Research Article
On the Relationship between Pollen Size and Genome Size

Charles A. Knight,1 Rachel B. Clancy,1 Lars Götzenberger,2 Leighton Dann,3 and Jeremy M. Beaulieu4

1 Department of Biological Sciences, California Polytechnic State University, San Luis Obispo, CA, USA
2 Department of Botany, Institute of Ecology and Earth Science, University of Tartu, Tartu, Estonia
3 Robinson College, Cambridge, CB3 9AN, UK
4 Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

Correspondence should be addressed to Charles A. Knight, knight@calpoly.edu

Received 22 December 2009; Accepted 2 April 2010

Academic Editor: Ilia Judith Leitch

Copyright © 2010 Charles A. Knight et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Here we test whether genome size is a predictor of pollen size. If it were, inferences of ancient genome size would be possible using the abundant paleo-palynological record. We performed regression analyses across 464 species of pollen width and genome size. We found a significant positive trend. However, regression analysis using phylogenetically independent contrasts did not support the correlated evolution of these traits. Instead, a large split between angiosperms and gymnosperms for both pollen width and genome size was revealed. Sister taxa were not more likely to show a positive contrast when compared to deeper nodes. However, significantly more congeneric species had a positive trend than expected by chance. These results may reflect the strong selection pressure for pollen to be small. Also, because pollen grains are not metabolically active when measured, their biology is different than other cells which have been shown to be strongly related to genome size, such as guard cells. Our findings contrast with previously published research. It was our hope that pollen size could be used as a proxy for inferring the genome size of ancient species. However, our results suggest pollen is not a good candidate for such endeavors.

1. Introduction

Pollen range in size by over three orders of magnitude [1, 2] (Figure 1). The variation in pollen size may stem from strong selection pressures related to pollen dispersal strategies. For example, wind-pollinated species may achieve long-distance transport by having pollen that are (1) small, (2) light weighed, (3) dehydrated, and (4) that have shapes conducive to wind capture [1–3]. However, some gymnosperms have large pollen but are also wind pollinated (Pinaceae and Podocarpaceae) [4]. Two air-filled sacs (sacci) facilitate wind dispersal in these groups [3, 4]. Pollen of species that use insect facilitated dispersal can sometimes be quite large, but we are not aware of any study showing that pollen dispersed by insects is generally larger than pollen dispersed abiotically. However, there is greater interspecific variability for pollen grain size in species that use insect dispersal [5–7]. Understanding what controls pollen size from a developmental perspective will enhance our understanding of the ecological significance of variation in pollen size.

It has frequently been observed that pollen size is related to the length of the style (see [8, 9] and citations therein). Delpino [8] suggested that larger pollen grains contain more resources for the growth of pollen tubes and therefore larger pollen is better suited to fertilize flowers with longer styles. Darwin [10] disagreed with this proposal, suggesting that pollen tube growth was facilitated by resources garnered from the style. Closely related species sometimes exhibit extreme variation in pollen size and style length. A change in style length may ensure reproductive isolation, especially if style length increases with pollen size, and larger pollen may be necessary for pollination of flowers with longer styles [8, 9]. Conversely, there may simply be inherent allometric determinants of organ size that are shared between both pollen, styles, and other plant parts (see [11, 12] for a current review of genetic determinants of organ size).

Recently, Beaulieu et al. [13] found a strong positive relationship between genome size and cell size, leaving open the possibility that genome size may partly determine, or be correlated with pollen size. A pollen grain consists of
Journal of Botany

Myosotis palustris
Boraginaceae
very small (<10 μm)

Lotus corniculatus
Fabaceae
small (10–25 μm)

Staphylea pinnata
Staphyleaceae
medium sized (25–50 μm)

Linum tenuifolium
Linaceae
large (50–100 μm)

Cucurbita pepo
Cucurbitaceae
very large (>100 μm)

Figure 1: Pollen varies considerably in size. (a) Images of pollen at the same scale and (b) (Inset on a): a histogram of pollen widths showing a log normal distribution.

2. Methods

Estimates of DNA content were compiled from the Plant DNA C-values database maintained at the Royal Botanical Gardens, Kew [18]. Equatorial diameters for spheroidal (or near spheroidal) pollen were compiled from various sources including: (1) The Northwest European Pollen Flora periodically monographed by family in the Review of Palaeobotany and Palynology (114 species) [20–23] and others, (2) the palynological database (http://www.paldat.org/) an online publication of the Society for the Promotion of Palynological Research in Austria (122 species), (3) direct measurements by Leighton Dann using light microscopy (157 species—water suspension), and (4) various primary literature sources (71 species). For gymnosperms equatorial diameters only included the central sphere, not the peripheral structures. All values of genome size and pollen width are listed in our supplementary table mentioned in SM available online at doi: 10.1155/2010/612017.

We used Phylomatic (tree version: R20080417.new, maintained by C. A. Webb, http://www.phylodiversity.net/phyllomatic) to construct a “mega-tree” hypothesis for the species in our sample. Phylomatic is a compilation of previously published phylogenies and its ordinal “backbone” and family resolutions are based on the Angiosperm Phylogeny Website (APweb) [24]. The program matches a species to a reference tree first by “genus”, then by “family”. Most relationships among and within “genera” are returned as a polytomy due to insufficient resolution within the reference tree at this phylogenetic scale. Branch length information is taken from the single fossil-calibrated molecular divergence time estimates mentioned [25]. We fixed these age estimates and provided dates to undated nodes by distributing...
them evenly between nodes with known ages and terminal taxa.

We used R (R Development Core Team, 2009) to obtain slope estimates and $R^2$ from regression models. Independent contrasts were calculated across our phylogeny using Phylocom (V.4.1; [26]). The method of independent contrasts iteratively calculates trait differences (termed “contrasts”) between extant “species” pairs, and subsequently their weighed internal node averages, starting at the tips and moving down to the root of a phylogeny [26]. This calculation transforms the data into $N-1$ independent data points, each representing an evolutionary divergence. For consistency, the sign of the contrast for the independent variable (e.g., genome size) is set to always be positive with the contrasts of the dependent variable (e.g., pollen width) being compared in the same direction. These contrasts are then standardized by their branch length information to ensure statistically independent data, drawn from a normal distribution with equal variances, which can be analyzed using conventional statistics [27, 28]. Note that since the direction of subtraction in an independent contrast analysis is arbitrary, reversing the direction of subtraction would result in a contrast of the opposite sign. This property gives the expected mean value of zero to all contrasts. Therefore, all regression analyses forced the line through the origin [28].

We calculated a contribution index to examine the proportion of the variation, each divergence contributes to the present-day variation observed in our pollen width data. The contribution index is the product of the amount of variation within a focal clade that is from a particular focal divergence and the amount of the total variation within that focal clade compared with the whole tree (for a detailed discussion, see [29]). That is, large divergences leading to a large number of descendents with a large spread in trait data typically result in higher contribution index scores. Contribution index scores were taken directly from the Phylocom output.

To test whether recent divergences were more likely to lead to dramatic changes in both genome size and pollen width, we preformed two separate but similar analyses. First, we analyzed the independent contrast output for contrasts involving sister tip taxa (i.e., node depth equals 1) and compared this to the complete independent contrast output. The advantage of this approach is that it is completely objective, however, the limitation is that tip taxa contrasts could really be quite divergent because of lack of sister group representation in our dataset. Second, we examined how genome size and pollen width varied genus by genus.

3. Results

Genome size and pollen width information for 464 species was obtained and is summarized in Table 1. The species comprised 50 orders and 85 families of Spermatophyta (seed plants; [30]). The angiosperms made up a majority of the dataset (437 out of 464 species) and contained representatives from the major clades: Magnoliidae (magnoliids; 2 species), Monocotyledonae (monocots; 76 species), and the Eudicotyledonae (eudicots; 359 species). Only three families (Cupressaceae, Pinaceae, and Taxodiaceae) represented the extant lineages of gymnosperms (Acrogynospermae; [26]) and all are from the Coniferae. The mean 1C DNA estimates for this sample (1C = 22,883.6 Mbp) is comparable to the mean of the acrogynosperms (1C = 18,111.2 Mbp) taken from the Plant DNA C-values database [19].

Pollen width varied nearly three orders of magnitude, or 2.4-fold, from 7 to 167 μm. The average pollen width was 39.5 μm. Oenothera biennis had the largest pollen size (167 μm), while Myosotis scorpioides had the smallest pollen size (7 μm) (Table 1). Unlike the 1C DNA data, the mean of the magnoliids was larger (pollen width = 59.0 μm) than the monocots (pollen width = 48.3 μm) and eudicots (pollen width = 35.4 μm). However, the mean of the gymnosperms (pollen width = 67.2 μm) was larger than all three major groups of flowering plants. Of the 21 families that had more than five species represented in our sample, Onagraceae had the largest mean pollen width at 113.9 μm, while Plantaginaceae had the smallest mean pollen width at 23.6 μm.

The combined data sources showed a significant positive trend ($n = 464$, slope $= 0.104$, $R^2 = 0.096$, $P$-value $< .001$, Figure 2(a)). However, our phylogenetically independent contrast analysis suggested that there was a large split between Angiospermae versus Acrogynospermae (gymnosperms) for both pollen width and genome size (Table 2), but otherwise, divergences in genome size and pollen width did not co-vary with evolutionary divergences ($n = 197$ contrasts, slope $= 0.04$, $P > .05$, Figure 2(b)). There were 71 congeneric species pairs in our dataset. Of these, there were significantly more with a positive relationship between genome size and pollen width (44/71, sign test $P < .05$). Twenty-seven of these congeneric pairs had either no relationship (slope $= 0$) or a negative relationship.

Our literature review of ploidy and pollen width showed consistent reports of pollen width increasing with ploidy (Table 3): results show that pollen size increased by 1.1x to 2x with a doubling of DNA content.

4. Discussion

The consistent strong positive trend that Beaulieu et al. [13] found between plant cell size and genome size is weakly reflected in our analysis of pollen grains. Our regression test was significant across 464 species, but phylogenetically independent species contrasts suggest that the relationship was largely driven by early major divergences during seed plant evolution (between the Angiospermae versus Acrogynospermae, e.g., see Table 2 for other significant divergences). At the more microevolutionary level, congeneric species did tend to support the trend of increasing pollen width with increasing genome size, but again, divergences across all taxonomic levels did not support a general evolutionary trend. Previous investigators have found repeated instances of increased pollen width with increasing ploidy levels (Table 3). Our conclusion from these observations is that (1) if there is a relationship between genome size and pollen width, it is more likely exposed at the microevolutionary level, especially when divergences involve variation in ploidy.
Table 1: Summary statistics for pollen size and genome size (1C Mbp) for the major groups of plants analyzed in this study.

<table>
<thead>
<tr>
<th></th>
<th>Gymnosperms</th>
<th></th>
<th>Angiosperms</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All Data (N = 464)</td>
<td>Coniferae (N = 27)</td>
<td>Magnoliidae (N = 2)</td>
<td>Monocots (N = 76)</td>
</tr>
<tr>
<td>pollen size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smallest</td>
<td>7.00 μm</td>
<td>15.0 μm</td>
<td>44.0 μm</td>
<td>17.0 μm</td>
</tr>
<tr>
<td>Largest</td>
<td>617.0 μm</td>
<td>108.0 μm</td>
<td>74.0 μm</td>
<td>150.0 μm</td>
</tr>
<tr>
<td>Mean</td>
<td>39.5 μm</td>
<td>67.2 μm</td>
<td>59.0 μm</td>
<td>48.3 μm</td>
</tr>
<tr>
<td>genome size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smallest</td>
<td>142 Mb</td>
<td>9727 Mb</td>
<td>784 Mb</td>
<td>294 Mb</td>
</tr>
<tr>
<td>Largest</td>
<td>80,262 Mb</td>
<td>31,674 Mb</td>
<td>4753 Mb</td>
<td>80,262 Mb</td>
</tr>
<tr>
<td>Mean</td>
<td>6540 Mb</td>
<td>22883 Mb</td>
<td>2768 Mb</td>
<td>16,414 Mb</td>
</tr>
</tbody>
</table>

Table 2: Contribution index scores (with rank) for divergences in pollen width and 1C DNA content for the species in our sample.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Pollen width contribution</th>
<th>Divergences making the largest contribution</th>
<th>1C DNA content rank</th>
<th>1C DNA content contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.134</td>
<td>Angiospermae versus Acrogynnospermae</td>
<td>1</td>
<td>.384</td>
</tr>
<tr>
<td>2</td>
<td>0.050</td>
<td>Polytomy at the origin of Coniferae</td>
<td>194</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>3</td>
<td>0.046</td>
<td>Divergence at the origin of Papilionoidae</td>
<td>99</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>4</td>
<td>0.041</td>
<td>Divergence between Lythraceae and Onagraceae</td>
<td>138</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>5</td>
<td>0.036</td>
<td>Magnoliidae versus Eudicotyledonae</td>
<td>3</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>6</td>
<td>0.032</td>
<td>Divergence of Fagaceae and the rest of Fagales</td>
<td>81</td>
<td>.002</td>
</tr>
<tr>
<td>7</td>
<td>0.031</td>
<td>Polytomy at the origin of eurosid II</td>
<td>127</td>
<td>.002</td>
</tr>
<tr>
<td>8</td>
<td>0.030</td>
<td>Divergence between Zingiberales and Poales</td>
<td>175</td>
<td>.002</td>
</tr>
<tr>
<td>9</td>
<td>0.026</td>
<td>Divergence between Solanales and Lamiales</td>
<td>35</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>10</td>
<td>0.026</td>
<td>Divergence at the origin of Malvaceae</td>
<td>136</td>
<td>.002</td>
</tr>
</tbody>
</table>

Figure 2: (a) Scatter plot of the significant positive associations between genome size and pollen width. The slope was estimated using conventional least-squares methods that do not incorporate the correlated error structure due to phylogeny. (b) Independent contrast results showing that divergences in 1C DNA content are not associated with divergences in pollen width (open and black points). This result was consistent when isolating the results to just bifurcating sister tip taxa (black points). The unfilled points represent deeper nodes. A line is not shown because the relationship was not significant.
level, and (2) there was a significant divergence in both genome size and pollen width with the basal divergence between Angiospermae versus Acrogynospermae. While the ploidy results suggest a mechanistic link between genome size (of bulk DNA content) and pollen width, the basal divergence between Angiospermae versus Acrogynospermae may simply be a coincidence. Our results could also be explained by strong selection for pollen to be small which overwhelms any direct mechanistic link between genome size and pollen size (if there is any).

Natural selection may act strongly on pollen size, especially in relation to pollen dispersal strategies. However, even within species that are primarily bee pollinated, there is considerable variation in pollen size, even though they have very similar genome sizes (e.g., Luffa and Lotus in Figure 3). In contrast, the sometimes wind-pollinated Brassica napus [31] has small pollen (compared to Luffa and the rest of our dataset, Figure 3), but Brassica napus is also frequently insect pollinated [32]. Complicating matters, in some cases plants are self-compatible and can complete pollination without a vector. Brassica napus also fits into this category, it is self-compatible and capable of autonomous pollination [33]. Even in the absence of pollinators, it is able to set half of its seeds in still air and 80% when the stem is shaken [33]. Furthermore, pollination efficiency is considerably affected by local and seasonal environmental conditions [3, 34].

In comparison to other plant phenotypic traits, pollen size varies somewhat less. Pollen size varied in our sample over three orders of magnitude. However, seed mass and genome size vary over ten and five orders of magnitude, respectively [18, 29]. Why is there so little variation in pollen size? There is strong selection favoring small pollen size (as noted above), and likewise, selection pressures favoring small pollen size continually reduces pollen size unless this pressure is relaxed.

The relationship between cell size and genome size may arise from the greater necessity of gene transcripts to service larger cytoplasms [35]. However, pollen is not metabolically active after dehiscence, but rather become so soon after imbibition and pollen germination. This quiescence makes them quite different from guard cells and other cell types whose sizes have previously been shown to be strongly related to genome size [13, 36]. Perhaps the maximal volume of the pollen tube, after its metabolically active growth stage, may be a better measure of pollen size in this context, not the recently hydrated sphere.

Several measurement errors could have contributed to our weak results. Some of our measurements may have come from unhydrated or incompletely hydrated pollen. In addition, methods of hydration varied from water to glycerine jelly, or silicon oil, each of which can result in different final volumes [37]. Further, various methods of imaging were used, including scanning electron microscopy and light microscopy. There was also no control for the type of pollen reserve (starch or lipids). Each time a new type of pollen reserve (starch or lipids). Each time a new measurement was made, the pollen sample was divided into two and one half was left hydrated and the other one quarter was left unhydrated. For each method of imaging, the pollen samples were divided into two and one half was left hydrated and the other one quarter was left unhydrated. For each method of imaging, the pollen samples were divided into two and one half was left hydrated and the other one quarter was left unhydrated.

Table 3: Examples of previous studies on the relationship between ploidy and pollen size reported by family, genus, ploidy variation and how doubling DNA content changed pollen volume. Primary literature sources are also given.

<table>
<thead>
<tr>
<th>Family</th>
<th>Genus</th>
<th>Species and chromosome numbers</th>
<th>2x the DNA led to:</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boraginaceae</td>
<td>Lappenda</td>
<td>deflexa (2n = 24) &amp; squarrosa (2n = 48)</td>
<td>1.5x to 2x larger pollen</td>
<td>[20]</td>
</tr>
<tr>
<td>Convolvulaceae</td>
<td>Cuscuta</td>
<td>epithymum (2n = 14) &amp; carapestris (2n = 56)</td>
<td>1.2x larger pollen</td>
<td>[22]</td>
</tr>
<tr>
<td>Papaveraceae</td>
<td>Fumaria</td>
<td>marialis (2n = 30) &amp; capreolata (2n = 60)</td>
<td>1.3x larger pollen</td>
<td>[21]</td>
</tr>
<tr>
<td>Poaceae</td>
<td>Andropogon</td>
<td>various species (2n = 60, 120, and 180)</td>
<td>1.2x larger pollen</td>
<td>[14]</td>
</tr>
<tr>
<td>Polygonaceae</td>
<td>Rumex</td>
<td>acetosella (2n = 14, 28, and 42)</td>
<td>1.1x to 1.3x larger pollen</td>
<td>[23]</td>
</tr>
<tr>
<td>Convolvulaceae</td>
<td>Ipomea</td>
<td>trifida with diploid pollen</td>
<td>1.3x larger pollen</td>
<td>[15]</td>
</tr>
<tr>
<td>Brassicaceae</td>
<td>Arabidopsis</td>
<td>thaliana with diploid and tetraploid pollen</td>
<td>1.7x larger pollen</td>
<td>[17]</td>
</tr>
</tbody>
</table>

![Figure 3](image3.png)

**Figure 3:** Pollen size can vary considerably (13–150 μm) within a narrow range of genome sizes (833–1078 Mbp).
Our results suggest that this effort would be difficult and perhaps misleading. Fortunately, the morphology of pollen grains seems to have enough stasis so that species or group level identification is accurate through the paleobotanical record.

Acknowledgments

The authors thank Professor Martina Weber from PalDat (http://www.paldat.org/) for the preparation of Figure 1 and the use of images presented in Figure 3. This project was supported by a sabbatical grant to C. K. and a research stipend for R. C. from the Biological Sciences Department at California Polytechnic State University, San Luis Obispo. They also thank Dr. Gerhard Leubner for hosting them at the University of Freiburg, Germany, where this project was completed.

References


[34] K. E. Hayter and J. E. Cresswell, “The influence of pollinator abundance on the dynamics and efficiency of pollination in agricultural *Brassica napus*: implications for landscape-scale

