RESEARCH ARTICLE

Linking seed size and number to trait syndromes in trees


1Institut National de Recherche pour Agriculture, Alimentation et Environnement (INRAE), Laboratoire EcoSystemes et Societes En Montagne (LESSEM), Université Grenoble Alpes, Grenoble, France
2Forest Biology Center, Faculty of Biology, Adam Mickiewicz University, Poznan, Poland
3Laboratorio de Recursos Agroforestales, Centro Austral de Investigaciones Científicas (CADIC – CONICET), Ushuaia, Argentina
4INRAE, UR EFNO, Nogent-sur-Vernisson, France
5Department of Agriculture, Forest and Food Sciences, University of Torino, Grugliasco, Italy
6Université du Québec en Abitibi-Témiscamingue, Rouyn-Noranda, Québec, Canada
7Université du Québec à Montréal, Montreal, Québec, Canada
8CNRS, Université Paris-Saclay, AgroParisTech, Paris, France
9INRAE, UR 629, Ecologie des Forêts Méditerranéennes, Avignon, France
10Department of Biodiversity, Ecology and Evolution, Faculty of Biological Sciences, Complutense University of Madrid, Madrid, Spain
11Institut National de Recherche pour Agriculture, Alimentation et Environnement (INRAE), Biodiversity, Genes, and Communities (BIOGECO), Université Bordeaux, Pessac, France
Abstract

Aim: Our understanding of the mechanisms that maintain forest diversity under changing climate can benefit from knowledge about traits that are closely linked to fitness. We tested whether the link between traits and seed number and seed size is consistent with two hypotheses, termed the leaf economics spectrum and the plant size syndrome, or whether reproduction represents an independent dimension related to a seed size–seed number trade-off.

Location: Most of the data come from Europe, North and Central America and East Asia. A minority of the data come from South America, Africa and Australia.


Major taxa studied: Trees.

Methods: We gathered 12 million observations of the number of seeds produced in 784 tree species. We estimated the number of seeds produced by individual trees and scaled it up to the species level. Next, we used principal components analysis and generalized joint attribute modelling (GJAM) to map seed number and size on the tree traits spectrum.

Results: Incorporating seed size and number into trait analysis while controlling for environment and phylogeny with GJAM exposes relationships in trees that might otherwise remain hidden. Production of the large total biomass of seeds [product of seed number and seed size; hereafter, species seed productivity (SSP)] is associated with high leaf area, low foliar nitrogen, low specific leaf area (SLA) and dense wood. Production of high seed numbers is associated with small seeds produced by nutrient-demanding species with softwood, small leaves and high SLA. Trait covariation is consistent with opposing strategies: one fast-growing, early successional, with high dispersal, and the other slow-growing, stress-tolerant, that recruit in shaded conditions.

Main conclusions: Earth system models currently assume that reproductive allocation is indifferent among plant functional types. Easily measurable seed size is a strong predictor of the seed number and species seed productivity. The connection of SSP with the functional traits can form the first basis of improved fecundity prediction across global forests.

Keywords: fecundity, functional traits, leaf economics, life history strategies, size syndrome, tree recruitment
1 | INTRODUCTION

Our understanding of the mechanisms that promote and maintain forest tree diversity under a warming climate can benefit from knowledge about the traits that are closely linked to fitness (Adler et al., 2014; Kelly et al., 2021; Paine et al., 2015; Yang et al., 2018). Adaptive evolution operates on the variation that affects survival and reproduction. Leaf traits, wood density and plant height are clearly important for trees, yet their connections to fitness are indirect (Chave et al., 2009; Díaz et al., 2016; Violle et al., 2007; Wright et al., 2004). For example, large, thin, short-lived leaves with high nitrogen content confer clear benefits in high-resource environments where long-lived, highly lignified leaves are less advantageous (Field & Mooney, 1986; Reich, 2014; Shreve, 1925). Fitness is the quantitative representation of individual reproductive success, the ability of an organism to pass its genetic material to its offspring. Thus, interpreting the fitness implications of traits often requires broad extrapolation, such as ecophysiological measurements describing minute-scale responses of leaves, roots or xylem elements that are integrated with many other responses to determine survival and/or reproduction over the lifetimes of whole plants.

Although no trait links directly to fitness in trees, many are so weakly tied to fitness that their utility for comparative studies remains uncertain. In that light, seedling recruitment at tropical Barro Colorado Island (BCI) provides a more direct link to fitness (Rüger et al., 2018, 2020). Nonetheless, given that recruitment varies for each species at each site, the species-level reproductive effort could be a valuable extension for understanding traits. Only recently have long-term and geographically extensive measurements of the number of seeds produced by trees needed for species-level synthesis become available (Clark et al., 2021; Journé et al., 2022; Qiu et al., 2022; Sharma et al., 2022). Here, we reexamine the hypotheses that describe the seed number and size as part of an omnibus syndrome that explains all traits (e.g., fast–slow plant economics spectrum) or, alternatively, as a separate axis of variation.

Principal components analysis (PCA) has been a primary tool for exploring combined trait variation, contributing to at least three interpretations for forest trees. One view sees the number of seeds produced and their size together with leaf traits as part of a “fast–slow” continuum of plant strategies (Reich, 2014). That dimension represents the trade-off between resource acquisition and processing, and it could be linked to a growth–survival trade-off (Poorter et al., 2008; Rüger et al., 2018; Wright et al., 2010). Cheaply constructed leaves that assimilate carbon quickly, together with low wood density, characterize species that are resource-demanding, grow fast and die young (Moles, 2018; Westoby et al., 2002). In such species, the production of a large number of seeds might offset mortality losses (Muller-Landau, 2010; Reich, 2014). Species with some or all of these traits might dominate early successional stages through effective colonization, and they might not persist under intense competition (Poorter et al., 2008; Wright et al., 2010).

Alternatively, the size hypothesis suggests that the seed number and seed size are part of a stature–recruitment trade-off (Díaz et al., 2016; Kohyama, 1993; Rüger et al., 2018, 2020). According to this hypothesis, large plant size maximizes canopy performance at the expense of recruitment, and vice versa. The analysis of 282 co-occurring tree species at tropical BCI emphasized a leaf–trait axis and a size–recruitment axis, with species characterized by small stature, small leaves and small seeds having high recruitment (Rüger et al., 2018). Follow-up studies indicated that the stature–recruitment trade-off extends to tropical forests more generally (Kambach et al., 2022).

Finally, seed number and seed size might represent a third, largely independent dimension of variation, as proposed by the twin-filter (TF) hypothesis (Grime & Pierce, 2012). According to the TF hypothesis, primary strategies, such as fast–slow, determine persistence for the climate/habitat norms, whereas traits involved in episodic events, which might include reproduction, affect fitness independent of other traits (Grime & Pierce, 2012; Pierce et al., 2014). The leaf–height–seed (LHS) scheme of Westoby (1998) hypothesizes that seed size plays a role similar to reproduction in the TF model. In both hypotheses, plants can produce either many or few seeds (TF) or small or large seeds (LHS), largely independent of other plant traits. All three of the foregoing hypotheses imply an important role of seed number and seed size, and they assume that all traits have some connection to fitness. The availability of species-level seed numbers can provide new insight to trait analysis owing to its close connection to recruitment, a major demographic and fitness indicator.

A limitation of summaries available from PCA comes from the fact that correlations include all the indirect ways that traits could be associated. For example, a correlation between seed size and wood density could occur if there was a need for high wood density in order to produce large seeds. If true, this would be a direct relationship. Alternatively, both variables might be driven by climate for reasons that do not depend on one another. In such a case, that would be an indirect relationship. Another indirect relationship is represented by phylogenetic conservatism. Some species groups tend to produce larger seeds or denser wood than others, even if environments that might select for one or both traits change. The correlation structure exposed by PCA does not discriminate between direct (conditional) and indirect (marginal) relationships. If relationships are indirect, then conditional estimates offer the most transparent view of their connections (Seyednasrollah & Clark, 2020). To quantify direct links between traits, the traditional study with PCA is supplemented here with conditional relationships between traits using generalized joint attribute modelling (GJAM; Clark et al., 2017). Including the environment as fixed effects and phylogenetic groups as random effects, GJAM decomposes trait relationships into direct and indirect relationships. Although we believe that GJAM is a valuable extension, we also present results of PCA to facilitate comparison with past studies.

In this study, we analyse trait syndromes in trees from a perspective that includes the number of seeds produced and seed size. The masting inference and forecasting (MASTIF) network includes 12 million tree-year observations of the number of seeds produced by 775,991 trees from 784 species from a broad range of biomes (Journé et al., 2022; Qiu et al., 2022). To control for variation within species and, thus, to sharpen our understanding of interspecific differences, we estimate seed numbers produced by trees within an analytical framework that includes the condition of trees (species, size and shading), habitat (soils) and climate (temperature and
moisture deficit), while accommodating dependence between and within trees across years (Clark et al., 2019). This large sample size is important for the notoriously noisy seed production in trees (Kelly et al., 2021), whereby tree-to-tree and year-to-year variation in seed number spans several orders of magnitude (Clark et al., 2004; Journé et al., 2022). By combining seed number with seed size into species seed productivity [seed size x seed number (SSP), developed by Qiu et al. (2022)], we show how reproductive traits relate to one another separately and in combination. Combining seed number and seed size into SSP provides more exhaustive information on reproductive investment, because species that invest in large seeds are producing more seeds than expected from the 1:1 trade-off (Qiu et al., 2022). For this reason, SSP should be more strongly aligned with seed size than with seed number. By standardizing these metrics for the tree size, we account for the variation in size distribution within the data and facilitate comparisons. For example, the SSP is the average annual species seed productivity per square metre basal area in average environmental conditions across the species range in the data (Qiu et al., 2022; Qiu, Aravena, et al., 2021). If large seeds confer an advantage in competitive, shaded understoreys, whereas many small seeds allow colonization of distant sites, then SSP provides a direct link to fitness. The hypothesized relationships between seed number, seed size, SSP and traits are summarized in Figure 1.

2 MATERIALS AND METHODS

2.1 Seed number, SSP and MASTIF model

Estimation of the number of seeds produced in perennial plants suffers from an extreme signal-to-noise problem, created by orders of magnitude variation from year to year and tree to tree (Clark et al., 2004; Pearse et al., 2020; Pesendorfer et al., 2021) that can bury any trend (Clark et al., 2021). There are as many time series as there are trees that must be modelled together because there is dependence created by among-tree synchrony in masting variation (Bogdziewicz et al., 2021; Crone et al., 2011). Masting patterns are complicated further by the spatio-temporal variation in habitat and climate (Pearse et al., 2020; Pesendorfer et al., 2021). The many sources of variation mean that estimation of a seed number produced by trees can be achieved only from broad coverage and large sample sizes while accounting for the condition of individual trees, their local habitat and climate (Clark et al., 2021; Qiu, Aravena, et al., 2021; Sharma et al., 2022). This is achieved here with the MASTIF model (Clark et al., 2019).

The MASTIF model and data from the MASTIF network are summarized here and have been described extensively in recent papers (Clark et al., 2019, 2021; Journé et al., 2022; Qiu, Aravena, et al., 2021; Sharma et al., 2022). The tree-year observations of seed numbers in the network come from seed traps and from crop counts. Data include longitudinal (repeated) observations on individual trees (99%) and opportunistic observations that come through the iNaturalist project (Clark et al., 2019). Seed trap data consist of numbers of seeds that accumulate annually in mapped seed traps on forest inventory plots. A fitted dispersal kernel relates seed counts to mapped trees, accounting for uncertainty in seed transport and Poisson seed counts (Clark et al., 2019). Crop counts include counts of reproductive structures with estimates of the fraction of the crop observed, and beta-binominal distribution accounts for uncertainty in the crop-fraction estimates (Clark et al., 2019). This study includes 12,008,722 tree-years from North America, South and Central America, Europe, Africa, Asia and Oceania, gathered from 5115 sites and 787,444 trees (Figure 2). The MASTIF model jointly estimates the number of seeds produced based on all the observations. The

FIGURE 1 Hypothetical associations between dimensions of plant life strategies represented by functional traits and the seed number, seed size and species seed productivity (the product of seed number and seed size, SSP). Seed production can be associated with: (a) fast life syndrome (slow–fast resource turnover axis); (b) size syndrome; (c) its own, largely independent axis of seed size–seed number trade-off. Both seed number and SSP are divided by tree basal area in our analyses. PC, principal component; SLA, specific leaf area.
seed number (SN) and SSP (Qiu et al., 2022) used in the analysis are calculated based on 297,690 mature individuals and 3,730,381 tree-years. The MASTIF model uses the whole dataset to estimate seed numbers produced annually, but the SN and SSP are calculated based on a mature tree subset of these estimates.

The MASTIF model, detailed by Clark et al. (2019), is a dynamic biophysical model for year-to-year and tree-to-tree seed production. The MASTIF model is a Bayesian hierarchical, state-space model that allows for conditional independence in crop-count and seed-trap data through latent states. The model estimates the number of seeds produced with conditional fecundity, which depends on the probability that the tree is sexually mature, tree size, shading (five classes from full sun to full shade), local climate and soil conditions. Random effects on individual and year allow for wide variation between trees and over time that is typical of seed production. The posterior covariance between trees and years can take any form, avoiding assumptions of standard time-series models, which is important owing to the quasiperiodic variation in time and varying levels of synchronicity between individual trees (Pesendorfer et al., 2021). Model structure and methodology were implemented with R v.4.0 (R Core Team, 2020) and the R package Mast Inference and Forecasting (MASTIF) (Clark et al., 2019).

### 2.2 Seed number and SSP at the tree and species levels

The MASTIF model incorporates the effects of tree attributes with the environment on maturation and conditional fecundity. To allow for an uncertain identification of seeds from trees of the same genus and for dependence within trees over time and between trees, all tree-years of a genus are modelled jointly (Clark et al., 2019, 2021). For each tree $i$ of species $s$ at stand $j$ in year $t$, the expected seed number is the product of maturation probability $\hat{p}$ and conditional fecundity $\hat{\psi}$:

$$E(f_{ijs,t}) = \hat{f}_{ijs,t} = \hat{p}_{ijs} \hat{\psi}_{ijs,t}. \quad (1)$$

Conditional fecundity depends on predictors, individual effects, year effects and error, as follows:

$$\log_{10}(\hat{\psi}_{ijs,t}) = x_{ijs,t}^T \beta_{ijs} + \mu_{ijs} + \gamma_{ijs,t} + \eta_{ijs,t}. \quad (2)$$

where $x_{ijs}$ is a matrix containing individual attributes and environmental conditions (see section 2.5), and $\beta_{ijs}$ are fixed-effects coefficients; $\mu_{ijs}$ is the random effect for tree $i$ of species $s$ at stand $j$; $\gamma_{ijs,t}$ are year effects that are random across groups $g$ and fixed for the year $t$ to account for interannual variation that is not captured fully by climate anomalies. Group membership for year effects ($\gamma_{ijs,t}$) is defined by species-ecoRegion (Clark et al., 2019). There is a noise term, $\eta_{ijs,t}$.

Maturation probability $\hat{p}_{ijs}$ accounts for the immature state (for small trees) and failed crop in larger trees. The model implementation is open access with the R package MASTIF, with algorithm details provided by Clark et al. (2019).

We estimated the investment of species into reproduction using two metrics, both scaled to the tree basal area: annual seed number (SN) and annual species seed productivity (SSP; seed number $\times$ seed mass; Qiu et al. (2022)). Estimation of both SN and SSP starts with the estimation of individual tree number of produced seeds, which depends on the location of each tree, accounts for effects of the environment and includes uncertainty for each year. The number of seeds of individual trees produced over a species is then summarized as SN or SSP, as explained below. The tree-level estimate of seed number [i.e., individual seed production (ISP)], is the product of seed size (its mass, $m_i$) and seed number, scaled to tree basal area per year (Journé et al., 2022). We quantify ISP as the mass of a tree's seed production relative to its basal area to standardize for tree size [intermediate trees produce more seeds than smaller ones;
2.3 | Traits

We selected six functional plant traits previously shown to capture plant life strategies well (Carmona et al., 2021; Diaz et al., 2016): plant height (in metres); leaf area (in millimetres); specific leaf area (SLA; in square millimetres per milligram; the inverse of leaf mass per area); leaf nitrogen concentration (in milligrams per gram); wood density (in grams per cubic metre); and seed size (in grams). The data were obtained from primary sources and supplemented with publicly available data from the latest version of the TRY plant trait database (Kattge et al., 2020) extracted from the study by Carmona et al. (2021). Missing values for the six traits were filled with genus-level means.

2.4 | Trait relationships

PCA summarizes correlation in the joint distribution of traits, written in bracket notation as \([T] = [T_1, ... , T_g]\). If the relationship between traits depends on phylogeny, summarized by phylogenetic groups \(P_{s=1,...,g}\) (taxonomic, e.g., genus or family), and on the environment, \(X\), then there is a joint distribution \([T, P, X]\). The indirect effects of environment and phylogeny might dominate the relationships between some or many traits. An alternative approach uses a conditional distribution,

\[
[T | P, X] = \frac{[T, P, X]}{[P, X]}, \tag{3}
\]

where the distribution of groups and environments \([P, X]\) is that which occurs in the dataset.

To determine trait relationships, we fitted a joint model to the conditional distribution \([T|P, X]\), which provides estimates of the phylogeny as random groups, \(g\), \(g = 1, ..., M\), for species \(s\) and \(X\) as a \(Q \times M\) matrix of coefficients \(B\) for \(Q\) predictors of \(M\) traits. We then decompose the distribution into (conditional) effects of other traits and the environment (Qiu, Sharma, et al., 2021; Seyednasrollah & Clark, 2020). The effect of trait \(m\) on the remaining \(-m\) traits is the conditional distribution, \([T_{-m}|T_m, P, X]\). Using the fitted model in GJAM (see below, 2.5) we decompose the conditional effect of \(m\) on other traits as follows:

\[
E(T_{-m} | T_m, P, X) = E(T_{-m} | T_m) + E(T_{-m} | P, X). \tag{4}
\]

The first term is a conditional influence of \(m\) as distinct from \((P, X)\).

2.5 | Generalized joint attribute modelling

To incorporate the effects of environment and phylogeny on the distribution of traits, we use GJAM (Clark et al., 2017). Environmental covariates include soil fertility (cation exchange capacity), mean annual temperature and annual accumulative moisture deficit (difference between potential evapotranspiration and precipitation), averaged at the species level for the MASTIF dataset. GJAM allowed us to accommodate the dependence between traits and phylogeny as random groups. A more detailed description of GJAM fitting is given in the Supporting Information. GJAM fitting is open access, with the R package gjam on CRAN https://cran.r-project.org/web/packages/gjam/index.html.

3 | RESULTS

Across the 784 species, foliar traits, wood density and seed number and size are the dominant sources of variation. In the PCA of our data that includes SSP, 54.2% of the variation is concentrated in two principal components (PCs) of equal importance (Figure 3a; Supporting Information Figure S1). PC1 is associated with leaf traits. At one end are species with thin, large, acquisitive leaves (large SLA, high area and high foliar nitrogen). Common examples include heaven lotus (Gustavia superba), Panama tree (Sterculia apetala), pawpaw (Asimina triloba) and eastern walnut (Juglans nigra). At the other end are species with low SLA, low foliar nitrogen and low leaf area, including evergreen conifers such as giant sequoia (Sequoiadendron giganteum), California redwood (Sequoia sempervirens), monkey puzzle tree (Araucaria araucana), Fraser fir (Abies fraseri) and white cedar (Thuja occidentalis). PC2 is dominated by seed size, SSP and wood density. Large seeds are associated with high SSP because species that produce large seeds tend to produce proportionally more of them than predicted by the strict trade-off between seed size and number (Qiu et al., 2022). Dense wood is associated with both seed size and SSP, with examples including African crabwood (Carapa procera) and Fagales, such as chestnuts (Castanea) and oaks (Quercus). At the opposite end, with low-density wood and small seeds, are willows (Salix), fuchsia (Fuchsia excorticata) and trumpet tree (Cecropia obtusa). Tree height is weakly associated with foliar attributes: small trees tend to have large, thin leaves.

A second PCA, in which SSP is replaced with seed number, yields similar results (Supporting Information Figure S3). As with the PCA using SSP (Figure 3a), the first axis of this second PCA is associated with foliar traits, with no contributions from seed numbers. The second PCA axis separates species according to seed size, seed number and wood density. Tree height is again weakly associated with foliar attributes but also with reproduction: small trees tend to produce small seeds in large numbers (Supporting Information Figure S2).
Using conditional prediction to control for the environment and taxonomic relatedness shows that seed size is positively related to SSP and negatively related to seed number (Figure 3b). Conditional prediction allows for uncertainty and the effects of the environment on all traits, but then isolates their direct (conditional) relationships to one another (see Materials and Methods). Nutrient-demanding species with high foliar nitrogen concentrations, high SLA and low-density wood produce small seeds in high numbers, a relationship that is not apparent in PCA. Large seeds are produced by trees with dense wood, few seeds, high leaf area, low foliar nitrogen and low SLA (Figure 4b–g). The relationship between high SSP and dense wood is suggested by PCA (Figure 3a), but that relationship is not significant after accounting for environment and phylogeny (Figure 3b). Rather, the PCA can be driven by indirect links between traits. Although the links between wood density and foliar traits are significant, they are weaker than the relationship between wood density and seed size (Figure 4d).

4 | DISCUSSION

Across 784 species spanning tropical to boreal environments, estimation of the number of seeds produced by trees brings new insight to trait analysis with a strong connection to fitness. Seed size and number make a dominant contribution to trait syndromes in trees, but one that is not strictly consistent with the fast–slow or stature trade-offs. Controlling for common ancestry and environment with GJAM indicates that large seed size is weakly associated with high leaf area, low foliar nitrogen, low SLA and dense wood. These associations were not detected by conventional PCA that does not condition on environmental dependencies. Thus, there is a weak, indirect link between these traits and SSP. Fast strategies, as captured by leaf traits, were not coupled with high seed numbers, although nutrient-demanding trees show a tendency to produce small seeds. Seed size and number were also not associated with tree height as in the stature–recruitment hypothesis at the tropical BCI (Rüger et al., 2018, 2020). Across all species and sites in this study, trees with dense wood, large leaves and low nutrient demands produce large but few seeds. These species invest heavily in SSP. In contrast, a high seed number is associated with small seeds, most common in species with low-density wood, low leaf area, high foliar nitrogen and high SLA.

Trait relationships identified here are consistent with some traditional trait concepts, including change of traits associated with species turnover through succession (Bazzaz, 1979; Falster & Westoby, 2005; Wilfahrt et al., 2014). Production of a large number of small seeds increases recruitment in distant, disturbed habitats (Muller-Landau, 2010). The $r$ strategy of the $r$–$K$ spectrum is associated with fast growth and high nutrient requirements (Bazzaz, 1979; Henery & Westoby, 2001; Huston & Smith, 1987; Muller-Landau, 2010).

FIGURE 3  Seed size, seed number and species seed productivity (SSP) on the spectrum of tree form. (a) Biplot. Arrow length indicates the loading of each considered trait onto the first two principal components analysis (PCA) axes. Points represent the position of species, coded blue for needle, black for broad deciduous, and yellow for broad evergreen leaf habit. Larger points indicate means for the groups. An extended version of the graph is given in the Supporting Information (Figure S2). (b) Summary of generalized joint attribute modelling (GJAM) coefficients presented in Figure 4. Significant associations between traits are highlighted by lines, coded red for negative and blue for positive relationships. Dashed lines indicate associations that are significant only in the model without phylogenetic control (see Supporting Information Figure S4). Extended PCA plots are available in the Supporting Information (Figure S3). The SSP is the product of seed size and seed number (Qiu et al., 2022). Both SSP and seed number are standardized to a tree basal area. Thicker lines separate main relationships qualitatively from the minute correlations among some foliar and other traits. Each trait has a unique colour to improve comparisons between (a) and (b). N, nitrogen; PC, principal component; SLA, specific leaf area; SSP, species seed productivity.
Species seed productivity is more strongly driven by seed size than by seed number, our analysis indicates that the traditional $r$ strategy, which might include low-density wood that often comes with fast growth (Chave et al., 2009), also includes the production of small seeds. High foliar nitrogen and cheap leaf construction (high SLA) align with high photosynthetic rates (Moles, 2018; Reich, 2014; Reich & Oleksyn, 2004). On the K side are species with dense wood and slow growth (Poorter et al., 2005; Westoby, 1998). Low foliar nitrogen and low SLA can align with low foliar Rubisco content, low photosynthetic capacity and, thus, low-maintenance respiration in low light (Moles, 2018; Poorter, 2015; Reich et al., 1998). Species with such conservative leaves are also selected for large seeds needed for seedling establishment in shade, at the expense of the many small seeds that would promote colonization of distant sites (Muller-Landau, 2010; Westoby et al., 2002).

Species seed productivity is more strongly driven by seed size than by seed number, which follows from the observation that the size–number trade-off in trees is not 1:1 (Qiu et al., 2022). Instead, species that produce large seeds more than compensate (on a mass basis) for fewer of them, resulting in higher species seed productivity (Qiu et al., 2022). Therefore, the estimates of SSP for a given tree size enrich the interpretation of plant reproductive strategies beyond the insights that come from seed size alone (Lebrija-Trejos et al., 2016; Muller-Landau, 2010; Westoby et al., 2002). On the other hand, the production of small, copious seeds increases recruitment opportunities at the cost of limited investment per individual seed. Small seeds can mean low abiotic stress tolerance in competitive sites (Fricke et al., 2019; Tilman, 1994; Westoby et al., 2002). On the other hand, large seeds come with a cost of producing fewer of them (Fricke et al., 2019; Henery & Westoby, 2001; Muller-Landau et al., 2008), each with a greater investment in seedling survival (Fricke et al., 2019; Muller-Landau et al., 2008). However, the high SSP in species that produce large seeds means that such species are selected for proportionally high total seed biomass investment per individual to maintain populations in low light conditions (Falster & Westoby, 2005; Kohyama et al., 2003). In other words, the production of a large number of seeds appears to generate a generally higher cost of reproduction. Testing whether SSP is a better indicator of reproductive success than seed number or size alone appears a fruitful avenue for future research.

The divergent results from the present study and those suggesting a stature–recruitment trade-off in tropical forests (Guillemot et al., 2022; Kambach et al., 2022; Rüger et al., 2018, 2020) are not
necessarily in conflict. The within-site covariation in traits, whereby short trees might be associated with small seeds and leaves in the shaded understory (Rüger et al., 2018), does not need to align with an among-site, species-level difference, which integrates over environments for each species at many sites. Moreover, Rüger et al. (2018) measured the recruitment of saplings, whereas our analysis included seed numbers. In turn, the lack of relationship between seed size and plant height reported by past studies (Díaz et al., 2016) might follow from the fact that the GJAM models control for phylogeny, whereas PCA does not. This is supported by the fact that both PCA and GJAM models that do not include shared ancestry indicate a positive relationship between seed size and tree height. This and other trait relationships that are present only in phylogenetically controlled GJAM indicate that conditional prediction to control for the environment and taxonomic relatedness might be a step forward for the subdiscipline.

Anticipating individual and combined effects of global change requires an understanding of the vulnerability not only of mature trees but also of seed number and recruitment (Bogdziewicz, 2022; Clark et al., 2021; Hanbury-Brown et al., 2022; Ohse et al., 2023; Qiu, Sharma, et al., 2021; Sharma et al., 2022). One major challenge that exists in ecology more generally (Nuñez et al., 2021) is to increase the data coverage to underrepresented regions, such as Africa or Southeast Asia in our case. Earth system models currently assume that reproductive allocation does not differ among plant functional types (Hanbury-Brown et al., 2022; Scholze et al., 2006). There is area and promise for improvement using functional trait data. A recent study at BCI predicted forest succession by replacing the hyperdiversity of tropical forests with only two trait axes associated with fast−slow and size dimensions (Rüger et al., 2020). Although the size of our seed production dataset is unprecedented, seed number is much more difficult to measure owing to its variable nature, in comparison, for example, to seed size. This could explain why we found stronger links between traits with seed size. The good news is that the easily measurable seed size is a strong predictor of SSP and seed number. The connection of SSP with the functional traits can form the first basis of improved fecundity prediction across global forests.

AUTHOR CONTRIBUTIONS

M.B. and J.S.C. performed analyses and co-wrote the paper. M.B. and J.S.C. designed the study. J.S.C. compiled the MASTIF network and wrote the MASTIF model and software. B.C., G.K., V.J. and T.Q. co-wrote the paper. All authors contributed data and revised the paper.

ACKNOWLEDGEMENTS

We thank the National Ecological Observatory Network (NEON) for access to sites and vegetation structure data. F. LeFèvre (INRAE, France), William Schlesinger (the Comstock Foundation) and W. Koenig provided additional data. The project has been funded by grants to J.S.C. from the National Science Foundation, most recently DEB-1754443, and by the Belmont Forum (1854976), NASA (AIST16-0052 and AIST18-0063) and the Programme d’Investissement d’Avenir under project FORBIC (18-MPGA-0004; Make Our Planet Great Again). Jerry Franklin’s data remain accessible through NSF LTER DEB-1440409. Data from Hubbard Brook (New Hampshire) were funded through NSF-LTER. Puerto Rico data were funded by NSF grants, most recently DEB 0963447 and LTREB 1122325. Data from the Andes Biodiversity and Ecosystem Research Group were funded by the Gordon and Betty Moore Foundation and NSF 727 LTREB 1754647. M.B. was supported by grant no. 2019/35/D/NZ8/00050 from the (Polish) National Science Centre and by Polish National Agency for Academic Exchange Bekker programme PPN/BEK/2020/1/00009/U/00001. Research by the USDA Forest Service and the USGS was funded by these agencies. Any use of trade, firm or product names does not imply endorsement by the US Government. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data supporting the results are attached in the Online Supplement.

ORCID

Michal Bogdziewicz https://orcid.org/0000-0002-6777-9034
Davide Ascoli https://orcid.org/0000-0002-2671-5122
Chia-Hao Chang-Yang https://orcid.org/0000-0003-3635-4946
Arthur Guignabert https://orcid.org/0000-0002-1512-6760
Valentin Journé https://orcid.org/0000-0001-7324-7002
Georges Kunstler https://orcid.org/0000-0002-2544-1940
James A. Lutz https://orcid.org/0000-0002-2560-0710
Anders Marell https://orcid.org/0000-0002-3328-4834
Jonathan A. Myers https://orcid.org/0000-0002-2058-8468
Shoji Naoe https://orcid.org/0000-0002-0605-4187
John Poulsen https://orcid.org/0000-0002-1532-9808
Tong Qiu https://orcid.org/0000-0003-4499-437X
C. Lane Scher https://orcid.org/0000-0003-3689-5769
Kai Zhu https://orcid.org/0000-0003-1587-3317

REFERENCES


Additional supporting information can be found online in the Supporting Information section at the end of this article.