**INTRODUCTION**

The functional significance of the profound variation in plant genome size is still largely unknown and represents one of the most significant unanswered questions in plant biology. There has been great interest in linking this variation to plant phenotypic traits. Early observations showed that genome size was positively correlated with cell size, and the duration of meiosis, formed the basis of hypothesized genome size consequences at higher phenotypic scales (Bennett, 1971, 1972, 1987). This scaling was supported by numerous studies showing a positive relationship between genome size and seed mass (see table 4 in Knight et al., 2005, for a list of studies), leaf anatomical traits (Castro-Jimenez et al., 1989; Chung et al., 1998; Wakamiya et al., 1993) and growth rate (see table 6 in Knight et al., 2005, for a list of studies). In addition, several studies have documented relationships between environmental conditions (temperature, water availability, latitude and elevation) and genome size (reviewed by Knight and Ackerly, 2002). These environmental predictors of genome size might be a consequence of genome size change.

To date, most studies have investigated limited subsets of species (often performed within individual genera or families). Therefore, the generality of genome size correlations with phenotypic traits is unknown. It is possible that the there is no direct, or general phenotypic consequences of variation in plant genome size (Oliver et al., 2007), or it may be that any significant associations are mitigated through some other third factor (Beaulieu et al., 2007b). Regardless, establishing the generality of any phenotypic correlation with genome size is a logical first step. The analyses presented here have benefited from large cross-species comparisons involving diverse assemblages of angiosperm and gymnosperm species (often involving 100 or more species). The goal is to generalize the phenotypic consequences (or correlations) of genome size variation at several different phenotypic scales (Fig. 1). It is our hope that after the generality of these patterns has been established, new mechanistic hypotheses will be proposed, and experiments performed to explain how genome size affects the phenotype.

Pattern searching across large comparative data sets is prone to mistakes when phylogenetic information is not taken into consideration. While a simple regression or correlation statistic may document a predictive relationship between two traits, evolutionary inferences should not be drawn from regression analyses. For evolutionary inferences, methods to incorporate the phylogenetic tree of a species set should be used (see Felsenstein, 1985; Harvey and Pagel, 1991; Garland et al., 1992; http://www.phylodiversity.net/phylocom/). Because of the centrality
of the independent contrast method to pattern searching in comparative biology, and because of the continued publication of studies in which independent contrasts should be used but are not, a brief tutorial on how to carry out these analyses is presented in Fig. 2.

Here we review a series of analyses on the relationship between genome size and the phenotype in order of increasing scale (roughly equating to the number of cells required to produce a particular phenotypic trait): cell size (guard cell and epidermal cell), stomata density, seed mass, leaf mass per unit area (LMA), wood density, photosynthetic rate and finally maximum plant height (Fig. 1). The analyses reviewed below involved two metrics of cellular DNA content: first, the 2C DNA content, which is the total amount of DNA in an unreplicated somatic cell; and second, the 1Cx DNA content, which is the ploidy-corrected monoploid genome size of a species (sensu Greilhuber et al., 2005). As similar results were found for both metrics we just give the data for 2C DNA amounts and refer to this as 'genome size' correlations throughout.

THE SCALING OF GENOME SIZE AND CELL SIZE

Guard cells are one of the smallest cell types in plants. They are good candidates for this study because they rarely undergo endoreduplication while other leaf cells may be highly endoreduplicated (i.e. Arabidopsis; Melaragno et al., 1993). Furthermore, it is conceivable that selection pressures operate strongly on guard cell size parameters, and thus, if genome size increases manifest with increased guard cell size too, the generality of the relationship is strengthened. We tested the hypothesis that genome size is correlated with cell size using three different cell types (guard cells, epidermal cells and unicellular diatom cell volumes) (Beaulieu et al., 2007c). Briefly, epidermal impressions (made using clear nail varnish) of a diverse assemblage of angiosperms were taken from the living collections at Royal Botanical Gardens, Kew (102 species in total). Guard cell length on both the abaxial and the adaxial surfaces of mature, fully expanded leaves was measured from digital photographs of these impressions (further description of the methods will be given in a forthcoming paper).

A strong positive association was found between genome size and guard cell length, with genome size explaining 61% of the variation in guard cell length (Figs 3 and 4A, Table 1). Independent contrast analyses supported this conclusion, albeit with a slightly lower $R^2$ value (0.61 vs. 0.42). This is a remarkably strong relationship. What is surprising is that guard cell length seems to scale directly with genome size rather than setting a minimum threshold (Figs 3 and 4A). It seems plausible that guard cells may be large for species with large genome sizes due to space constraints imposed by an increase in the amounts of bulk DNA. However, it is not obvious why species with small genome sizes have small guard cell sizes, as there could be many other factors that may increase guard cell size irrespective of genome size. However, that is not what the data show.

Epidermal cell areas are intricately margined and therefore no simple dimensional analysis other than two-dimensional projected area would suffice for cell size. A strong positive association was found between genome size and epidermal cell area, with genome size explaining 59% of the total variation in area (Fig. 4B, Table 1). Again, this relationship was quite linear, with no indication of threshold effects. Species with small genome size had small epidermal cell areas, and species with large genome size had large epidermal cell areas. The observation was supported by independent contrast statistics, albeit with a lower percentage of the variation explained (0.59 vs. 0.22).

Recently, a significant relationship between genome size and unicellular diatom cell volume was reported for unicellular diatoms (Connolly et al., 2007). They studied 16 species and found a significant positive relationship that explained 69% of the variation in cell volume for these diatoms. This result was supported using independent contrasts. These results, combined with similar trends in the animal world (see Discussion), suggest that cell volume scaling with genome size is a general phenomenon for all life.

GENOME SIZE SCALING TO STOMATAL DENSITY AND PHOTOSYNTHETIC RATE

Stomatal density on the abaxial surface was measured using the same images collected to measure guard cell length (Beaulieu et al., 2007c). There was a strong link between stomatal density and genome size (Fig. 4C). As genome
size increases, stomatal density decreases. However, genome size explains less of the variation in stomatal density ($R^2 = 0.34$) than either guard cell length or epidermal cell size (Table 1). This is the first hint that genome size effects diminish as we move up in phenotypic scale. The link between genome size and stomatal density was supported with independent contrast statistics, although again with lower percentage of the variation explained (0.32 versus 0.18). Decreased stomatal density is partly determined by increasing epidermal cell area (Beerling and Chaloner, 1993). Furthermore, stomatal density also decreases with increasing guard cell size (Heatherington and Woodward, 2003; Beaulieu et al., 2007c).

It seems possible that changes in stomatal density could affect the gas exchange characteristics of a species, including the transpiration and photosynthetic rate. Previously, we measured photosynthetic rate for 112 angiosperm species with known genome size that were growing in the living collections at the Royal Botanical Gardens, Kew (Beaulieu et al., 2007b). The data for photosynthetic rate highlight the importance of using independent contrast methodology. There is a large basal divergence between angiosperms and gymnosperms (Fig. 4D). If a correlation is plotted across all species, it is significant, and negative (Table 1). However, this is largely due to the fact that gymnosperms as a group have, on average, larger genomes and lower photosynthetic rates, while angiosperms typically have smaller genome sizes, and higher photosynthetic rates. Within angiosperms alone, there is no relationship between genome size and photosynthetic rate. There is a weak negative relationship within gymnosperms alone. Splitting the data into these two groups is the first step to incorporating the evolutionary history of species; however, the problem spans every divergence in the phylogeny. Independent contrast analyses reveal that there has not been correlated evolution between genome size and photosynthetic rate both across all species and within angiosperms alone.

**GENOME SIZE SCALING WITH SEED MASS**

Beaulieu et al. (2007a) examined the relationship between genome size and seed mass by testing the relationship across 1222 species, from 139 families and 48 orders of seed plants using information from the Seed Information Database (SID; Flynn et al., 2004). They found that there was no significant linear regression relationship between genome size and seed mass across 1222 species despite the multitude of studies that have documented such trends in smaller subsets of species. However, they did discover a unique threshold effect of genome size variation. Species with very large genome sizes never had small...
seeds, while species with small genome sizes had a large range of seed sizes (Fig. 5A). Independent contrast analyses increased the percentage variation in seed size explained by genome size, although the percentages were very small (Table 1). By plotting the slopes within all congeneric sets of species, both positive and negative slopes were apparent (Fig. 5A). However, by sign test, there were significantly more positive slopes than negative slopes. Interestingly, of all predictors of seed mass, genome size ranks quite highly on a recent list compiled by Moles et al. (2005).

**Table 1. Regression and independent contrast statistics for the relationship between genome size (2C DNA content) and each of the eight phenotypic traits analysed**

<table>
<thead>
<tr>
<th>Trait</th>
<th>Regression</th>
<th>Independent contrasts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slope</td>
<td>$R^2$</td>
</tr>
<tr>
<td>Maximum plant height (m)</td>
<td>-0.21</td>
<td>0.03</td>
</tr>
<tr>
<td>$A_{\text{mass}}$ (nmol g$^{-1}$ s$^{-1}$)</td>
<td>-0.02</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Wood density (kg m$^{-3}$)</td>
<td>-0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>LMA (g m$^{-2}$)</td>
<td>-0.10</td>
<td>0.03</td>
</tr>
<tr>
<td>Seed mass (mg)</td>
<td>-0.02</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Stomatal density (no. mm$^{-2}$)</td>
<td>-0.36</td>
<td>0.34</td>
</tr>
<tr>
<td>Epidermal cell area ($\mu$m$^2$)</td>
<td>0.56</td>
<td>0.59</td>
</tr>
<tr>
<td>Guard cell length ($\mu$m)</td>
<td>0.20</td>
<td>0.61</td>
</tr>
</tbody>
</table>

These data are for relationships within angiosperms only. Traits are arranged in the same order as in Fig. 1.

**Genome Size Scaling with LMA and Wood Density**

Following on from genome size to cell size the next logical phenotypic level is at the density of plant material. The fact that much of the biomass of plant material is composed of cell walls and that larger cells have a smaller ratio of cell wall per unit volume lead to the prediction that increasing cell size should lead to decreasing cell and mass density (on a dry weight basis). Likewise, because the relationship between genome size and cell volume is so robust, density...
parameters such as LMA and wood density should be correlated with genome size. We envisaged that these links would be negative; as genome size increases, density would decrease. LMA is perhaps the most predictive trait for plant physiology. If you know the LMA of a species, you can make a reasonable prediction of its leaf life span, growth rate, photosynthetic rate, nitrogen content and many other traits (Wright et al., 2004). Nevertheless, the genetic basis of variation in LMA is largely unknown; therefore, it was of interest to us to test whether genome size variation was associated with LMA variation.

LMA data were collected for all species for which we had measured the photosynthetic, and additional observations were added from Glopnet (Wright et al., 2004). Across 274 species of both angiosperms and gymnosperms there was a weak positive relationship that was significant. However, the relationship was significant and negative within angiosperms and significant and positive within gymnosperms (Fig. 5B). Interestingly, independent contrast results showed that there has been significant positive correlated evolution between LMA and genome size that was driven by divergences in angiosperms. Again, however, the relationship was quite weak ($R^2 = 0.05$; Table 1).

Wood density information was generously provided by Nathan Swenson, who previously reported the database in Swenson and Enquist (2007). Across 200 species, wood density followed a similar trend, with a marked difference between angiosperms and gymnosperms, with gymnosperms having marginally less dense wood on average and significantly larger genome sizes (Fig. 5C). However, both regression and independent contrast analyses failed to uncover any significant relationship between the two traits (Table 1).

**GENOME SIZE SCALING WITH PLANT HEIGHT**

Maximum plant height information was obtained from both Glopnet (Wright et al., 2004) and the SID (Flynn et al., 2004). There was a triangular relationship between genome size and maximum plant height across 324 species of angiosperms. As genome size increases maximum plant height decreases within angiosperms (Fig. 5D). This relationship was significant for regression analyses but not for independent contrasts.

**CONCLUSIONS**

The relationship between genome size and phenotypic traits decreases at higher phenotypic scales. This is somewhat surprising given the strength of the relationship at the cellular level. By contrast, it seems that compensatory mutations have occurred such that leaf and wood density are largely
unaffected by changes in cell size, and likewise, variation in seed mass is only marginally affected by changes in genome size, and there is no relationship with photosynthetic rate. Interestingly, there is a significant association with maximum plant height.

The genome size effect on cell size is not unique to seed plants, as other investigators have documented positive relationships with animal cell sizes, such as for red blood cell size in fishes, amphibians, reptiles, birds and mammals (Gregory, 2005). The generality of this phenomenon begs for a better mechanistic understanding for why it exists. Because of the linearity of the response, it appears that there is a functional relationship. In addition, guard cell sizes proportionally increase in polyploid series (Masterson, 1994), also suggesting a direct DNA content effect on cell size. We suggest a functional hypothesis for why this relationship exists that involves an osmotic effect of DNA. Nucleotides are charged solutes that may decrease the osmotic potential of plant cells and draw in more water, increasing turgor pressure, and perhaps resulting in larger cells.

Discrepancies between regression and independent contrast analyses can reveal important patterns in the evolution of phenotypic traits. In the work presented here, for angiosperms (Table 1), independent contrast analyses had uniformly lower variation explained by genome size for five of the eight traits considered. Strong regression and weak independent contrast results arise when large divergences deep in the phylogeny are highly influential and more recent divergences are much smaller. This result can be an indication of significant trait shifts at higher taxonomic levels with subsequent trait conservatism operating among more closely related species (Ackerly and Donoghue, 1998; Ackerly and Reich, 1999). The most obvious example of this process is the evolutionary divergence between angiosperms and gymnosperms. Not only is this divergence important in shaping genome size variation among extant species, but it also resulted in significant trait shifts coinciding with an apparent reduction in genome size within the angiosperms (Leitch et al., 1998; Beaulieu et al., 2007a, b). The functional and physiological significance of genome size during these important evolutionary events may provide insights into understanding the influence of genome size evolution at higher phenotypic scales.

The strength of correlations between guard cell length and epidermal cell area is remarkably strong and allows for the possibility of genome size inferences of extinct fossilized species. Early work by Thomson (1972), Thomson and Muraszko (1978) and Conway Morris and Harper (1988) demonstrated the strength of this kind of approach by identifying intriguing evolutionary trends in genome size within and among various vertebrate groups (reviewed by Leitch, 2007). The recent study of Organ et al. (2007) has ignited a renewed interest in this area of research by clearly showing that the small genomes of bird species were a pre-existing trait within a lineage of dinosaurs. For plant biology, this same approach can be used to determine the response of genome size to climatic catastrophe (i.e. the KT extinction event). Which species survived? Was there a genome size filter? There is evidence suggesting that species with smaller genome size are more invasive (Bennett et al., 1998) and grow in more extreme environments (Knight and Ackerly, 2002). Therefore, these species may be more likely to survive a climatic catastrophe such as the KT event. We propose a concerted effort to align databases of fossilized leaf impressions with known fossil ages to address this question. We also encourage the continued effort in joining the Plant DNA C-values database (Bennett and Leitch, 2005) and other functional trait databases, such as GlopNet (Wright et al., 2004), SfD (Flynn et al., 2004), wood density (Swenson and Enquist, 2007), The Ecological Flora of California (http://ucjeps.berkeley.edu/efc) and BioFlor (www.ufz.de/bioflor). Through this effort, the patterns will soon be fully described, and processed-based studies can be initiated for traits that are significantly correlated with genome size.

LITERATURE CITED


Beaulieu JM, Leitch IJ, Patel S, Pendharkar A, Knight CA. 2007e. Genome size is a strong predictor of cell size and stomatal density in angiosperms.


