5 are arrayed along the lines given by the above formulae for centroid locations. Note the large distance between centroids of the blue ellipses if data are mass-normalized (because S_{Mi} is close to −1 for Amax, Rdark, N and P), but the small distances if data are area-normalized (because S_{Ai} is close to zero).

Again, the red ellipses in figs. S4-S5 show the 95% isodensity ellipse for the $\log_{10}$-transformed, normalized GLOPNET data for a given pair of traits. These red ellipses nearly encompass the three blue ellipses in each graph in Fig. S5 because the data contain the full range of values of logLMA. One can also think of changes in logLMA as stretching the central blue ellipse into the red ellipse along the lines given by the formulae above for centroid locations. A key point is that most of the structure in the mass-normalized LES reflects variation in LMA, rather than correlations between trait values (normalized or not) among leaves with similar LMA.

In summary, normalization stretches the centroids of the different constant-LMA groups along the black lines shown in Fig. S4, approximately from one arrowhead to the other, and thus alters the normalization-independent correlations. The greater alteration induced by mass-normalization (compared to area-normalization) is reflected in empirical estimates of S (see main-text Equations 1-2), which are closer to zero than to one (Table S1). The lines are typically aligned along the long axis of the mass-normalized red ellipses, because LMA-induced covariance dominates the mass-normalized LES.

We now present additional details that provide further insights into the normalization-independent LES described by Equations 1-2 in main text and Model-LN. As explained above, the blue ellipses in Fig. S4 show the pattern of interspecific scatter expected around the trait means for a group of species with the same LMA (no matter what the LMA). If we were to plot the actual values of two log-transformed traits a and b expected within a constant-LMA group in Fig. S4, instead of the residuals around the means for the group, then the blue ellipse would still have the same shape and size, but its centroid (mean) would no longer be at (0, 0). Instead the center of the blue ellipse would be at the point ($S_a\Delta\log_{10}$LMA, $S_b\Delta\log_{10}$LMA) if area-normalized and

[$(S_a - 1)\Delta\log_{10}LMA$, $(S_b - 1)\Delta\log_{10}LMA$] if mass-normalized, where $\Delta\log_{10}LMA$ is the logarithm of the group's LMA minus the mean of $\log_{10}LMA$ across all species. The centroids of blue ellipses for different constant-LMA groups will be arranged along a line shown by the black line in each panel of Fig. S4. The flanking blue ellipses in Fig. S5, for example, are for constant-LMA groups with $\Delta\log_{10}LMA = \pm 2\sigma_{LMA}$, where σ_{LMA} is the observed standard deviation of $\log_{10}LMA$. The slope of the black line in each panel in Fig. S4 and Fig. S5 is $(S_a - 1)/(S_b - 1)$ for mass-normalized traits and S_a/S_b for area-normalized traits. The distance between the arrowheads on each line encompasses the centroids of constant-LMA groups containing 95% of $\log_{10}LMA$ values (i.e., $\Delta\log_{10}LMA = \pm 2\sigma_{LMA}$). The formulae for the inter-arrowhead distance is:

$$4\sigma_{LMA}\sqrt{(S_a - 1)^2 + (S_b - 1)^2}$$

for mass-normalized traits and:

$$4\sigma_{LMA}\sqrt{S_a^2 + S_b^2}$$

for area-normalized traits.

The correlation predicted by Model-LN between two log-transformed traits (*a* and *b*) is:

$$Correlation_{Mass} = \frac{(1 - S_{Aa})(1 - S_{Ab})\sigma_{LMA}^2 + Cov_{ab}}{\sqrt{\left[(1 - S_{Aa})^2 \sigma_{LMA}^2 + Var_a\right]\left[(1 - S_{Ab})^2 \sigma_{LMA}^2 + Var_b\right]}}$$

if data are mass-normalized, and:

$$Correlation_{Area} = \frac{S_{Aa}S_{Ab}\sigma_{LMA}^2 + Cov_{ab}}{\sqrt{\left(S_{Aa}^2\sigma_{LMA}^2 + Var_a\right)\left(S_{Ab}^2\sigma_{LMA}^2 + Var_b\right)}}$$

if data are area-normalized, where S_{Aa} and S_{Ab} correspond to S_i in Equations 1-2 in the main text for traits-*a* and -*b*; Cov_{ab}, Var_a and Var_b are the covariances and variances of the residuals from Equations 1-2 in main text for traits-*a* and -*b*; and σ^2_{LMA} is the variance of $\log_{10}LMA$. In contrast the normalization-independent correlation is:

$$Correlation_{Ind} = \frac{Cov_{ab}}{\sqrt{Var_a Var_b}}$$

If the traits are purely proportional to area (S_{Aa}=S_{Ab}=0), then:

$$Correlation_{Mass} = \frac{\sigma^2_{LMA} + Cov_{ab}}{\sqrt{\left[\sigma^2_{LMA} + Var_a\right]\left[\sigma^2_{LMA} + Var_b\right]}}$$

which approaches one as σ^2_{LMA} becomes large and:

$$Correlation_{Area} = \frac{Cov_{ab}}{\sqrt{Var_a Var_b}} = Correlation_{Ind}$$

Thus, if traits are proportional to area, mass-normalization stretches the data towards a correlation of one, whereas area-normalization reveals the correlations of the normalization-independent LES. The differences between $Correlation_{Mass}$ and $Correlation_{Ind}$ are shown in Table S4.

The preceding analysis of correlations is easily extended to an analysis of OLS regression slopes. The OLS slope of log-transformed trait-*b* as a function of log-transformed trait-*a* is:

$$OLS\ slope\ area-normalized = \frac{S_{Aa}S_{Ab}\sigma^2_{LMA} + Cov_{ab}}{S^2_{Aa}\sigma^2_{LMA} + Var_a}$$

if the data are area-normalized,

$$OLS\ slope\ mass-normalized = \frac{(1-S_{Aa})(1-S_{Ab})\sigma^2_{LMA} + Cov_{ab}}{(1-S_{Aa})^2\sigma^2_{LMA} + Var_a}$$

if the data are mass-normalized, and

$$OLS\ slope\ size-independent = \frac{Cov_{ab}}{Var_a}$$

if the regression involves normalization-independent residuals. Three examples illustrate the significance of these formulae:

(1) Trait-*a* is Amax or P, so that S_{Aa} is approximately zero in GLOPNET. Then the slope for area-normalization is similar to the normalization-independent slope (identical if $S_{Aa} = 0$), and the mass-normalized slope is:

$$\frac{(1-S_{Ab})\sigma^2_{LMA} + Cov_{ab}}{\sigma^2_{LMA} + Var_a}$$

which will tend toward 1 as the LMA variance increases, assuming that S_{Ab} is small (as it is for most traits in GLOPNET).

(2) Trait-*a* is N or Rdark, and trait-*b* is Amax or P. Because S_A for Amax and P is approximately zero in GLOPNET, the, the area-normalized slope is approximately:

$$\frac{Cov_{ab}}{S_{Aa}^2\sigma_{LMA}^2 + Var_a}$$

which is less than the normalization-independent slope. The mass-normalized slope is approximately:

$$\frac{(1-S_{Aa})\sigma_{LMA}^2 + Cov_{ab}}{(1-S_{Aa})^2\sigma_{LMA}^2 + Var_a}$$

which will tend toward 1as the LMA increases, assuming S_{Ab} is small (as it is for most traits in GLOPNET).

(3) Trait-*a* is N or Rdark and trait-*b* is N or Rdark. In this case S_{Aa} and S_{Ab} are small but significantly greater than zero, so the area-normalized slope is the basic formula:

$$\frac{S_{Aa}S_{Ab}\sigma_{LMA}^2 + Cov_{ab}}{S_{Aa}^2\sigma_{LMA}^2 + Var_a}$$

which could be larger or smaller than the normalization-independent slope depending on the parameter values. The mass normalized slope is again the basic formula:

$$\frac{(1-S_{Aa})(1-S_{Ab})\sigma_{LMA}^2 + Cov_{ab}}{(1-S_{Aa})^2\sigma_{LMA}^2 + Var_a}$$

which will tend toward 1 as the LMA variance increases, assuming S_{Aa} and S_{Ab} are small.

<u>When should mass-normalized, area-normalized, or normalization-independent values be used?</u>

The appropriateness of normalized data or normalization-independent values (residuals from Equations 1 or 2 in Main Text) depends on the application. Normalization-independent values should be used to quantify interspecific trait relationships while controlling for LMA (equivalent to quantifying relationships among species that share the same LMA), because relationships between mass-normalized traits or between area-normalized traits are affected by hidden changes in LMA. Mass-normalized traits may be appropriate for (1) calculating the amount of time required for a leaf to return through photosynthesis the amount of carbon required for its construction (but see Cost Model analysis of annualized construction costs below); or (2) expressions of nutrient investment tradeoffs, such as N/mass required to make a leaf of a particular LMA or LL. Additionally, instantaneous carbon gain for a light-saturated plant can be estimated by multiplying Amax/mass by total leaf mass. Area-normalized quantities are useful for understanding gas exchange and nutrient investment for a given area of light-intercepting leaf surface. For some applications, either mass- or area-normalized traits will yield the same results. For example, either mass- or area-normalized quantities can be used to

estimate ecosystem nutrient stocks: N per m^2 ground area is equal to N/mass multiplied by total leaf mass per m^2 ground area, or N/area multiplied by total leaf area per m^2 ground area.

For analyses of *inter*specific trait relationships, area-normalized (as opposed to mass-normalized) traits will typically yield results closer to the normalization-independent analysis (Table S2, SOM section entitled "Pairwise trait relationships…."). In contrast, for *intra*specific analyses across contrasting light environments, mass-normalized traits are probably to be preferred, because per-unit-mass trait values remain roughly constant as LMA decreases from sun to shade (*13, 16-18*).

Although mass-normalized quantities may be appropriate for some economic (cost-benefit) analyses, we emphasize that leaf mass may not always be an appropriate index for construction costs, depending on the contrasts in question. In many contexts LL should be considered in economic analyses, because the cost of an object declines (in some sense) with its longevity. Does Amax/mass or Amax/area provide the better benefit-cost ratio when costs are annualized? In other words, is a leaf's mass or its area a better measure of its annualized cost? A leaf's annualized carbon construction cost is proportional to mass/LL. Because LL and LMA are roughly proportional to each other (OLS slope of logLL on logLMA is 1.11 ± 0.05 s.e.), mass/LL is roughly proportional to mass/LMA, which is equal to area. This rough analysis suggests that when comparing across species, Amax/area, not Amax/mass, is a better benefit-cost ratio when costs are annualized. To evaluate this question quantitatively, we tested the alternative hypotheses that mass/LL is proportional to mass or area by considering two alternative models that are analogous to Model-M and Model-A above:

(Cost-model-M) $Mass_k LL_k^{-1} = C_M Mass_k \varepsilon_{Mk}$

which reduces to: $LL_k^{-1} = C_M \ \varepsilon_{Mk}$, and:

(Cost-model-A) $Mass_k LL_k^{-1} = C_A Area_k \varepsilon_{Ak}$

which reduces to: $LL_k^{-1} = C_A \ LMA_k^{-1} \ \varepsilon_{Ak}$

where the subscript k refers to species identity; C_M and C_A are mass and area proportionality constants, respectively; and ε_{Mk} and ε_{Ak} are species specific, unitless, unit-mean lognormal random variables that represent interspecific variation in per-unit-mass and per-unit-area construction costs, respectively. For Cost-model-M, D_{ik} is LL_k^{-1} and $\mu_i(LMA_k)$ is C_M. For Cost-model-A, D_{ik} is again LL_k^{-1} and $\mu_i(LMA_k)$ is $C_A \, LMA_k^{-1}$. The posterior probabilities of Cost-models-M and -A are calculated as above for Models-M and -A: the posterior probability of Cost-model-M is $1/[1+\exp(\Delta_L)]$, where Δ_L is the log-likelihood of Cost-model-A evaluated at its maximum likelihood parameter estimates minus the corresponding quantity for Cost-model-M. The difference between the maximum log-likelihoods, Δ_L, for Cost-models-A and -M is 305.9, and so, assuming equal prior probabilities, the posterior probability of the hypothesis that mass/LL is proportional to mass instead of area ($1/[1+\exp(\Delta_L)$) is less than 10^{-133}. Thus, Amax/area is a much better benefit-cost ratio for a leaf's annual carbon budget than Amax/mass.

Statistics for the major axis (MA) regression slopes in Figure 4 of the main text

A test of homogeneity of slopes indicates that the major axis slopes of the 5 colored LMA-bins in Fig. 4A are (i) not significantly different from one another (p-value = 0.487, Likelihood ratio statistic = 3.438, df = 4); (ii) not significantly different from the pooled-data (black line) slope of 0.458 in Fig. 4C (p-value = 0.619, Likelihood ratio statistic = 3.529, df = 5); and (iii) significantly more shallow than the all-data (black line) slope of 0.621 in Fig. 4A (p-value < 0.00001, Likelihood ratio statistic = 71.15, df = 5). Similarly, the major axis slopes of the 5 bins in Fig. 4B are (i) not significantly different from one another (p-value = .135, Likelihood ratio statistic = 7.025, df = 4); (ii) not significantly different from the all-data (black line) slope of 0.458 in Fig. 4C (p-value = 0.179, Likelihood ratio statistic = 7.606, df = 5); and (iii) significantly more shallow than the all-data (black line) slope of 0.555 in Fig. 4B (p-value < 0.00001, Likelihood ratio statistic = 35.97, df = 5). The above tests were performed with the "sma" function in the SMATR package in R (*22*).

Pairwise trait relationships in the area- and mass-normalized LES and the normalization-independent LES

All regression statistics and other statistics characterizing bivariate trait relationships using area- or mass-normalized data or the normalization-independent LES values were calculated using the statistics package R (*20*). Table S2 contains the full Table of all statistics characterizing the pairwise relationships. Ordinary least squares (OLS) regressions characterizing mass-, area- and normalization-independent pairwise relationships were fit and their slopes compared to one another using a linear model (function "lm" in R). Similarly, major axis (MA) and standardized major axis (SMA) regressions were fit and their slopes compared to one another using the likelihood ratio statistic using the "sma" function in the "smatr" package for R (*22*). Here, as in the main text, the normalization-independent LES is defined as the residuals from the Model-LN fits to $\log_{10}$-transformed trait data.

Here we evaluate whether significant differences exist in the slope of bivariate trait relationships calculated using $\log_{10}$-transformed area- or mass-normalized data or normalization-independent values. After a brief summary of the results of these slopes comparison tests for relationships of particular interest, we present slopes comparison results for OLS and SMA regression for all pairwise relationships between traits LL, N, P, Rdark and Amax. Unless otherwise noted, all slope comparisons discussed here in the summary, between the mass- and area-normalized and normalization-independent LES, are from ordinary least squares (OLS) regression. The trait relationships here are labeled as "Y vs. X," where Y is the dependent variable and X is the independent variable.

For N vs. P, the mass-normalized slope is significantly steeper than both the normalization-independent and the area-normalized slope, and the normalization-independent slope is not significantly different from the area-normalized slope. Using standardized major axis regression (SMA), the well-discussed mass-normalized slope of 2/3 (*15*) is significantly steeper than the normalization-independent slope of 0.59, meaning that across the global flora, leaves accrue proportionately less N for a given level of P than would be expected from metabolic scaling theory.

The normalization-independent spectrum indicates that a robust relationship exists between Amax and N, and the slope of this relationship is intermediate between those

predicted from the area- and mass-normalized LES. For Amax vs P, the mass-normalized slope is significantly steeper than both the area-normalized and normalization-independent slopes, which are not significantly different from one another. The same is true for Amax vs Rdark, Rdark vs N, and Rdark vs P. Therefore, the area-normalized slope is an excellent statistical approximation of the normalization-independent relationships between Amax and P or Rdark; and Rdark vs N or P.

For all relationships between a trait and LL, the mass-normalized and normalization-independent slopes are negative, the mass-normalized slope significantly more steeply so for N, Amax and Rdark. The area-normalized slopes of both the P vs LL and Rdark vs LL relationships are not significantly different from zero and are significantly different from either the normalization-independent or mass-normalized slopes. The area-normalized slope of N vs LL is slightly significantly greater than zero. The area-normalized slope of Amax vs LL is negative, and the magnitude of the normalization-independent slope is intermediate between those of the area- and mass-normalized negative slopes.

Below we present slopes comparisons results between the area- and mass-normalized and normalization-independent LES from both OLS and SMA regressions. The slopes in question are printed in Table S2 and are also printed in the relevant text here with the 95% confidence limits in parentheses. Detailed results of slopes comparison tests beyond those presented here available upon request.

Y = N, X = P:

SMA: The mass- (0.663 (0.638, 0.689)) and area-normalized (0.705 (0.664, 0.748)) SMA slopes are marginally not significantly different form one another (test statistic = 2.908, p = 0.089). Both the area- (test statistic = 3.87e-05, p = 16.934) and mass-normalized (test statistic = 0.001, p = 10.2) SMA slopes, however, are significantly different from the normalization-independent SMA slope (0.594 (0.561, 0.628)).

OLS: Area- (0.404 (0.362, 0.446)) and mass-normalized (0.566 (0.541, 0.591)) OLS slopes are significantly different from one another (df = 2193, t = 6.995, p = 3.5e-12). The normalization-independent OLS slope (0.376 (0.342, 0.409)) is not significantly different from the area-normalized slope (t = -1.032, p = 0.302). The normalization-independent slope is, however, significantly different from the mass-normalized slope (t = 8.193, p = 4.28e-16). The mass-normalized slope is significantly steeper than both the area-normalized slope and the normalization-independent slope.

MA: The area-normalized, mass-normalized and normalization-independent MA slopes are all significantly different form one another. The area- and mass-normalized slopes are both significantly steeper than the normalization-independent slope.

Y = A, X = N:

SMA: The area-normalized SMA slope (1.221 (1.14, 1.308)) is significantly different from the mass-normalized slope (1.71 (1.625, 1.799)) (test statistic = 59.1, p = 1.47e-14). Both the area- (test statistic = 22.36751, p = 2.2515e-06) and mass-normalized (test statistic = 6.167, p = 0.013) SMA slopes are significantly different from the normalization-independent SMA slope (1.538 (1.439, 1.644)).

OLS: The mass-normalized OLS slope (1.245 (1.158, 1.332)) is significantly different from the area-normalized slope (0.454 (0.37, 0.538)) (df = 2109, t = 13.149, p <

< 2e-16). The normalization-independent OLS slope (0.679 (0.577, 0.781)) is significantly different from the area-normalized slope (t = 3.213, p = 0.001). The mass-normalized OLS slope is significantly different from the normalization-independent slope (t = 8.229, p = 3.25e-16). The normalization-independent OLS slope is intermediate between the mass-normalized slope (the steepest slope) and the area-normalized slope (most shallow).

Y = A, X = P:

SMA: The area-normalized SMA slope (0.73 (0.638, 0.834))is significantly different from the mass-normalized SMA slope (1.016 (0.897, 1.152)) (test statistic = 12.5, p = 0.0004). The area-normalized SMA slope is not significantly different from the normalization-independent SMA slope (0.738 (0.645, 0.845)) (test statistic = 0.014, p = 0.905). The mass-normalized SMA slope is significantly different from the normalization-independent slope (test statistic = 11.6, p = 0.0007).

OLS: The mass-normalized OLS slope (0.421 (0.293, 0.548)) is significantly different from the area-normalized slope (0.154 (0.056, 0.253)) (t = 3.42, p = 0.0007). The normalization-independent OLS slope (0.14 (0.04, 0.24)) is not significantly different from the area-normalized slope (df = 615, t = -0.182, p = 0.856). The mass-normalized OLS slope is significantly different from the normalization-independent slope (t = 3.57, p = 0.0004). The mass-normalized slope is significantly steeper than both the area-normalized and normalization-independent OLS slopes.

Y = A, X = Rd:

SMA: The area-normalized SMA slope (0.93 (0.832, 1.038)) is significantly different from the mass-normalized SMA slope (1.177 (1.088, 1.274)) (test statistic = 11.6, p = 0.0007). The area-normalized SMA slope is not significantly different from the normalization-independent SMA slope (0.986 (0.884, 1.099)) (test statistic = 0.551, p = 0.457). The mass-normalized SMA slope is significantly different from the normalization-independent slope (test statistic = 6.692, p = 0.010).

OLS: The mass-normalized OLS slope (0.901 (0.808, 0.994)) is significantly different from the area-normalized OLS slope (0.401 (0.298, 0.504)) (df = 771, t = 7.162, p = 1.86e-12). The normalization-independent OLS slope (0.45 (0.342, 0.558)) is not significantly different from area-normalized OLS slope (t = 0.611, p = 0.541). The mass-normalized slope is significantly different from the normalization-independent OLS slope (t =6.225 , p = 7.89e-10). The mass-normalized slope is significantly steeper than both the normalization-independent and area-normalized OLS slopes.

Y = Rd, X = P:

SMA: The area-normalized SMA slope (0.93 (0.832, 1.038)) is not significantly different from the mass-normalized slope (1.177 (1.088, 1.274)) (test stat = 2.07, p = 0.150). The area-normalized SMA slope is not significantly different from the normalization-independent slope (0.986 (0.884, 1.099)) (test stat = 0.09, p = 0.765), nor is the mass-normalized SMA slope significantly different from the normalization-independent slope (test stat = 2.90, p = 0.09).

OLS: The area-normalized OLS slope (0.402 (0.246, 0.558)) is not significantly different from the mass-normalized slope (0.563 (0.385, 0.742)) (t = 1.382, p = 0.168).

The normalization-independent slope (0.313 (0.155, 0.47)) is not significantly different form the area-normalized slope (t = -0.783, p = .437). The mass-normalized OLS slope is significantly steeper than the normalization-independent slope (t = 2.128, p = 0.035).

Y = Rd, X = N:

SMA: The area-normalized SMA slope (1.129 (1.023, 1.246)) is significantly different from both the mass-normalized (1.431 (1.32, 1.552)) (test stat = 13.3, p = 0.0002) and normalization-independent (1.371 (1.233, 1.524)) (test stat = 6.9, p =0.009) SMA slopes. The mass-normalized slope is not significantly different from the normalization-independent slope (test stat = 0.403, p = 0.525).

OLS: The area- (0.654 (0.542, 0.765)) and mass-normalized (1.063 (0.947, 1.179)) OLS slopes are significantly different (t = 5.03, p = 6.14e-07) from one another. The area-normalized slope is not significantly different from the normalization-independent slope (0.655 (0.509, 0.801)) (t = 0.015, p = 0.988). The mass-normalized slope is significantly different from the normalization-independent slope (t = 4.32, p = 1.73e-05). The mass-normalized OLS slope is significantly steeper than both the area-normalized and normalization-independent OLS slopes.

Y = N, X = LL:

SMA: There is no significant difference between any combination of the area-normalized (0.425 (0.394, 0.458)), mass-normalized (-0.444 (-0.47, -0.42)) and normalization-independent (-0.45 (-0.484, -0.419)) SMA slopes.

OLS: The area-normalized OLS slope (0.086 (0.054, 0.118)) is significantly different from both the mass-normalized slope (-0.3 (-0.325, -0.275)) (t = -19.7, p < 2e-16) and the normalization-independent slope (-0.159 (-0.191, -0.127)) (t = -10.5, p < 2e-16). The mass-normalized slope is significantly different from the normalization-independent slope (t = -6.6, p = 1.62e-09). The mass-normalized slope is significantly more steeply negative than the normalization-independent slope, and the area-normalized slope is positive.

Y = P, X = LL:

SMA: The area-normalized SMA slope (0.901 (0.785, 1.036)) is not significantly different from the mass-normalized SMA slope (-0.96 (-1.083, -0.851)). The area-normalized SMA slope is significantly different from the normalization-independent slope (-1.144 (-1.304, -1.003)), and the mass-normalized SMA slope is marginally not significantly different from the normalization-independent slope (p = 0.053).

OLS: The area-normalized OLS slope (0.011 (-0.115, 0.137)) is significantly different from both the mass-normalized (-0.481 (-0.597, -0.365)) and normalization-independent slopes (-0.384 (-0.534, -0.234)). The mass-normalized OLS slope is not significantly different from the normalization-independent OLS slope. The area-normalized OLS slope is not significantly different from zero.

Y = A, X = LL:

SMA: The area-normalized (-0.471 (-0.511, -0.434)), mass-normalized (-0.728 (-0.765, -0.693)) and normalization-independent (-0.66 (-0.707, -0.616)) SMA slopes are all significantly different form one another.

OLS: The area-normalized OLS slope (-0.169 (-0.207, -0.13)) is significantly different from both the mass-normalized (-0.603 (-0.638, -0.567)) and normalization-independent (-0.409 (-0.454, -0.363)) OLS slopes. The mass-normalized OLS slope is significantly different from the normalization-independent OLS slope. Mass-normalized, area-normalized and normalization-independent OLS slopes are all negative, and the normalization-independent slope is intermediate between the area- and mass-normalized slopes.

Y = Rd, X = LL:

SMA: The area-normalized SMA slope (-0.47 (-0.538, -0.412)) is significantly different from both the mass-normalized (-0.6 (-0.653, -0.551)) and normalization-independent (-0.627 (-0.703, -0.56)) SMA slopes, but the mass-normalized SMA slope is not significantly different from the normalization-independent SMA slope.

OLS: The area-normalized (-0.04 (-0.103, 0.023)), mass-normalized (-0.464 (-0.515, -0.413)) and normalization-independent (-0.331 (-0.403, -0.26)) OLS slopes are all significantly different form one another. The area-normalized slope is not significantly different from zero, and the mass-normalized slope is significantly more steeply negative than the normalization-independent slope.

<u>Principal Component Analyses: comparing the area- and mass-normalized LES and the normalization-independent LES</u>

Wright et al. (*2*) used principal components analysis (PCA) to quantify multivariate relationships among the six GLOPNET traits: Amax, Rdark, N, P, LL, and LMA. The first principal component axis (PC1) explains 74% and 43%, respectively, of the variation in the mass- and area-normalized log-transformed data (*2*). The high percent variance explained by PC1 of the mass-normalized data, along with the apparently clear interpretation of this axis (leaves with high Amax, Rdark, N and P, and low LMA and LL at one end of the axis; leaves with the opposite properties at the other end), led to the notion of the Leaf Economics Spectrum, ranging from leaves with fast returns on investments to leaves with slow returns (*2*). However, as we have shown, the tight correlations between two mass-normalized traits, or between a mass-normalized trait and LMA or LL, largely reflect normalization-induced covariances (i.e., the correlation between LMA and itself; see SOM section above, "Effects of normalization on covariances"). These results imply that the high percent variance explained by PC1 of the mass-normalized LES is partially due to normalization-induced covariance.

Here, we re-evaluate the multivariate trait relationships reported by Wright et al. (*2*). We performed PCA on the normalization-independent residuals from Model-LN (or, equivalently, from Equations 1-2 in the main text) to quantify normalization-independent multivariate trait relationships. We compare these results to PCAs of mass- and area-normalized data. We focus on the following questions: How much variation is explained by PC1 of the normalization-independent LES relative to PC1 obtained from mass- and area-normalized data? What is the biological interpretation of PC1 of the normalization-independent LES?

Table S3 reports PCAs for the following (all variables are $\log_{10}$-transformed):

1. Mass- (Table S3a) and area-normalized (Table S3b) data for all six variables (LMA, LL, N, P, Rdark and Amax). These replicate PCAs reported in (*2*) and are

reproduced here for convenience. Dividing each of the loadings reported in (*2*) by its component length yields the unit vector representation reported here.

2. Mass- (Table S3c) and area-normalized (Table S3d) data for LL, N, P, Rdark and Amax, but with LMA excluded from the analysis. These PCAs were performed for direct comparison with the version of the normalization-independent PCA that excluded LMA (see below).
3. Normalization-independent residuals for N, P, Rdark and Amax combined with each of the following:
    a. Residuals of LL from OLS regression on LMA (Table S3e; here, LL is treated exactly like N, P, Rdark, and Amax in the normalization-independent LES). LMA is not included in this PCA, because it is by definition orthogonal to the residuals for the other five traits.
    b. LL values (not residuals) and LMA (Table S3f).
    c. LL residuals and LMA (Table S3g). As noted above, LMA is orthogonal to the residuals for the other five traits. This PCA is reported primarily for the heuristic purpose of clarifying the meaning of the normalization-independent LES, and to show that our qualitative results do not depend on how we treat LL and LMA in PCA.

All three of the normalization-independent PCAs we report are derived from mass-normalized data for Amax, Rdark, N, and P. As noted elsewhere, the normalization-independent LES is mathematically identical whether calculated from mass- or area-normalized data for single leaves. However, the quantitative results obtained from mass- vs. area-normalized GLOPNET data differ slightly because mass- and area-normalized values are not inter-converted exactly using species mean LMA values reported in GLOPNET. The normalization-independent PCA obtained from area-normalized data (not reported) is quantitatively very similar to its mass-normalized counterpart reported here, with the area-normalized PC1 explaining slightly less variance than the mass-normalized PC1.

Removing LMA from the PCAs of mass- and area-normalized data has little effect on the explanatory power of PC1 (which still accounts for 75% of the total variance in the mass-normalized data and 47% of that of the area-normalized data) and also little effect on the loadings (compare Tables S3c-d to Tables S3a-b). The minimal effect on these PCAs of removing LMA suggests that multivariate associations in the mass- and area-normalized data do not depend on LMA per se and can be simply explained by three things: (i) the positive relationship between LL and LMA; (ii) normalization-induced correlations; and (iii) normalization-independent relationships.

In the normalization-independent PCA with LL residuals and LMA excluded (Table S3e), PC1 explained 58% of the variance, which is intermediate between the variance explained by PC1 for the mass- and area-normalized data (either with or without LMA included; see Tables S3a-d). The loadings of each trait on this normalization-independent PC1 (Table S3e) are similar to those for the mass-normalized PC1 (Tables S3a and S3c): Amax, Rdark, N and P have similar loadings to each other, and these loadings are opposite in sign to that for LL. When LL values (as opposed to residuals from regression on LMA) and LMA are included in the normalization-independent PCA (Table S3f), PC1 explains 44% of the variance in the data (similar to the area-normalized PCA, Table S3b), the PC1 loadings for all but LMA are similar to those in the mass-

normalized PCA (Table S3a), and PC2 (26% of variance) is primarily a LMA-LL axis. Finally, when LL residuals are included with LMA in the normalization-independent PCA (Table S3g), LMA has a loading of 0 on all PCs except for PC2 (where it has a loading of 1), and the other five traits have identical loadings as in the case where LMA is excluded (Table S3e). This third version of the normalization-independent PCA illustrates how LMA and LMA-independent variation can be treated separately. In this case, the LMA-independent PC1 explains 49% of the variation in the data (similar to the area-normalized PC1, Table S3b), and LMA explains 1/6 of the variation in the data (as it must, since it is orthogonal to the other five variables).

In summary, as in the mass-normalized LES described by Wright et al. (*2*), the dominant axis of variation in all three versions of the normalization-independent LES reported here (Tables S3e-g) describes a spectrum from leaves with fast economic returns to leaves with slow returns. However, in the normalization-independent LES, this spectrum occurs at any given LMA, whereas in the mass-normalized LES described by Wright et al. (*2*), LMA varies along the dominant axis. Furthermore, our analyses are inconsistent with the conclusion of Wright et al. (*2*) that global trait variation can be accurately summarized by a single dominant axis of variation because (i) the normalization-independent PC1 is either orthogonal to LMA (Tables S3e and g) or nearly so (Table S3f), which implies either a separate LMA axis (PC2 in Table S3g) or a separate LMA-LL axis (PC2 in Table S3f); and (ii) the more modest percent variation explained by PC1 in the normalization-independent compared to the mass-normalized PCA (e.g., 58% in Table S3e compared to 75% in Table S3c) suggests that there are additional, important axes of variation that are altered by normalization-induced correlations.

Comparisons of Models-LN and -C

Models-C and -M&A have similar mathematical forms and similar parameter estimates. Here, we focus on comparing Models-LN and -C. Although Models-LN and -C have different functional forms, the two models, when fit to data, are quantitatively similar. This implies that the assumption that the data are lognormally distributed is, in practical terms, compatible with the assumption that Amax, Rdark, N and P can each be partitioned into two fractions, one proportional to mass and the other proportional to area. When evaluated at their maximum likelihood estimates, the two models have very similar predictions for each trait's "LMA profile", which shows how much new trait is added to a unit of leaf area for each 10 g/m^2 increase in LMA, assuming the area-proportional component is contained in the first 10 g/m^2 (Fig. S6). (LMA ranges in GLOPNET from roughly 10 g/m^2 to 1500 g/m^2.) Although none of our analyses make any explicit assumptions about the physical location of traits within leaves, it is nevertheless heuristically useful to think of an LMA profile as a depth profile giving the concentration of a trait as a function of depth beneath the leaf's surface. Both models imply that traits are distributed approximately per unit area, because the LMA profiles drop so precipitously (Fig. S6). The primary difference is that Model-LN predicts a steep but smooth drop, while Model-C assumes a step function.

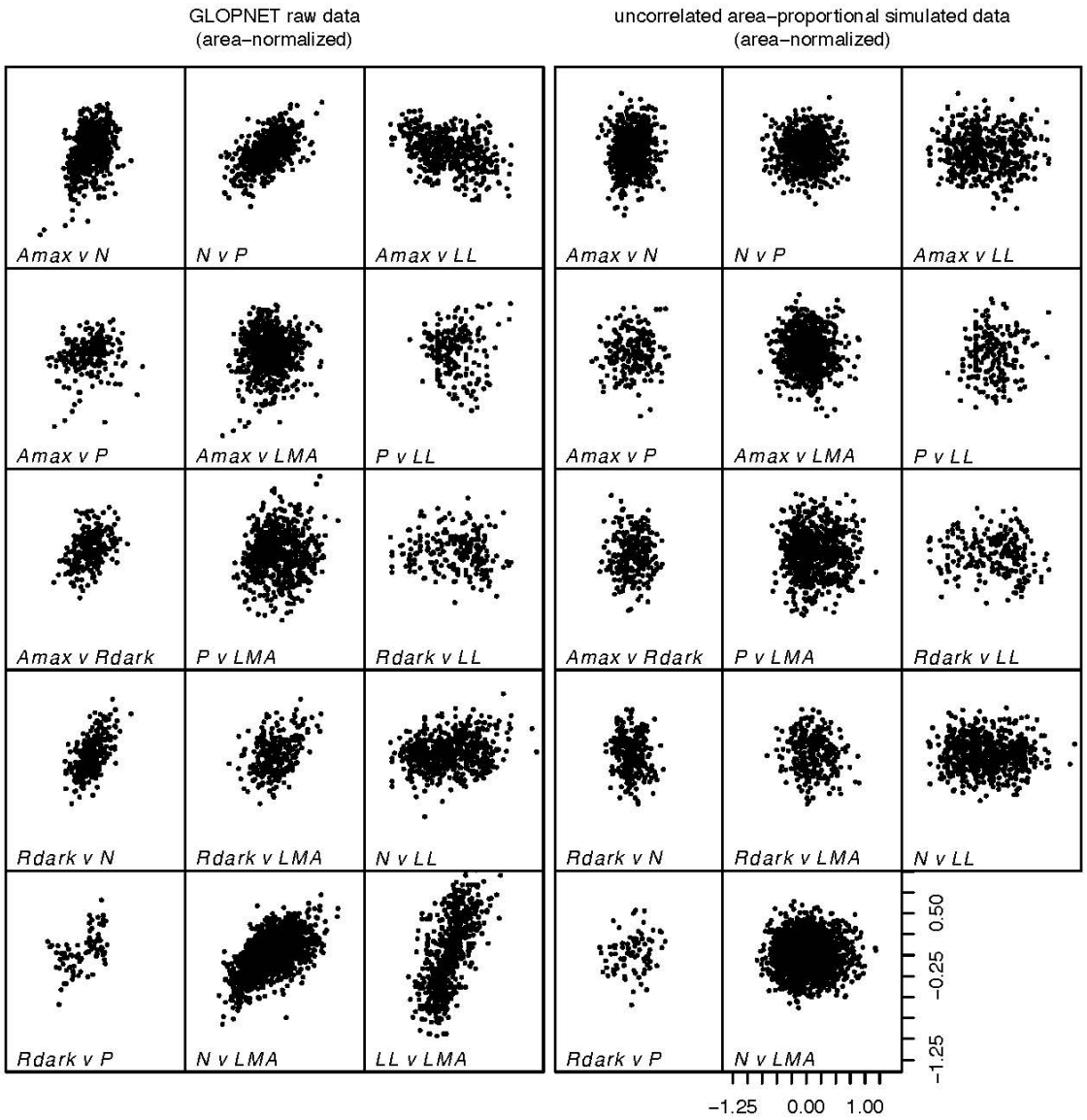

Fig. S1. The area-normalized Leaf Economics Spectrum (left three columns) and area-normalized predictions of null Model-A (right three columns). The plots are all $\log_{10}$-$\log_{10}$, with means rescaled to (0,0).

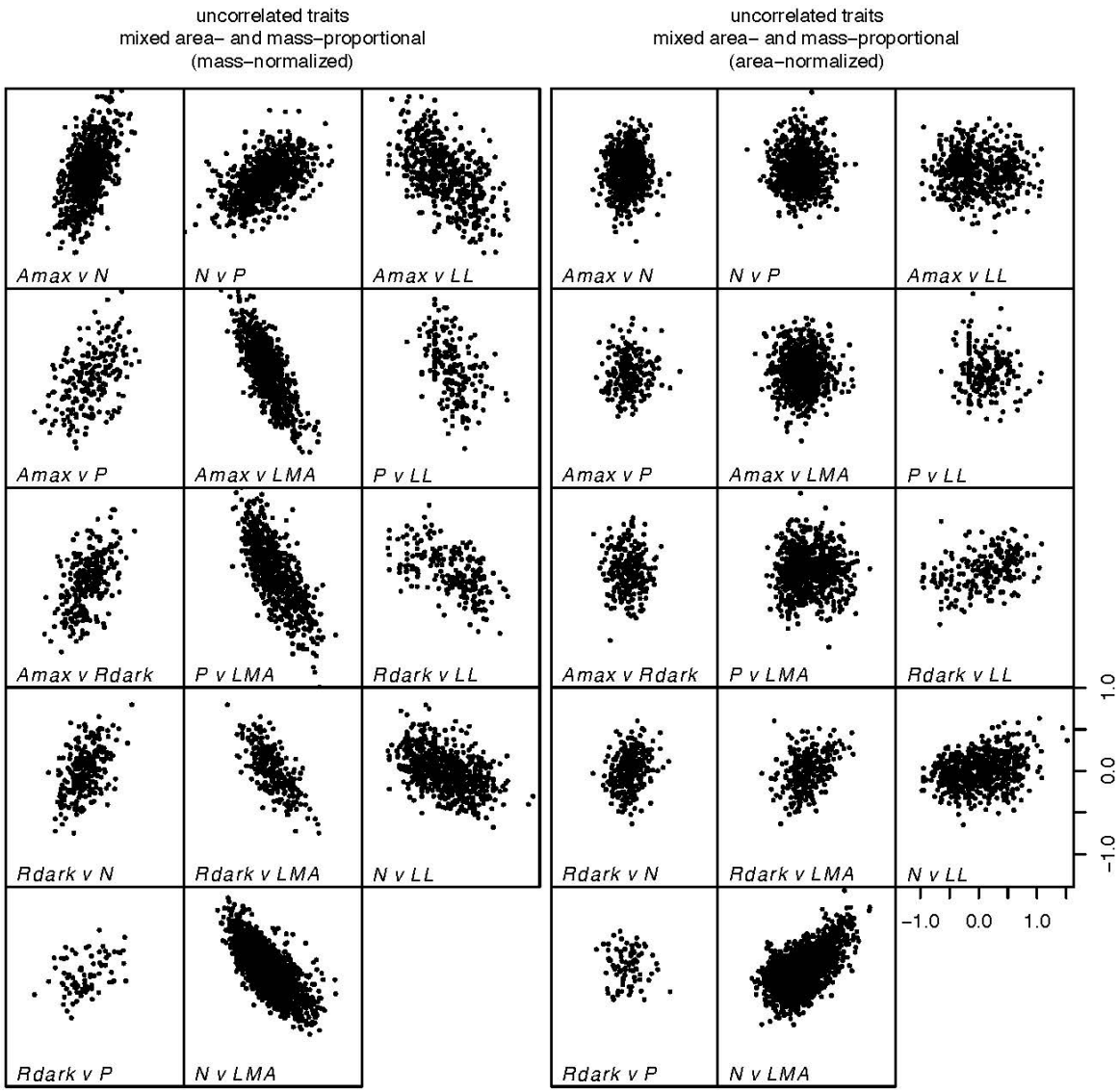

Fig. S2. Mass- (left three columns) and area-normalized (right three columns) predictions of null Model-C. The plots are all $\log_{10}$-$\log_{10}$, with means rescaled to (0,0).

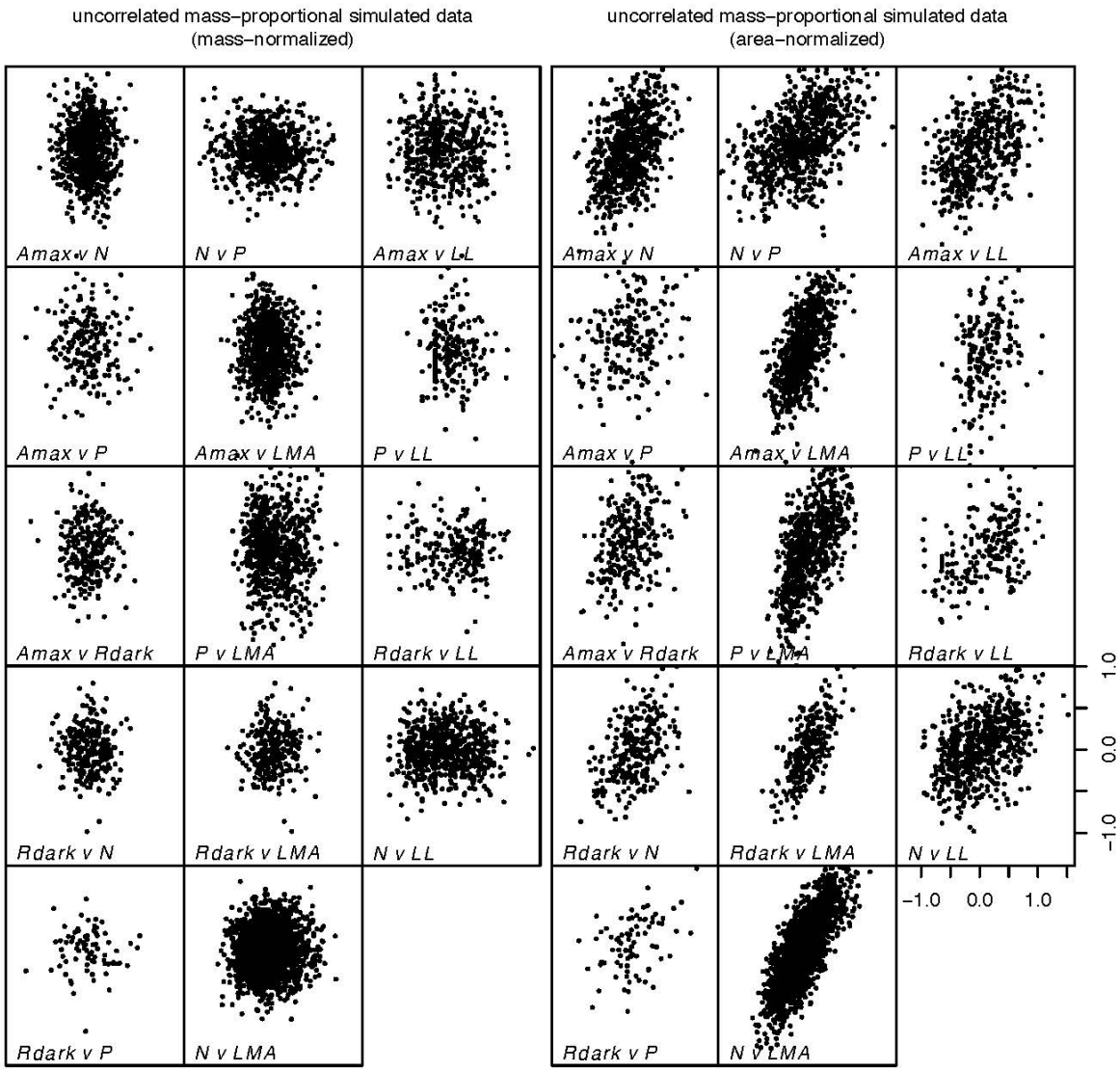

Fig. S3. Mass- (left three columns) and area-normalized (right three columns) predictions of null Model-M. The plots are all $\log_{10}$-$\log_{10}$, with means rescaled to (0,0).

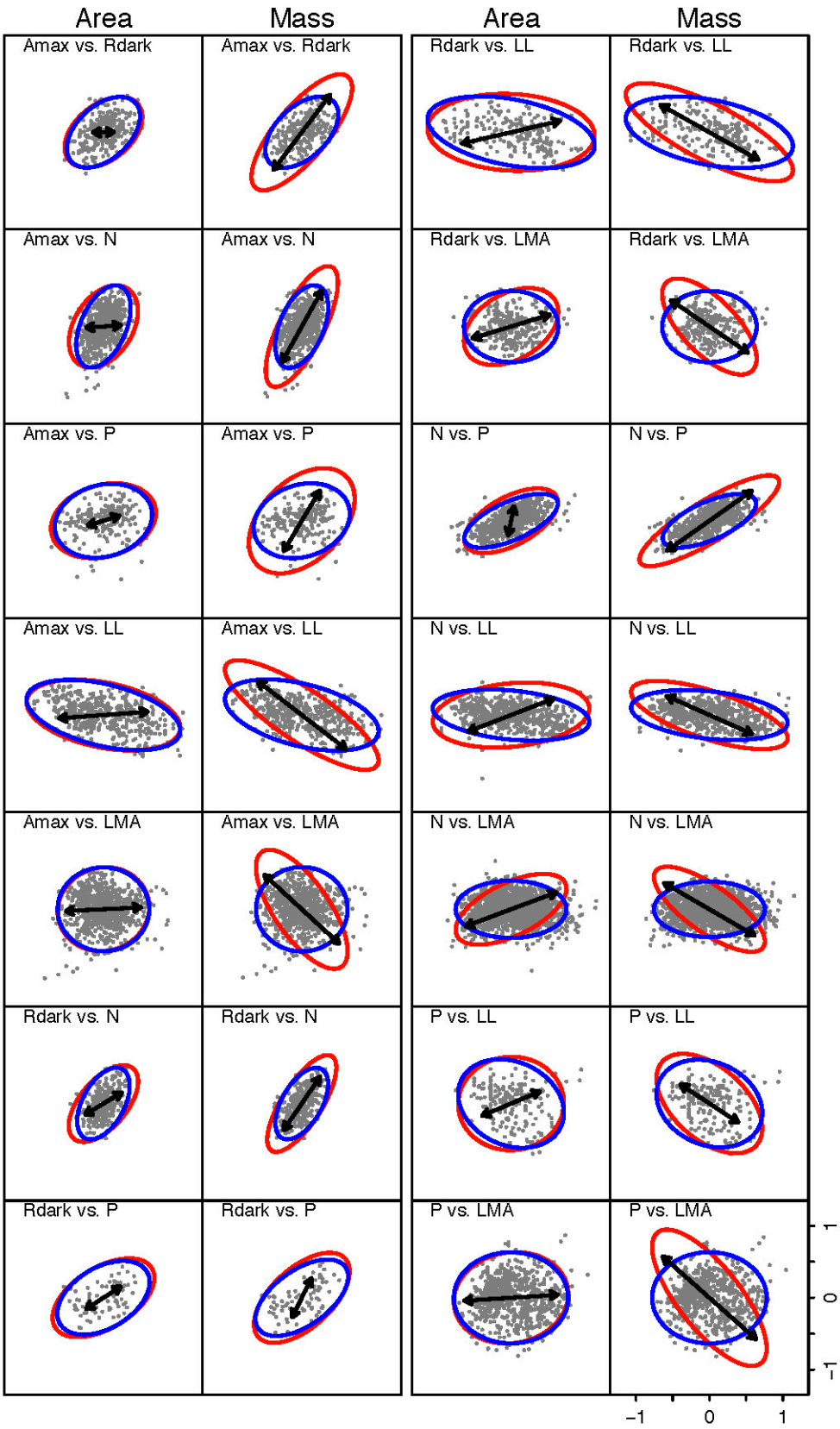
Area
Mass
Area
Mass
Amax vs. Rdark
Amax vs. Rdark
Rdark vs. LL
Rdark vs. LL
Amax vs. N
Amax vs. N
Rdark vs. LMA
Rdark vs. LMA
Amax vs. P
Amax vs. P
N vs. P
N vs. P
Amax vs. LL
Amax vs. LL
N vs. LL
N vs. LL
Amax vs. LMA
Amax vs. LMA
N vs. LMA
N vs. LMA
Rdark vs. N
Rdark vs. N
P vs. LL
P vs. LL
Rdark vs. P
Rdark vs. P
P vs. LMA
P vs. LMA
1
0
−1
−1
0
1

Fig. S4. Comparison of the normalization-independent LES (blue ellipses, which are the same whether data are area- or mass-normalized) with the area- and mass-normalized LESs (red ellipses in the"Area" and "Mass" columns, respectively). Scatter-plots show residuals from main-text Equations 1 and 2 for the logarithms of Amax, Rdark, N and P, and mean-centered values for the logarithms of LL and LMA. Area and Mass residuals differ slightly from each other because area- and mass-normalized traits are not always inter-converted exactly by species mean LMA in GLOPNET. Ellipses are 95% bivariate normal density contours derived from estimated covariance matrices. Red ellipses are estimated to encompass 95% of the mean-centered normalized trait data in $\log_{10}$-space (area-normalized data in left side of Fig. S1; mass-normalized data in left side of Fig. 1). Blue ellipses are derived from the estimated covariance matrices among residuals from main-text Equations 1 or 2, logLL, and logLMA. Differences between red and blue ellipses show effects of normalization, which are much stronger for mass- than for area-normalization. Effects of area-normalization are relatively weak because traits are primarily area-proportional. If a trait were purely area-proportional (as Amax and P nearly are), then the area-normalized and normalization-independent relationships would be identical to each other (as is nearly the case for relationships involving Amax, P, and the non-normalized quantities LMA and LL). See also Fig. S5.

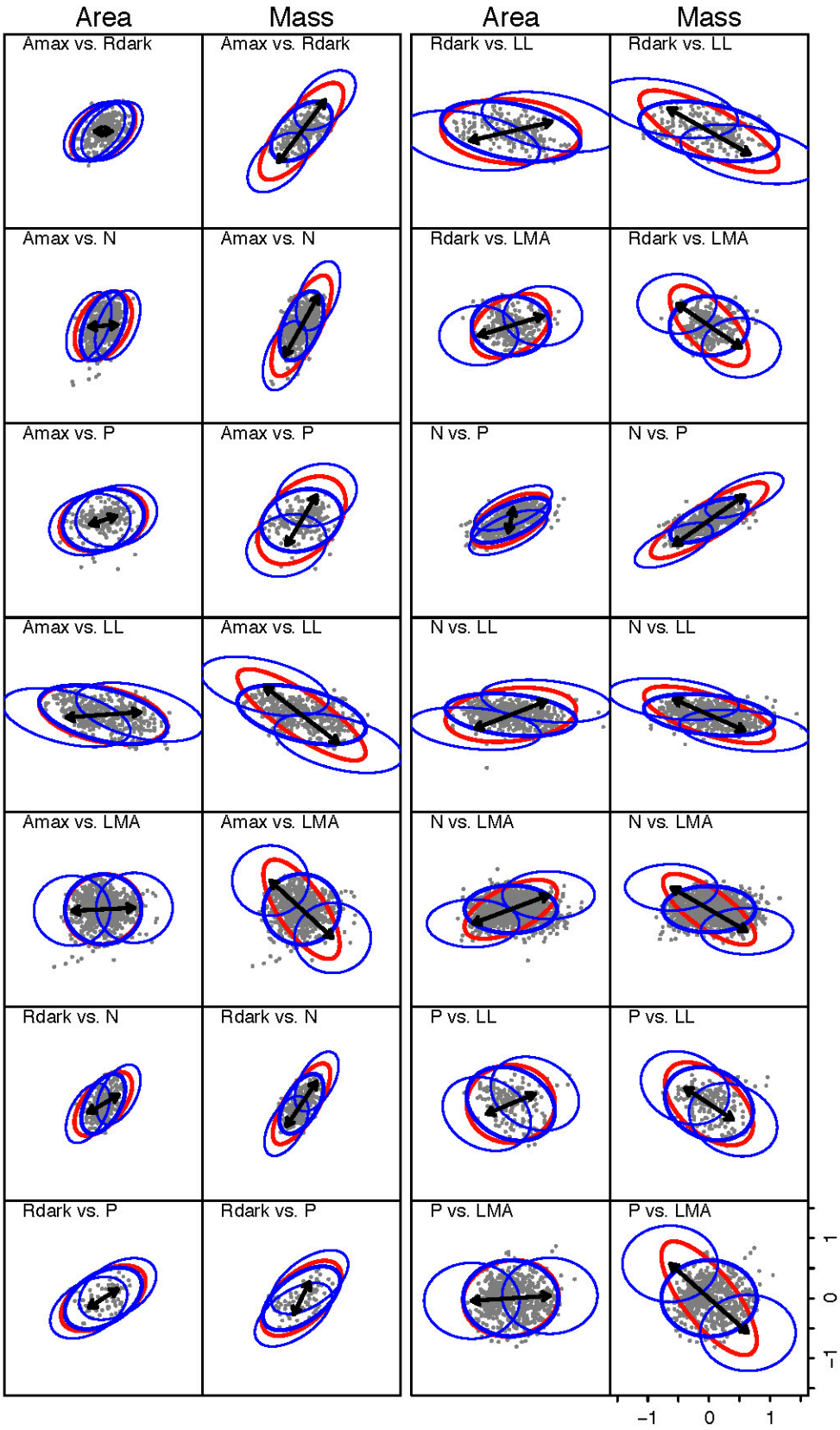
Area
Mass
Area
Mass
Amax vs. Rdark
Amax vs. Rdark
Rdark vs. LL
Rdark vs. LL
Amax vs. N
Amax vs. N
Rdark vs. LMA
Rdark vs. LMA
Amax vs. P
Amax vs. P
N vs. P
N vs. P
Amax vs. LL
Amax vs. LL
N vs. LL
N vs. LL
Amax vs. LMA
Amax vs. LMA
N vs. LMA
N vs. LMA
Rdark vs. N
Rdark vs. N
P vs. LL
P vs. LL
Rdark vs. P
Rdark vs. P
P vs. LMA
P vs. LMA
−1
0
1
−1
0
1

Fig. S5. Same as Fig. S4, but with additional blue ellipses representing the normalization-independent relationships at different LMA values. The bold blue ellipse in each panel is estimated to encompass 95% of the normalization-independent variation for a given pair of traits when evaluated at the mean logLMA (standardized to zero in the figure). The thin blue ellipses, which reflect the identical covariances indicated by the bold ellipses, show how the normalization-independent LES moves in bivariate trait space as LMA varies. The centroids of the thin blue ellipses (located at the arrow tips) indicate the centers of mass of the normalization-independent relationships for leaves with logLMA ± 2 standard deviations from the mean logLMA.

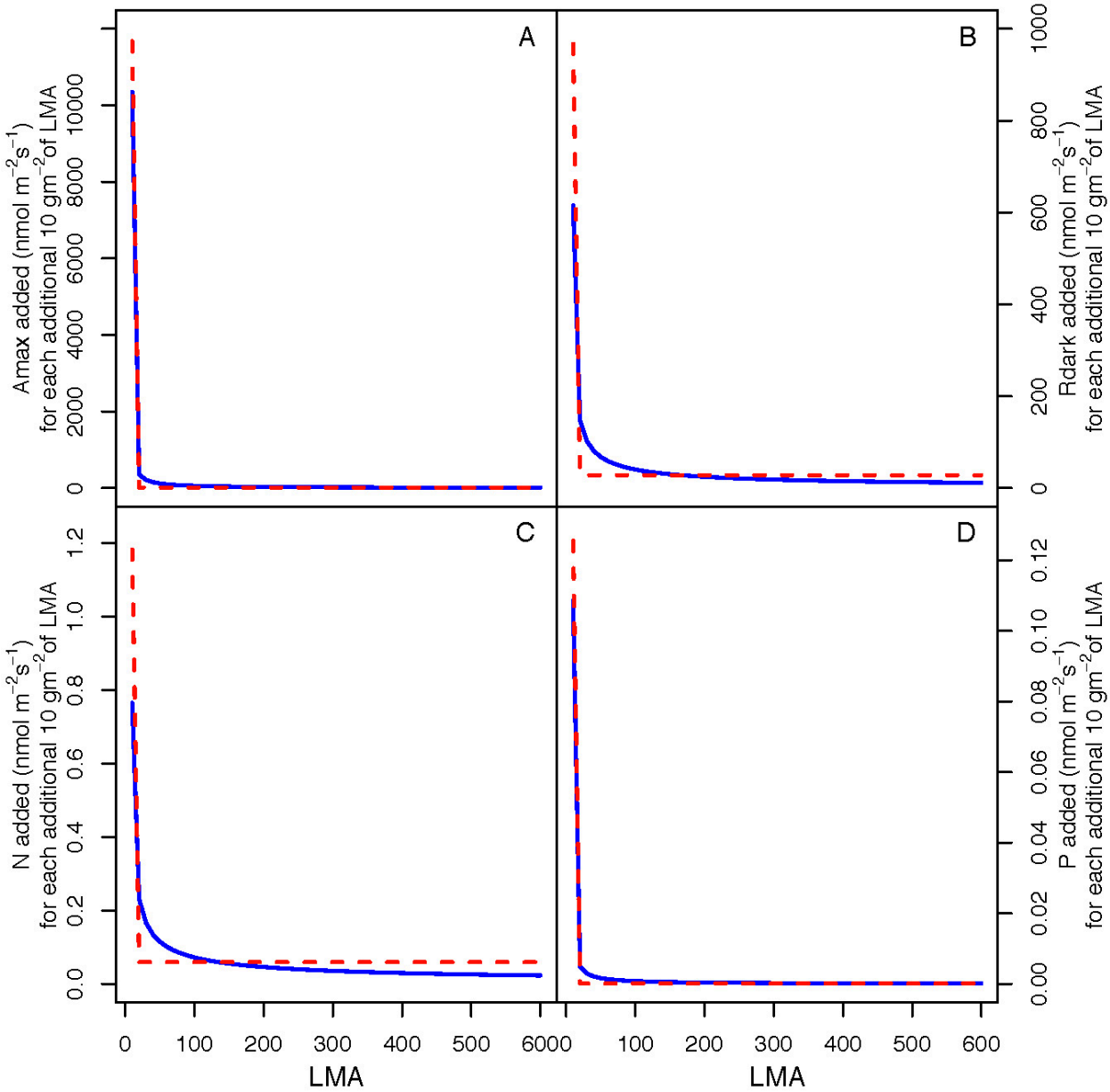

Fig. S6. LMA profiles (amount of area-normalized trait added to a leaf per 10 g m^{-2} LMA) predicted by Model-LN (blue curve) and Model-C (red curve) for: A Amax; B Rdark; C N; and D P. The LMA profile for Model-LN is obtained by calculating the expected amount of area-normalized trait (X_{Aik}) added to a leaf for each 10 g m^{-2} of LMA. The area-normalized form of Model-LN is equivalent to main-text Equation 1 transformed to the linear scale. The LMA profile for Model-C is obtained in analogous fashion, assuming that the area-associated trait amount is contained in the first 10 g m^{-2} of LMA. We make this assumption here to facilitate comparison between Models-LN and -C, but none of our other analyses make any assumptions about the physical location of traits within leaves.

Table S1. Results for all trait models (Model-M&A, -C -M, -A, and -LN) fit to area- and mass-normalized data, including estimates from Models -C and -M&A for p_A, the fraction of each trait associated with leaf area (the fraction $1 - p_A$ is associated with leaf mass). Statistics for p_A reported here for Model-C ignore interspecific variation (ε_A and ε_M), and therefore differ slightly from the results in Fig. 2. The column definitions are:

- trait = Amax, Rdark, N, or P
- normalization = mass- or area-normalized data
- model = M&A, C, M, A, or LN
- n = number of observations analyzed (number of GLOPNET records for which a trait value and LMA were both reported)
- npar = number of free parameters in model
- loglik = log(likelihood)
- AIC = 2*npar – 2*loglik
- pA.mu = average p_A from distribution that propagates parameter uncertainty and the distribution of LMA values from GLOPNET.
- pA_xxx = quantile of p_A species distribution. For example, p_A _0.01 is the 1st percentile, and p_A _0.5 is the 50th percentile. If p_A _0.01 = 0.7, this means that 1% of the species have a p_A value less than 0.7; or, equivalently, 99% of the species have a p_A value greater than 0.7.
- c1_mle = maximum likelihood estimate (MLE) for parameter c1, where c1 = $\bar{X}_{Mi}$ for trait *i* in Models-M&A, -C, and -M, and c1 = *K* in Model-LN. c1_mle is reported as "NA" for Model-A.
- c2_mle = MLE for parameter c2, where c2 = $\bar{X}_{Ai}$ for trait *i* in Models-M&A, -C, and -A, and c2 = *S* in Model-LN. c2_mle is reported as "NA" for Model-M.
- sdlog1_mle = MLE for parameter sdlog1, where sdlog1 is the standard deviation of log(ε) in Models-M&A and -LN, and the standard deviation of log(ε_M) in Models-C, and -M. sdlog1_mle is reported as "NA" for Model-A.
- sdlog2_mle = MLE for parameter sdlog2, where sdlog2 is the standard deviation of log(ε_A) in Models-C, and -A. sdlog2_mle is reported as "NA" for Models-M&A, -LN, and -M.
- Columns with suffixes _2.5 and _97.5 are the lower and upper 95% confidence limits fo*r* the named parameter. The value is "NA" if the parameter is not included in the model.

Table S1 is available in a separate file , "Osnas_etal_Table_S1.xlsx".

Table S2. Pairwise regression statistics for $\log_{10}$-transformed normalized data and normalization-independent quantities (residuals from main-text Equations 1 or 2) among six traits from the GLOPNET dataset: Amax, N, P, Rdark, LMA and LL. The Table reports slopes and intercepts (95% CIs in parentheses) for three different regressions for all possible pairwise trait combinations: Ordinary Least Squares (OLS), Standardized Major Axis (SMA), and Major Axis (MA). "*r*" is the Pearson correlation coefficient. "SMA p-value" is for the null hypothesis that the SMA slope equals one, and "OLS p-value" is for the null hypothesis that the OLS slope equals zero. The covariance and sample size for each analysis are also reported. Row labels are independent variables ("X trait"), and column headings are dependent variables ("Y trait"). The terms "independent" and "dependent" are used here only in the statistical sense. We make no assumptions here about the direction of causality, and performed tests in both directions for a given pair. Column headings indicate if the X and Y variables were normalized by mass or area, except for LMA and LL (which were not normalized). Normalization-independent analyses were performed with both mass- and area-normalized data by regressing logX and logY on logLMA, and then regressing the logY residuals on the logX residuals. As noted elsewhere, normalization-independent analyses are mathematically identical whether performed with mass- or area-normalized data, but the statistics differ slightly because mass- and area-normalized traits in GLOPNET are not inter-converted exactly via species mean LMA. Units (prior to $\log_{10}$ transformation) and the arithmetic mean and variance of the $\log_{10}$-transformed Y variables appear above each column. Statistics for the normalization-independent LES are reported here for the first time. Many of the other statistics were reported in (*2*) and are reproduced here for convenience.

Table S2 is available in a separate file , "Osnas_etal_Table_S2.xlsx".

Table S3. Principal Component Analysis (PCA) for seven sets of $\log_{10}$-transformed trait values (Tables S3a-S3g). Normalization-independent PCAs (Tables S3e-g) were based on normalization-independent residuals obtained from mass-normalized data and main-text Equation 2; very similar results were obtained from area-normalized data and main-text Equation 1. Rows labeled by traits report the loadings for the trait on each principal component axis ("PC"). Note that the direction of each PC is arbitrary, so that the interpretation of a given PC would be unchanged if the signs of all loadings were reversed. The bottom row of each Table ("eig.") reports the eigenvalues, which are the proportions of variance explained by each PC. The dataset used to construct each Table is described above each Table.

Table S3a. Mass-normalized GLOPNET data, including LMA.

Trait	**PCA Loadings (Unit Vector)**					
	PC 1	**PC 2**	**PC 3**	**PC 4**	**PC 5**	**PC 6**
LMA	-0.413	0.202	0.740	0.406	0.128	-0.246
LL	-0.405	-0.416	-0.112	-0.324	0.722	-0.156
N	0.434	-0.271	0.201	0.422	0.433	0.576
P	0.372	-0.669	0.186	0.101	-0.224	-0.565
Rdark	0.415	0.220	0.532	-0.691	0.134	0.011
Amax	0.407	0.466	-0.287	0.251	0.455	-0.514
eig.	**0.746**	**0.129**	**0.050**	**0.036**	**0.034**	**0.006**

Table S3b. Area-normalized GLOPNET data, including LMA.

Trait	**PCA Loadings (Unit Vector)**					
	PC 1	**PC 2**	**PC 3**	**PC 4**	**PC 5**	**PC 6**
LMA	-0.405	0.484	0.359	-0.125	-0.108	0.668
LL	-0.144	0.689	-0.057	0.223	0.515	-0.431
N	-0.557	0.035	-0.029	0.283	-0.667	-0.403
P	-0.414	-0.185	-0.736	0.240	0.250	0.364
Rdark	-0.496	-0.206	0.030	-0.761	0.243	-0.271
Amax	-0.296	-0.462	0.569	0.467	0.396	0.000
eig.	**0.436**	**0.288**	**0.133**	**0.074**	**0.049**	**0.020**

Table S3c. Mass-normalized GLOPNET data, excluding LMA.

Trait	PCA Loadings (Unit Vector)				
	PC 1	**PC 2**	**PC 3**	**PC 4**	**PC 5**
LL	0.452	-0.374	-0.081	0.783	-0.192
N	-0.475	-0.338	0.304	0.313	0.686
P	-0.394	-0.721	-0.002	-0.243	-0.516
Rdark	-0.462	0.161	-0.832	0.260	0.010
Amax	-0.448	0.448	0.456	0.404	-0.476
eig.	**0.752**	**0.150**	**0.047**	**0.040**	**0.010**

Table S3d. Area-normalized GLOPNET data, excluding LMA.

Trait	PCA Loadings (Unit Vector)				
	PC 1	**PC 2**	**PC 3**	**PC 4**	**PC 5**
LL	-0.057	0.798	0.433	0.011	0.414
N	0.540	0.288	0.230	0.212	-0.727
P	0.492	0.206	-0.661	0.393	0.352
Rdark	0.544	-0.021	-0.012	-0.825	0.151
Amax	0.410	-0.487	0.568	0.346	0.391
eig.	**0.473**	**0.259**	**0.124**	**0.086**	**0.057**

Table S3e. Normalization-independent residuals from main-text Equation 2 for N, P, Rdark, and Amax; and LMA-independent variation in LL obtained in the same manner (i.e., residuals from ordinary least squares regression of LL on LMA).

Trait	PCA Loadings (Unit Vector)				
	PC 1	**PC 2**	**PC 3**	**PC 4**	**PC 5**
LL	0.449	-0.436	0.319	-0.629	-0.333
N	-0.464	-0.352	0.607	-0.063	0.537
P	-0.412	-0.623	-0.158	0.347	-0.545
Rdark	-0.481	-0.019	-0.551	-0.671	0.117
Amax	-0.428	0.545	0.448	-0.171	-0.539
eig.	**0.584**	**0.190**	**0.096**	**0.077**	**0.052**

Table S3f: Normalization-independent residuals from main-text Equation 2 for N, P, Rdark, and Amax; and $\log_{10}$ values (not residuals) for LL and LMA.

Trait	**PCA Loadings (Unit Vector)**					
	PC 1	**PC 2**	**PC 3**	**PC 4**	**PC 5**	**PC 6**
LL	0.353	-0.600	0.105	0.124	-0.253	0.651
N	-0.491	-0.240	0.177	0.520	0.613	0.156
P	-0.409	-0.350	0.558	-0.031	-0.503	-0.380
Rdark	-0.508	-0.105	-0.112	-0.746	0.119	0.385
Amax	-0.434	0.020	-0.655	0.366	-0.487	0.101
LMA	0.140	-0.669	-0.452	-0.153	0.237	-0.499
eig.	**0.440**	**0.264**	**0.141**	**0.079**	**0.051**	**0.025**

Table S3g: Normalization-independent residuals from main-text Equation 2 for N, P, Rdark, and Amax; LMA-independent variation in LL (as in Table S3e); and $\log_{10}$ LMA values.

Trait	**PCA Loadings (Unit Vector)**					
	PC 1	**PC 2**	**PC 3**	**PC 4**	**PC 5**	**PC 6**
LL	0.449	0.000	-0.436	-0.319	0.629	-0.333
N	-0.464	0.000	-0.352	-0.607	0.063	0.537
P	-0.412	0.000	-0.623	0.158	-0.347	-0.545
Rdark	-0.481	0.000	-0.019	0.551	0.671	0.117
Amax	-0.428	0.000	0.545	-0.448	0.171	-0.539
LMA	0.000	1.000	0.000	0.000	0.000	0.000
eig.	**0.487**	**0.167**	**0.158**	**0.080**	**0.064**	**0.043**

Table S4. Pearson correlation coefficients ("*r*") for pairs of $\log_{10}$-transformed traits expressed either as mass-normalized values or as normalization-independent residuals from main-text Equation 2. Mass-normalized trait correlations are larger than normalization-independent correlations for all pairwise relationships between N, P, Amax and Rdark. These larger correlations reflect "inflation" induced by mass-normalization of traits that are in fact primarily area-proportional (Fig. 2). The inflation of the correlation coefficient due to mass-normalization ranges from 11.5% for P vs. Rdark to 96.5% for P vs. Amax. *n* is the sample size for each correlation.

X *vs* Y	Mass-normalized *r*	Normalization-independent *r*	Inflation (%)	*n*
N *vs* Amax	0.73	0.46	59.1	705
N *vs* P	0.85	0.64	34.1	733
N *vs* Rdark	0.74	0.53	40.3	267
P *vs* Amax	0.41	0.21	96.5	207
P *vs* Rdark	0.59	0.53	11.5	78
Amax *vs* Rdark	0.77	0.5	53.0	259

References and Notes

1. P. B. Reich, M. B. Walters, D. S. Ellsworth, From tropics to tundra: Global convergence in plant functioning. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 13730 (1997). doi:10.1073/pnas.94.25.13730 Medline

2. I. J. Wright *et al*., The worldwide leaf economics spectrum. *Nature* **428**, 821 (2004). doi:10.1038/nature02403 Medline

3. G. B. Bonan, S. Levis, L. Kergoat, K. W. Oleson, Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. *Global Biogeochem. Cycles* **16**, 5-1 (2002). doi:10.1029/2000GB001360

4. G. C. Hurtt *et al*., Projecting the future of the U.S. carbon sink. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1389 (2002). doi:10.1073/pnas.012249999 Medline

5. P. E. Thornton, N. E. Zimmermann, An improved canopy integration scheme for a land surface model with prognostic canopy structure. *J. Clim.* **20**, 3902 (2007). doi:10.1175/JCLI4222.1

6. S. Zaehle, A. D. Friend, Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates. *Global Biogeochem. Cycles* **24**, (Feb 11 2010). doi:10.1029/2009GB003521

7. Please see the supplementary materials on *Science* Online.

8. C. Field, H. Mooney, in *On the Economy of Plant Form and Function: Proceedings of the Sixth Maria Moors Cabot Symposium, Evolutionary Constraints on Primary Productivity, Adaptive Patterns of Energy Capture in Plants, Harvard Forest, August 1983*, T. J. Givnish, Ed. (Cambridge Univ. Press, Cambridge [Cambridgeshire], New York, 1986), pp. 25–55.

9. C. Field, H. A. Mooney, Leaf age and seasonal effects on light, water, and nitrogen use efficiency in a California shrub. *Oecologia* **56**, 348 (1983). doi:10.1007/BF00379711

10. D. S. Falster *et al*., Lifetime return on investment increases with leaf lifespan among 10 Australian woodland species. *New Phytol.* **193**, 409 (2012). doi:10.1111/j.1469-8137.2011.03940.x Medline

11. H. Poorter, Ü. Niinemets, L. Poorter, I. J. Wright, R. Villar, Causes and consequences of variation in leaf mass per area (LMA): A meta-analysis. *New Phytol.* **182**, 565 (2009). doi:10.1111/j.1469-8137.2009.02830.x Medline

12. I. J. Wright *et al*., Assessing the generality of global leaf trait relationships. *New Phytol.* **166**, 485 (2005). doi:10.1111/j.1469-8137.2005.01349.x Medline

13. J. R. Evans, H. Poorter, Photosynthetic acclimation of plants to growth irradiance: The relative importance of specific leaf area and nitrogen partitioning in maximizing carbon gain. *Plant Cell Environ.* **24**, 755 (2001). doi:10.1046/j.1365-3040.2001.00724.x

14. M. E. McGroddy, T. Daufresne, L. O. Hedin, Scaling of C:N:P stoichiometry in forests worldwide: Implications of terrestrial redfield-type ratios. *Ecology* **85**, 2390 (2004). doi:10.1890/03-0351

15. P. B. Reich *et al*., Evidence of a general 2/3-power law of scaling leaf nitrogen to phosphorus among major plant groups and biomes. *Proc. R. Soc. London Ser. B* **277**, 877 (2010). doi:10.1098/rspb.2009.1818 Medline

16. I. Aranda, F. Pardo, L. Gil, J. A. Pardos, Anatomical basis of the change in leaf mass per area and nitrogen investment with relative irradiance within the canopy of eight temperate tree species. *Acta Oecol.* **25**, 187 (2004). doi:10.1016/j.actao.2004.01.003

17. T. Koike, M. Kitao, Y. Maruyama, S. Mori, T. T. Lei, Leaf morphology and photosynthetic adjustments among deciduous broad-leaved trees within the vertical canopy profile. *Tree Physiol.* **21**, 951 (2001). doi:10.1093/treephys/21.12-13.951 Medline

18. U. Niinemets, J. D. Tenhunen, A model separating leaf structural and physiological effects on carbon gain along light gradients for the shade-tolerant species *Acer saccharum*. *Plant Cell Environ.* **20**, 845 (1997). doi:10.1046/j.1365-3040.1997.d01-133.x

19. J. Kattge *et al*., TRY - a global database of plant traits. *Glob. Change Biol.* **17**, 2905 (2011). doi:10.1111/j.1365-2486.2011.02451.x

20. R Core Development Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, version 2.8, 2010).

21. B. M. Bolker, bbmle: Maximum Likelihood Estimation (2007).

22. D. S. Falster, D. I. Warton, I. J. Wright, (S)MATR: Standardised Major Axis Tests and Routines (2003).