Global patterns in seed size
Angela T. Moles1,2,*, David D. Ackerly3, John C. Tweddle4,5, John B. Dickie4, Roger Smith1, Michelle R. Leishman2, Margaret M. Mayfield6, Andy Pitman7, Jeff T. Wood8 and Mark Westoby2

ABSTRACT

Aim To provide the first global quantification of the slope and shape of the latitudinal gradient in seed mass, and to determine whether global patterns in seed mass are best explained by growth form, vegetation type, seed dispersal syndrome, or net primary productivity (NPP).

Location Global.

Methods We collected seed mass data for 11,481 species × site combinations from around the world. We used regression to describe the latitudinal gradient in seed mass, then applied general linear models to quantify the relative explanatory power of each of the variables hypothesized to underlie the latitudinal gradient in seed size.

Results There is a 320-fold decline in geometric mean seed mass between the equator and 60°. This decline is not linear. At the edge of the tropics, there is a sudden 7-fold drop in mean seed mass. The strongest correlates of the latitudinal gradient in seed mass are plant growth form, and vegetation type, followed by dispersal syndrome and NPP. A model including growth form, vegetation type, dispersal syndrome and NPP explains 51% of the variation in seed mass. Latitude explains just 0.2% of the residual variation from this model.

Main conclusions This is the first demonstration of a major decrease in seed size at the edge of the tropics. This drop in seed mass is most closely correlated with changes in plant growth form and vegetation type. This suggests that the drop in seed mass might be part of a sudden change in plant strategy at the edge of the tropics.

Keywords Growth form, latitudinal gradient, plant traits, seed dispersal syndrome, seed mass.

INTRODUCTION

In the last few years both ecologists and the general public have become much more aware of global-scale environmental issues. However, our understanding of global patterns in the ecological traits that determine the way our ecosystems function is embarrassingly weak. Ecologists have not even quantified latitudinal gradients in readily observed, ecologically crucial plant traits such as leaf size and plant height, let alone reached an understanding of the factors that underlie these major biogeographical gradients.

This situation is about to change. Databases for some ecologically important plant traits now cover thousands or tens of thousands of species from around the world. This means that it is now possible to quantify patterns and test hypotheses regarding plant trait variation at a global scale. The present manuscript does just this, with seed mass.

Seed mass is a key ecological trait that influences many aspects of a species’ regeneration strategy, including seedling survival rates, seed dispersal syndrome, and the number of seeds that can be produced for a given amount of energy (Leishman et al., 2000). Seed size exhibits a striking global pattern, declining by 2–3 orders of magnitude between the equator and 60° (Moles & Westoby, 2003).

Several theories have been proposed to explain this latitudinal gradient in seed size.

1 Tropical species might need large seeds in order to cope with the shaded understorey conditions associated with tropical rain forests (Salisbury, 1942). This theory is supported by strong experimental data showing that large-seeded species have higher
survival under low-light conditions than do small-seeded species (Grime & Jeffrey, 1965; Leishman & Westoby, 1994; Osunkoya et al., 1994; Walters & Reich, 2000), by strong relationships between seed size and the light environment under which species establish (Salisbury, 1942; Foster & Janson, 1985; Mazer, 1989), and by the high leaf area index of much tropical vegetation (Asner et al., 2003).

2 The abundance of vertebrate seed dispersal agents in tropical ecosystems might have allowed tropical species to radiate into larger seed masses than was possible for species growing in the temperate zone, where the predominantly abiotic dispersal agents might favour small seeds (Lord et al., 1997). Empirical data confirm that species with wind and unassisted dispersal have smaller seeds than species with animal dispersal (Leishman et al., 1995) and that there is a latitudinal gradient in the proportion of vertebrate-dispersed fruits (Willson et al., 1989; Lord et al., 1997).

3 The relationship between seed size and latitude could result secondarily from a latitudinal gradient in plant size and life form, with a shift towards smaller plants at high latitudes together with a positive relationship between seed size and plant size (Baker, 1972; Moles et al., 2004).

4 A latitudinal gradient in the length of the growing season could place an upper limit on the size of seeds that can be produced at a given latitude (Stebbins, 1974). This idea is supported by data showing that larger seeds take longer to complete development than do smaller seeds (Eriksson & Ehrlen, 1991; Moles & Westoby, 2003).

5 The higher net primary productivity (NPP) of tropical ecosystems (Bondeau et al., 1999; Schloss et al., 1999) might translate to higher total seed production, thus increasing competition for space and favouring larger seeds (Moles & Westoby, 2003).

Previous work has shown latitudinal gradients in seed mass within vegetation types (Lord et al., 1997), within growth forms (Levin, 1974; Lord et al., 1997) and even within growth form–dispersal syndrome combinations (Lord et al., 1997). Thus, these variables are not the sole determinants of the latitudinal gradient in seed size. However, none of the five theories listed above are mutually exclusive. It is not known whether the latitudinal gradient in seed mass is explained by a combination of some or all of these mechanisms, or whether there is an additional component of latitudinal variation that remains unexplained.

Here, we describe the shape and slope of the relationship between seed mass and latitude across an unprecedented number of species and ecosystems from around the world. The large size of the data set allows us to compare the proportion of variation in seed mass explained by plant growth form, seed dispersal syndrome, vegetation type and NPP, in a way that was not possible with smaller data sets.

MATERIALS AND METHODS

We collated seed mass data for 11,481 species × site combinations. Approximately half of the seed mass data are from the Kew Gardens Seed Information Database (http://www.rbgkew.org.uk/data/sid/). The other half were assembled by A. Moles. In both cases, data were collected opportunistically from the published literature, and through personal communications (see Acknowledgements for details of contributors). Seed masses reported in the literature were assumed to be dry masses unless otherwise stated. Fresh masses were converted to approximate dry masses [dry mass = (0.92 × fresh mass)0.66, \textit{R}^2 = 0.97, n = 418 species]. It was not possible to distinguish reliably between seeds and diaspores because of the great variety in what different authors call a ‘seed’. Therefore, no attempt was made to convert ‘diaspore’ masses into ‘seed’ masses. However, we recorded seed mass rather than diaspore mass wherever there was a choice. Obsolete genus names were replaced with valid synonyms from the Vascular Plant Families and Genera data base (http://www.rbge.org.uk/data/vascplnt.html). Species names were then checked against the International Plant Names Index (http://www.ipni.org/index.html). Subspecies and varieties were not recognized in analyses.

Latitude data represent the locations at which species were sampled, rather than the midpoint of species ranges. These data were taken from site descriptions in source papers where possible. Where necessary, latitudes from nearby locations were used in place of exact readings for field sites. Each site was assigned to one of 10 vegetation type categories: rain forest; evergreen sclerophyllous; mixed mountain and highland ecosystems; temperate broadleaf forest; warm desert/semi-desert; grassland; wetland; grassland with emergent shrubs/trees; other; unknown. These categories are based on the scheme provided by Udvardy (1975).

We determined the vegetation type for each location primarily on the basis of the site description in the papers from which seed mass data were taken, but also on the basis of their location. Where species occurred in more than one vegetation type, they were assigned to the category ‘other’.

Growth form and dispersal syndrome data were collected opportunistically from the published literature, from personal communications with researchers (see Acknowledgements) from the Kew Gardens Seed Information Database and from the United States Department of Agriculture ‘Plants’ data base (http://plants.usda.gov/).

Statistics

Seed mass was log-transformed before analyses. Because our main questions were about the global patterns in seed mass, rather than about species traits, the data points analysed were species × site combinations, not species means. However, each of the 7629 species involved was assigned to a single vegetation type, dispersal syndrome and growth form category.

The relationship between seed mass and latitude turned out to be highly non-linear, with a major shift in mean and variance at the edge of the tropics. The unusual shape of this relationship forced us to use slightly unusual statistics to describe it. The model we chose is a stepped model (not to be confused with stepwise regression). This model can be seen in Fig. 1 (the kinked line). Stepped models were calculated in a composite manner. First, we chose the location of the step, using an iterative method to find the latitude at which the \textit{R}^2 of a two-group (temperate and tropical) model was maximized. We then used \textit{spss} to calculate
within-group (temperate and tropical) regressions. These within-group models were combined, and the overall $R^2$ values and $P$ values were calculated in the same way as in a linear regression. That is, $R^2$ values were calculated as the sum of the squared differences between the model’s predicted values and the observed values, divided by the total sum of squares; and $P$ values were calculated from standard $F$ values, which were generated by dividing the mean square of the model by the mean square of the residuals (Sokal & Rohlf, 1995).

Random effects regressions were performed using Genstat version 7.2. The phylogenetic regression was performed using phylocom (http://www.phylodiversity.net/phylocom/). All other analyses were performed with spss version 11.01.

**RESULTS AND DISCUSSION**

**Latitude and seed mass**

There is a negative relationship between seed mass and latitude (Fig. 1a), with geometric mean seed mass decreasing 320-fold between the equator and 60° ($R^2 = 0.24$ with linear regression). The direction and magnitude of the decline in seed mass with latitude is similar to that found previously with smaller samples (Moles & Westoby, 2003).

We had previously thought that there was a relatively continuous decrease in seed size with increasing distance from the tropics (Moles & Westoby, 2003). However, dividing the data into 5-degree
latitudinal bands, and plotting frequency histograms of seed mass distributions (Fig. 1b) clearly reveals a major step down in seed mass, somewhere between 20° and 25° from the equator. We therefore divided the data into two latitudinal groups. These groups have the greatest predictive power over seed mass (R² = 0.25) when the step is located at 23°, very close to the geographical edge of the tropics. Seed mass declines 7.3-fold at this step. However, the magnitude of this step is slightly exaggerated by an over-representation of forest ecosystems in the tropical portion of our data set. Weighting data from different tropical vegetation types according to the area they occupy (Walter, 1973) reduces the step at the tropics to a 5.5-fold decrease.

Not only is there a sudden drop in the mean seed mass at the edge of the tropics, there is also a massive decrease in the variance of seed mass strategies. The average variance in log seed mass in 5° bands between 0° and 20° (1.93) is over twice as high as the average variance in 5° bands between 25° and 60° (0.86). This decrease in variance is largely due to the scarcity of species with very large seeds in temperate ecosystems (Fig. 1b).

Seed mass declines with latitude within both temperate and tropical zones (P < 0.001; Fig. 1b, stepped line). There is a 1.8-fold decrease in mean seed mass for every 10° moved away from the equator in the tropical regression, and a 1.5-fold decrease in mean seed mass for every 10° in the temperate regression. However, the predictive power of the within-zone regressions is very low (tropical R² = 0.01; temperate R² = 0.02) compared to that of the full stepped model, which has an R² of 0.26. Although 26% may not seem like a huge proportion of the variation, it is actually remarkably high, given that only 56% of the variation in seed mass in this data set lies at the cross-site level.

Our main questions in this paper are at the cross-species level, so most of our analyses treat species × site combinations as replicates. However, there is significant site to site variation, presumably due to environmental differences between sites at the same latitude, which cannot be explained purely in terms of species to species variation. We therefore also fitted a mixed model (McCulloch & Searle, 2001) containing terms both for random variation from site to site, and for random variation from species to species, as well as the systematic effect of latitude. The only notable difference between cross-species and mixed model regressions is that the relationship between seed size and latitude within the tropics is not significant in the mixed model. This drop in significance is almost certainly due to the much smaller effective sample size in the mixed model regression (the data come from 1752 sites, only 95 of which are tropical). Both regressions show that over 90% of the variation explained by latitude is due to the term indicating whether a species comes from a tropical or temperate latitude. Mixed model regression also shows that 68% of the variation in seed mass that is not explained by latitude lies at the between species level, compared to 31% at the between-site level.

Finally, we used phylogenetic regression to ask whether divergences in the latitude at which species occur have been consistently associated with divergences in log seed mass. We found a highly significant, yet weak, negative relationship between evolutionary divergence in seed mass and evolutionary divergence in latitude (P < 0.001; R² = 0.03; slope = −0.026). That is, shifts away from the equator have tended to be associated with decreases in seed mass throughout seed plant evolution.

Correlates of the latitudinal gradient in seed mass

General (not generalized) linear models (GLMs, performed in sss version 13.0) revealed significant (P < 0.001) relationships between seed mass and each of the variables predicted to influence the latitudinal gradient in seed mass: plant growth form (R² = 0.40; n = 9605), vegetation type (R² = 0.31; n = 7480), seed dispersal syndrome (R² = 0.19; n = 5145), and NPP (R² = 0.14; n = 10,428).

There are latitudinal gradients in NPP, plant growth form, seed dispersal syndrome and vegetation type (Fig. 2). Growth form and vegetation type show major shifts near the edge of the tropics (Fig. 2b,c), while NPP declines more gradually (Fig. 2a). Thus, the abrupt change in seed mass at the edge of the tropics (Fig. 1) is more closely associated with a shift in plant traits than with a shift in NPP. However, any/all of these variables could contribute to the latitudinal gradient in seed mass.

Next, we asked how much predictive power latitude had over seed size after growth form, dispersal syndrome, vegetation type and NPP had been accounted for. We performed analyses on the 3902 species × site combinations for which all these variables were known. We omitted altitude from these models in order to maintain a large sample size. We began with general linear models that had seed mass as the dependent variable, and different combinations of the other variables as predictors (Table 1). Residuals from these GLMs were then regressed on latitude, using both linear and stepped models. In all cases, stepped models provided significantly better fits than linear models (P ≤ 0.0008). Thus, we report only the results from stepped models here.

Latitude accounts for a significant (P ≤ 0.0001) proportion of the residual variation in seed mass from all single-term models (Table 1), explaining 20% of the residual variation after dispersal syndrome, 6% after NPP, 3% after growth form, and a tiny 1% after vegetation type. The difference in mean residuals between temperate and tropical regions is 2.0-fold after growth form and 1.6-fold after vegetation type, compared to a 26-fold difference in mean seed mass between tropical and temperate regions when neither of these variables has been accounted for. That is, 92% of the step in seed mass at the edge of the tropics is explained by differences in growth form, and 94% by differences in vegetation type. The degree of concordance between changes in vegetation

---

1We do not have data from enough sites between 20° and 25° to be certain of the exact shape of the relationship between latitude and seed mass in this region. However, there is no evidence for a gradual decrease in seed mass in this range.

2This value differs to the step size reported above mostly because it is the difference between the mean tropical and temperate seed masses rather than the size of the difference at 23 degrees, but also because it considers only a subset of the data — there is a 29-fold difference between mean temperate and tropical seed masses in the full data set.
Figure 2 Relationships between latitude and some of the variables hypothesized to affect seed mass. The relationship between latitude and: (a) net primary productivity, (b) plant growth form, (c) vegetation type and (d) seed dispersal syndrome. Each point in (a) represents one study site, while (b–d) show the proportion of species within each 5° latitudinal band recorded as having a given growth form or dispersal syndrome, or as coming from a given vegetation type.

Table 1 Correlates of the latitudinal gradient in seed mass. This table shows the amount of variation in \( \log_{10} \) seed mass explained by general linear models containing the specified terms. \( n = 3902 \) species × site combinations in all models. The fourth and fifth columns show the results of stepped regressions between latitude and residuals from the specified model. The final model listed in the first column was performed using a linear model for all terms.

<table>
<thead>
<tr>
<th>Model terms (main effects only)</th>
<th>Type III sum of squares</th>
<th>( R^2 ) of latitude vs. residuals</th>
<th>( R^2 ) of latitude vs. residuals</th>
<th>P-value of latitude vs. residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude</td>
<td>1517</td>
<td>0.240</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Dispersal syndrome</td>
<td>1361</td>
<td>0.202</td>
<td>0.199</td>
<td>( P &lt; 0.001 )</td>
</tr>
<tr>
<td>NPP</td>
<td>1503</td>
<td>0.224</td>
<td>0.063</td>
<td>( P &lt; 0.001 )</td>
</tr>
<tr>
<td>Growth form</td>
<td>2934</td>
<td>0.436</td>
<td>0.034</td>
<td>( P &lt; 0.001 )</td>
</tr>
<tr>
<td>Vegetation type</td>
<td>1953</td>
<td>0.290</td>
<td>0.012</td>
<td>( P &lt; 0.001 )</td>
</tr>
<tr>
<td>Vegetation type + growth form</td>
<td>3028</td>
<td>0.450</td>
<td>0.003</td>
<td>( P = 0.002 )</td>
</tr>
<tr>
<td>Vegetation type + NPP</td>
<td>2091</td>
<td>0.311</td>
<td>0.003</td>
<td>( P = 0.001 )</td>
</tr>
<tr>
<td>Vegetation type + dispersal syndrome</td>
<td>2470</td>
<td>0.367</td>
<td>0.006</td>
<td>( P &lt; 0.001 )</td>
</tr>
<tr>
<td>Growth form + dispersal syndrome</td>
<td>3315</td>
<td>0.493</td>
<td>0.011</td>
<td>( P &lt; 0.001 )</td>
</tr>
<tr>
<td>Growth form + NPP</td>
<td>3036</td>
<td>0.452</td>
<td>0.004</td>
<td>( P &lt; 0.001 )</td>
</tr>
<tr>
<td>Vegetation type + growth form + dispersal syndrome</td>
<td>3366</td>
<td>0.500</td>
<td>0.003</td>
<td>( P = 0.002 )</td>
</tr>
<tr>
<td>Vegetation type + growth form + dispersal syndrome + NPP*</td>
<td>3393</td>
<td>0.505</td>
<td>0.002</td>
<td>( P = 0.011 )</td>
</tr>
<tr>
<td>Latitude + vegetation type + growth form + dispersal syndrome + NPP</td>
<td>3416</td>
<td>0.508</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>TOTAL</td>
<td>6728</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*A model containing terms for all interactions explained 59% of the variation in seed mass. However, this ‘all interactions’ model has vastly more degrees of freedom than the main effects model, and not all of these terms are biologically meaningful.*
type, growth form and dispersal syndrome, and the latitudinal gradient in seed mass can be seen graphically in Fig. 3.

We wondered whether differences in altitude between sites might be confounding our results. However, the predictive power of altitude is very low ($R^2 = 0.06; n = 8512$), and adding altitude to a model after latitude increases $R^2$ by only 1.2%.

Although vegetation type explains more of the latitudinal gradient in seed mass than any other variable, growth form explains much more of the total variation in seed mass than does vegetation type. This is probably because vegetation type and latitude address only between-site variation, while growth form can account for seed mass differences between species at a site, as well as between species at different sites.

The relationships between latitude and residual variation in seed mass after growth form and dispersal syndrome are not inconsistent with Lord et al.'s (1997) finding of latitudinal gradients in seed mass within dispersal modes and growth forms. However, we would expect within-category relationships to be relatively weak. Here, we were able to compare the strength of relationships between seed mass and latitude within growth forms and dispersal syndromes with the strength of cross category relationships (see Appendix S1 in Supplementary Material).

The predictive power of relationships within-growth form/dispersal syndrome categories is quite low (most $R^2 < 0.1$), and all but one has a shallower slope than the cross-species regression (many less than half as steep; Appendix S1). As expected, the relationships between seed size and latitude are much stronger within dispersal syndromes than within growth forms. Interestingly, the relationship between seed mass and latitude is extremely weak within shrubs and herbs ($R^2 < 0.005$; slope less than a tenth as steep as the cross-species slope), but relatively strong within trees ($R^2 = 0.16$, slope 80% as steep as the cross-species slope).

Finally, we ask how much of a latitudinal gradient in seed mass remained after variation due to growth form, vegetation type, dispersal syndrome and NPP had been accounted for. A model including the main effects of all these terms explains 51% of the

Figure 3 The relationship between seed mass and latitude. Species coded by (a) vegetation type, (b) growth form and (c) dispersal syndrome. Sample sizes as in Fig. 1(b).
variation in seed mass. The relationship between latitude and the residual variation from this model is significant \((P = 0.011;\ \text{Table 1})\), but its explanatory power is extremely small \((R^2 = 0.002)\). Some of this residual variation is probably attributable to the latitudinal gradient in the length of the growing season — a factor that we were unable to investigate here. The difference in mean residuals between temperate and tropical regions is just 1.2-fold: 60% of the difference in residuals from a model only containing growth form, and less than 5% of the overall difference in mean seed mass between tropical and temperate regions. Thus, the latitudinal gradient in seed mass (including the step at the edge of the tropics) is almost entirely explained by changes in seed dispersal syndrome, vegetation type, NPP and plant growth form.

The fact that growth form is the strongest correlate of seed mass in this global data set is consistent with previous studies on the evolutionary history of seed mass, and with previous cross-species compilations (Leishman et al., 1995; Moles et al., 2005a, 2005b). Thus, this pattern is consistent through both time and space. Recent work suggests that the positive correlation between seed mass and plant size results from a coordination of life-history variables (Charnov, 1993; Moles & Westoby, 2006). Species with large adult size require a long juvenile stage to reach maturity (Moles et al., 2004). In order to survive this long juvenile phase, offspring need high rates of survival. This is achieved by producing larger seeds (Moles & Westoby, 2003). Combining the present results with this life-history theory raises two important questions for the future.

1. Does the latitudinal gradient in seed size reflect a latitudinal gradient in the ratio of r-selected to K-selected species (Pianka, 1970), or do relationships among body size, offspring size, growth rates and lifespans shift along the latitudinal gradient? To answer this question, we will need to quantify the latitudinal gradients in plant height, plant lifespan and longevity, and determine whether these traits display a similar shift at the edge of the tropics to that found here with seed mass.

2. Why is there a shift in plant strategy at the edge of the tropics? Primary productivity (Fig. 2a) is relatively constant between 20° and 30°, yet we see a major shift in seed mass — a trait central to many aspects of plant ecology. Could this strategy shift reflect lower levels of environmental variability in tropical environments?

Answering these, and similar questions will greatly increase our understanding of the mechanisms underlying large-scale ecological patterns.

Plant ecology is now entering a phase where global data sets are gathered together. Because there have been smaller studies previously, it is not expected that this scale-shift will uncover wholly unforeseen patterns. The benefits lie in other directions. First, global coverage overcomes the problem of knowing whether particular results extrapolate to other continents or climates. Second, we are building a firm picture of the relative strengths of different patterns. While local studies remain important for our understanding of individual species and communities, we are entering a time when the tools and information are available to refocus ecological studies onto the strongest and most consistent patterns defining our natural world. We hope that such compilations will become more common. By increasing our understanding of how ecological traits and processes vary along present-day climatic gradients, we will be much more likely to be able to predict accurately how global warming, and perhaps other human-induced environmental changes are going to influence natural plant and animal communities in the future.

ACKNOWLEDGEMENTS

Thanks to R. Condit, S. Diaz, P. Juniper, M. Leishman, J. Lord, M. Mayfield, B. Rice, K. Thompson, I. Wright and S. J. Wright, who gave us access to unpublished data. Thanks to two anonymous referees, who provided helpful comments on an earlier version of this manuscript. This work was supported by the National Center for Ecological Analysis and Synthesis, Australian Research Council Funding to M.W., and a National Science Foundation grant to C.O. Webb, M.J. Donoghue and D.D.A. The Millennium Seed Bank Project is funded by the UK Millennium Commission, The Wellcome Trust and Orange PLC. The Royal Botanic Gardens, Kew are partially funded by the UK Department of Environment Food and Rural Affairs.

REFERENCES


**BIOSKETCH**

**Angela Moles** is a postdoctoral researcher from New Zealand. The present paper brings together her two main research interests: understanding the ecology of seed mass, and quantifying global patterns in the ways different species of plants make a living. Since finishing this compilation, Angela has become an eco-nomad, in order to work on ‘The World Herbivory Project’. In this project, Angela and her collaborators are establishing study sites in 80 ecosystems all around the world, to quantify the latitudinal gradient in the strength of plant–animal interactions.

**SUPPLEMENTARY MATERIAL**

The following material is available online at www.blackwell-synergy.com/loi/geb

**Appendix S1** Relationships between seed mass and latitude within growth form, dispersal syndrome and growth form-dispersal syndrome combinations.

Editor: David Currie