The scaling of genome size and cell size limits maximum rates of photosynthesis with implications for ecological strategies

Adam B. Roddy¹, Guillaume Théroux-Rancourt², Tito Abbo³, Craig R. Brodersen¹, Brook Jensen⁴, Guo-Feng Jiang⁵,⁶, R. Alexander Thompson³, Sara E. Kuebbing⁷, Kevin A. Simonin³

¹ School of Forestry & Environmental Studies, Yale University, New Haven, CT 06511 USA
² Institute of Botany, Universität für Bodenkultur, Vienna, Austria
³ Department of Biology, San Francisco State University, San Francisco, CA, 94132 USA
⁴ Department of Biological Sciences, California State University-Stanislaus, Turlock, CA 95382 USA
⁵ State Key Laboratory of Conservation and Utilization of Subtropical Agrobioreources, Guangxi University, Nanning, Guangxi 530004, China
⁶ Guangxi Key Laboratory of Forest Ecology and Conservation, College of Forestry, Guangxi University, Nanning, Guangxi 530004, China
⁷ Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260 USA

Author for contact:
Email: adam.roddy@yale.edu
Phone: +1.510.224.4432
ORCID: 0000-0002-4423-8729
Abstract

A central challenge in plant ecology is to define the major axes of plant functional variation with direct consequences for fitness. Central to the three main components of plant fitness (growth, survival, and reproduction) is the rate of metabolic conversion of CO₂ into carbon that can be allocated to various structures and functions. Here we (1) argue that a primary constraint on the maximum rate of photosynthesis per unit leaf area is the size and packing density of cells and (2) show that variation in genome size is a strong predictor of cell sizes, packing densities, and the maximum rate of photosynthesis across terrestrial vascular plants. Regardless of the genic content associated with variation in genome size, the simple biophysical constraints of encapsulating the genome define the lower limit of cell size and the upper limit of cell packing densities, as well as the range of possible cell sizes and densities. Genome size, therefore, acts as a first-order constraint on carbon gain and defines the upper limits of allocation to growth, reproduction, and defense. The strong effects of genome size on metabolism, therefore, have broad implications for plant biogeography and for other theories of plant ecology.
Introduction

Quantifying major axes of plant functional variation has given rise to an ever-growing list of traits that impact growth, reproduction, and survival, the three components of individual fitness (Violle et al. 2007). These traits have traditionally been viewed from a reductionist perspective that scales form-function relationships of individual plant organs, e.g. leaves, stems, and roots, to whole organism ecological strategies. As the ultimate source of energy and matter for growth and reproduction, photosynthetic capacity represents a first-order constraint on the emergent properties between whole plant form and function and individual fitness. Here we provide evidence that genome-cellular allometry directly influences interspecific variation in photosynthetic metabolism and provide a mechanistic framework that links genome size and metabolism to other aspects of plant ecology and evolution.

A major limitation on photosynthetic capacity is the ability to deliver resources to, and exporting products from, the sites of metabolic processing (Enquist et al. 1998; West et al. 1999a; Brown et al. 2004). At the level of an individual cell—the fundamental unit of living organisms—rates of resource transport are strongly influenced by cell size because the ratio of cell surface area to cell volume increases exponentially with decreasing cell size. Building on an extensive body of literature we show how variation in the size of the nuclear genome fundamentally limits the sizes and packing densities of cells, and, as a result, the maximum rate of photosynthetic metabolism across vascular plants.

One of the three components of fitness is growth which is ultimately limited by photosynthetic capacity. Relative growth rate (RGR) varies considerably across species and is driven by photosynthetic rate and the resource investment to support photosynthesis as:

$$RGR = A_{mass} \cdot LMR$$

where $A_{mass}$ is the photosynthetic rate per unit leaf biomass and LMR is the leaf mass ratio (the proportion of leaf dry mass to total plant dry mass). $A_{mass}$ is, therefore, frequently
considered a major plant strategy axis (Poorter and Remkes 1990; Poorter et al. 1990; Reich et al. 1992). However, $A_{mass}$ can be decomposed as:

$$A_{mass} = SLA \cdot A_{area}$$

where SLA is the specific leaf area (leaf area per leaf dry mass) and $A_{area}$ is the net carbon assimilation rate per unit canopy leaf area. Because of its direct effect on $A_{mass}$, SLA is often considered a major predictor of interspecific variation in RGR. $A_{area}$, on the other hand, varies orthogonally to SLA (Wright et al. 2004), and, therefore, determines the upper limit of the relationship between $A_{mass}$, SLA, and RGR. Maximum potential $A_{area}$ represents, then, a fundamental limitation on the maximum amount of carbon available for allocation to growth, reproduction, and survival relative to species ecological strategies. The centrality of $A_{area}$ to plant ecological strategy suggests two questions:

- First, what are the fundamental features of plant structure that determine maximum potential $A_{area}$?
- Second, to what extent do these relationships scale up to affect plant ecological strategies and evolutionary dynamics?

Here we present a mechanistic framework to address both of these questions, that is based on a long-recognized but often overlooked relationship: positive scaling between genome size and cell size. Although the relationship between genome size (i.e. nuclear volume) and cell size has long been of interest (Sachs 1893), the mechanisms are still not fully understood (Doyle and Coate 2019). Yet, we show that the allometry between genome size and cell size directly influences rates of photosynthetic metabolism and argue that the scaling of genome size and metabolism affects ecological distributions and evolutionary dynamics. In this way, any factor affecting rates of metabolism is a potential agent of selection on the genome, if only via selection on genome size.

It is now widely recognized that variation in genome size can have significant consequences for organismal structure and function, independent of the genes that define the genotype (Bennett 1971). Positive scaling between genome size and cell size across terrestrial plants has given rise to numerous studies characterizing the many other
phenotypic correlates of genome size independent of variation in genome structure, commonly referred to as “nucleotype” effects (Bennett 1971; Cavalier-Smith 1978; 1982; Bennett and Leitch 2005). Correlates of genome size encompass an incredible diversity of plant phenotypes, including, for example, the sizes of plant structures, rates of cell division, rates of physiological processes, and tolerances and responses to abiotic conditions (Table 1).

Our goal is not to recapitulate the many reviews about the nucleotype-phenotype relationship but, instead, to align these studies more systematically with the field of plant functional biology. We believe that the diverse impacts of genome-cellular allometry on the body plan of terrestrial vascular plants strongly influences the coordination between plant functional traits and, ultimately, whole organism form-function relationships. Here we summarize previous research, perform new analyses of existing data, and present new data to show how genome size may, through its impacts on cell size and tissue structure, determine the biophysical limits of plant metabolic rates, and, therefore, influence other aspects of ecology and evolution. That genome size may be a key functional trait is not a new idea (Grime 1998). Yet, despite numerous reports of the phenotypic and ecological correlates of genome size (Table 1), it has not been fully identified and integrated into the functional trait literature. Our goal, therefore, is to more directly show how genome size influences plant traits that impact maximum rates of photosynthetic metabolism.

Metabolism is central to all three aspects of plant fitness, providing the carbon necessary for allocation to growth, reproduction, and survival. As such, genome size may not itself be a functional trait but instead may define the limits of variation in numerous other functional traits.

**Genome-cellular allometry limits rates of resource transport and metabolism**

**Allometry of genome size and cell size**

The role of the genome in limiting cell size has been postulated since at least the late 1800s (Sachs 1893) and was critical in shaping early modern views of the evolution of plant
vascular systems (Bailey and Tupper 1918). At a minimum, a cell must contain its genome, and there is a strong relationship between the volumes of meristematic cells and genome size (Šimová and Herben 2012). Cellular expansion from this meristematic minimum size is cell type-specific (Doyle and Coate 2019). Within a cell type, size can be influenced by various environmental and developmental factors (Melaragno et al. 1993). Despite this substantial growth in cell volume during development, there remains a significant effect of genome size on cell size, particularly for stomatal guard cells (Beaulieu et al. 2008; Knight and Beaulieu 2008; Lomax et al. 2013; Simonin and Roddy 2018). For example, stomatal guard cell size and density, which regulate the fluxes of water and CO₂ between the biosphere and atmosphere, vary within species depending on light, water availability, and atmospheric CO₂ concentration (Hetherington and Woodward 2003; Franks and Beerling 2009). Furthermore, in the vascular transport network, the sizes of xylem conduits and their density in the leaf are also affected by variation in genome size (Maherali et al. 2009; Hao et al. 2013; De Baerdemaeker et al. 2018; Simonin and Roddy 2018). Yet why genome size and final cell size are correlated within a cell type remains unclear (Doyle and Coate 2019).

We tested whether smaller genomes allow not only for smaller initial and final cell sizes but also for a greater range in final cell size using published data for terrestrial C3 plants. We used data for stomatal guard cells because they are the most commonly measured cell sizes in plants and because their sizes and abundance determine the leaf surface conductance to CO₂ and water vapor and, therefore, directly control rates of resource transport for use in photosynthetic metabolism. Sizes of guard cells for angiosperms (Beaulieu et al. 2008), gymnosperms, and ferns were compiled previously by Simonin and Roddy (2018), and here we include data for mosses and hornworts from Feild et al. (2015) and Renzaglia et al. (2017). We assumed that stomatal guard cells are shaped as capsules, which are composed of a central cylinder with hemispherical ends, such that cell volume could be estimated from cell length as:

\[ V = \pi \cdot r^2 \cdot (\frac{4}{3} r + a) \]
where \( r \) is the radius of the cylinder and \( a \) is the height of the cylinder. We assumed that \( a \) is equal to \( 2r \). The dumbbell-shaped guard cells common among grasses would likely violate these assumptions about cell shape and so we have not included in this analysis data for lineages known to have dumbbell-shaped guard cells. Data for meristematic cell volume and genome size were taken from Simova and Herben (2012). We used standard major axis regression (R package `smatr`) to fit the mean response and quantile regression (R package `rq`) to test whether there was greater variation in cell volume among taxa with smaller genomes (i.e. heteroskedasticity).

For a given genome size, mature guard cells were at least three orders of magnitude more voluminous than meristematic cells (Figure 1). Genome size was a strong and significant predictor of meristematic cell volume (\( \log(\text{volume}) = 1.24 \log(\text{genome size}) - 0.23; R^2 = 0.98, P < 0.001; \) Šimová and Herben 2012). Though it explained less of the variation, genome size was a significant predictor of final guard cell volume (\( \log(\text{cell volume}) = 0.80 \log(\text{genome size}) + 3.43; R^2 = 0.09, P < 0.001 \)). Quantile regression revealed that the slope through the 10% quantile was steeper (0.58) than the slope through the 90% quantile (-0.07), reflecting both the greater variation in final cell volume among species and the smaller minimum possible cell volume that can be obtained with smaller genomes (i.e. heteroskedasticity in cell volume with variation in genome size). Thus, smaller genomes allow for smaller minimum cell size with little or no effect on maximum cell size, together enabling greater potential variation in final cell size. The greater variation among species with smaller genomes implies that smaller genomes allow for greater plasticity in cell sizes and cell packing densities which directly influence maximum rates of leaf surface conductance to \( \text{CO}_2 \) and water and ultimately photosynthetic metabolism per unit leaf surface area (Simonin and Roddy 2018). Further, the greater diversity of cell sizes observed in plants with small genomes is supporting evidence that the correlation between genome size and cell size is simply the result of occupying available space within the cell. A small genome can be housed in either a small or a large cell, but a large genome requires a cell large enough to contain its volume.

The greater variation in cell volume allowed by smaller genomes across terrestrial C3 plants (Figure 1) further suggests that smaller genomes allow for greater variation in cell
packing densities. For guard cell lengths, stomatal densities, and vein densities, smaller
genomes allowed for greater variation in traits across ferns, gymnosperms and
angiosperms (Simonin and Roddy 2018). Species with smaller genomes in these datasets
are predominantly angiosperms, and these analyses compared distantly related species.
We further tested for greater variation in cell sizes and packing densities with smaller
genomes among closely related species using taxa in Rhododendron sect. Schistanthe Schltr.
(= sect. Vireya Blume). This monophyletic clade has a stepwise phylogeographic history,
having radiated eastward from the Malay Peninsula and reached New Guinea within the
last 15 Ma (Goetsch et al. 2011). We sampled leaves from 19 taxa growing under common
garden conditions at the Rhododendron Species Foundation Botanical Garden in Federal
Way, WA, USA. Genome sizes were measured following standard protocols (Dolezel et al.
2007) at the Benaroya Research Institute in Seattle, WA, USA. For measurements of
stomatal size and density, epidermal impressions were made on fresh leaves using dental
putty (Coltene Whaledent President Light Body), transferred using clear nail polish,
mounted in water, and imaged using a light microscope. Measurements of leaf vein density
were made on leaf sections cleared by soaking in 4% NaOH, 3% sodium hypochlorite,
stained with 1% Safranin O, counterstained with 1% Fast Green, mounted in ethanol, and
imaged with a light microscope. Stomatal traits were averaged across ten images per taxon,
and leaf vein density was averaged across five images per taxon. Quantile regression (R
package ‘rq’) through the 10th and 90th percentile of the species means were used to
quantify the variation in traits associated with variation in genome size. Consistent with
previous results across vascular plants (Simonin and Roddy 2018), among the closely
related Rhododendron sect. Schistanthe, there was greater variation in the sizes and packing
densities of veins and stomata among species with smaller genomes (Figure 2).
Furthermore, the steeper quantile slopes for the lower limit of cell size and the upper limit
of packing densities are consistent with the hypothesized constraints imposed by genome
size variation.

Genome size limits maximum photosynthetic metabolism

Previous work has hypothesized that genome size would be linked to maximum
photo synthetic rate but found little support (Knight et al. 2005; Beaulieu et al. 2007). We
hypothesize that one major reason for not finding support is that these previous studies attempted to predict variation in $A_{\text{mass}}$, which accounts for the construction costs of leaves, rather than $A_{\text{area}}$, which is the maximum metabolic rate regardless of the construction costs. As we described above, $A_{\text{area}}$ would define the maximum amount of carbon assimilated, but how the plant allocates the total assimilated carbon—to growth, reproduction, defense, more durable leaves, etc. (Bazzaz et al. 1987)—would reflect the numerous factors that influence plant form and other aspects of plant function. Thus, $A_{\text{area}}$, which is orthogonal to SLA and $A_{\text{mass}}$ (Wright et al. 2004), is predicted to be constrained by cell and genome sizes.

Consistent with this prediction, recent work has shown that genome size is a strong predictor of the sizes and densities of stomatal guard cells and leaf veins across vascular plants (Simonin and Roddy 2018), and we predicted, therefore, that genome size would, via its effects on the sizes and packing densities of cells, predict $A_{\text{area}}$.

Data for area-based maximum photosynthetic rate were compiled from the primary literature (Supplemental Table 1) and merged with the Kew Plant DNA C-Values Database (Bennett and Leitch). This dataset included 210 species, of which 138 were angiosperms, 46 were gymnosperms, and 26 were ferns. We tested for a relationship between genome size and $A_{\text{area}}$ both for the mean response (standard major axis regression, SMA; R package ‘smatr’) and for the upper 90% quantile of the data using quantile regression (R package ‘rq’). To test whether the SMA regression was driven by phylogenetic history, we used a phylogenetically corrected version (R package ‘phytools’), which provides a P-value only for the test against a specific slope, for which we used the non-phylogenetically corrected SMA slope. For this analysis, we generated a dated supertree using the methods of Simonin and Roddy (2018), which included generating a family-level supertree using the R package ‘bbranching’ (Chamberlain 2016) and applying node ages from Magallón et al (2015). Of the 128 nodes, 43 with node names had ages. These ages were applied and intermediate branch lengths smoothed using ‘bladj’ in Phylocom (Webb et al. 2008). We randomly resolved polytomies by adding 5 million years to each branch and subtracting an equivalent amount from the subtending branches to keep the tree ultrametric.

Species with smaller genomes had higher maximum photosynthetic rates across and within major plant clades (Figure 3). Standard major axis regression on log-transformed data
revealed a significant negative relationship between genome size and $A_{area}$ across all species ($\log(A_{area}) = -0.42 \log(\text{genome size}) + 1.29; R^2 = 0.23, P < 0.001$). The phylogenetic major axis regression was significantly steeper than the non-phylogenetic major axis regression ($\log(A_{area}) = -0.75 \log(\text{genome size}) + 1.76; \text{slope test } P < 0.001$), suggesting that the evolution of genome size across land plants is strongly associated with changes in $A_{area}$ across vascular plants. A similarly negative relationship was apparent among the upper (90%) quantile, which showed that $A_{area}$ was lower among species with larger genomes. Despite the large difference in genome size between angiosperms and gymnosperms, the 90% quantile regression slopes were nearly identical for the entire dataset (-0.19) as for the angiosperms alone (-0.18).

The scaling relationship between $A_{area}$ and genome size follows naturally from the relationships between genome size and the sizes and densities of veins and stomata. However, veins and stomata are not the only cells responsible for driving variation in photosynthetic rates. While the maximum rate of CO$_2$ diffusion into the leaf is defined by the sizes and densities of stomata (Franks and Beerling 2009), once inside the leaf, CO$_2$ must diffuse through the leaf intercellular airspace and into the chloroplasts lining the interior surfaces of mesophyll cells. The diffusive surfaces of the leaf mesophyll are the interface between liquid-phase and vapor-phase resource movement, where liquid water delivered by the veins evaporates and where CO$_2$ must diffuse into the liquid phase. Thus, the three-dimensional structure and organization of the mesophyll is predicted to be a prime target for selection on photosynthetic metabolism (Tholen et al. 2012; Ren et al. 2019) and to be critical to leaf photosynthetic function (Earles et al. 2019), and the limited evidence on Arabidopsis thaliana mutants with altered cell diameter and porosities point in that direction (Lehmeier et al. 2017). We predict, therefore, that the scaling relationships between genome size and cell size that coordinate veins and stomata (Simonin and Roddy 2018) extend also to the photosynthesizing cell size, structure, and organization within the leaf mesophyll. The significant negative scaling relationship between $A_{area}$ and genome size would emerge from the effects of genome size on the sizes and packing densities in each of these tissues.
Genome size limits the rate of metabolic up- or down-regulation

Although maximum potential rate of leaf gas exchange is an important parameter determining a species’ physiological capacity, the actual rate of leaf gas exchange at any given moment is often substantially lower, depending on a variety of physiological and environmental factors (e.g. light level, atmospheric humidity, leaf temperature, plant water status). Changes in sun angle, shading by passing clouds, and self-shading by fluttering leaves all drive changes in incoming solar radiation, and these rapid dynamics have influenced the evolution of photosynthetic biochemistry (Pearcy 1990). Under naturally varying conditions, leaf gas exchange fluctuates dramatically and rarely reaches its maximum rate, with greater variation occurring at the top of the plant canopy. How frequently a leaf can reach its maximum gas exchange rate and how well optimize its physiological processes to environmental conditions depend on how rapidly the leaf can respond to dynamic, fluctuating conditions.

An emerging consensus is forming that smaller stomata respond more rapidly to fluctuating conditions than larger stomata, allowing leaves with smaller stomata to more closely tune their physiological rates with environmental conditions (Drake et al. 2013; Lawson and Blatt 2014; Lawson and Viallet-Chabrand 2019). Leaf physiological processes change at different rates, with changes in stomatal conductance occurring an order of magnitude more slowly than changes in photosynthesis (McAusland et al. 2016). This difference in response times between physiological processes (e.g. photosynthetic assimilation rate and stomatal conductance) can reduce water use efficiency when stomata are closing and reduce photosynthetic efficiency when stomata are opening (Lawson and Viallet-Chabrand 2019), limiting total photosynthesis by up to 20% (Lawson and Blatt 2014). If stomatal response times are directly limited by the size of stomata then genome-cellular allometry fundamentally limits not only the maximum rate of metabolism but also how quickly metabolism can respond to environmental change. Of the species for which stomatal response times were measured by McAusland et al. (2016) and Drake et al. (2013), twelve were included in the Kew Plant DNA C-Values database. Consistent with previous reports, there was a positive correlation between genome size and guard cell length (Figure 4a), and stomatal response rate exhibited a triangular relationship with
genome size such that smaller genomes exhibited both higher maximum stomatal response rates but also a greater variation in stomatal response rate. While the available data on stomatal response rates measured using standard protocols are limited, these preliminary results suggest that genome size indirectly limits the maximum rate of stomatal opening and closing via its effects on the sizes and densities of stomata.

Implications of genome size-metabolism scaling on plant biogeography

Polyplody thought to increase habitat breadth

Variation in genome size and structure associated with polyploidization has long been considered to be an important driver of plant evolution and to be associated with environmental tolerances, habitat breadth, trait variation, and interspecific interactions (Stebbins 1940; Otto and Whitton 2000; Soltis et al. 2003; Soltis et al. 2014; Michael S Barker et al. 2016; 2016), all of which influence species biogeography. Describing the types of polyploids and how they are generated is outside the scope of the current manuscript and have been thoroughly reviewed elsewhere (e.g. Stebbins Jr 1947; Soltis et al. 2015). Rather, we focus our discussion here on how and why ploidy—via its relationship with genome size—may or may not correlate with species distributions and habitat breadth.

Polyploids have been hypothesized to be better adapted to extreme habitats, to have greater hardiness, and to have greater ecological adaptability [reviewed by (Stebbins 1985: @Brochmann:2004aa). The possible mechanisms for these effects can be roughly grouped into two categories: one involving the genetic and genic content of the polyploid genome and one involving the nucleotypic effects of ploidy and genome size. Because polyploid genomes have additional genome copies, they have higher absolute genic contents, would enable neo-functionalization of duplicated genes, and typically have higher heterozygosity, all of which can promote higher tolerances of environmental conditions. The nucleotypic effects of ploidal variation, though long recognized (Stebbins 1940), are often confounded with nucleotypic effects of genome size variation.
While ploidy and genome size are commonly assumed to be synonymous, at broad phylogenetic scales there is frequently no relationship between genome size and ploidy (Leitch and Bennett 2004), reflecting the complex history of both ancient and contemporary whole genome duplications, particularly among the angiosperms (Jiao et al. 2011; Clark and Donoghue 2018; Landis et al. 2018; Ren et al. 2018). Angiosperm genomes readily rediploidize after polyploidization such that genome size and ploidy are positively correlated only for narrowly defined phylogenetic groups (i.e. within genera and families, Figure 5; Leitch and Bennett 2004; Dodsworth et al. 2016). Ferns, by contrast, the other major group of vascular plants that readily undergoes polyploidization events seem not to rediploidize as readily (Clark et al. 2016), as evidence by the positive correlation between genome size and ploidy (Figure 5). If leaf and plant structure and function influence ecological tolerances and habitat breadth (i.e. if plant structure-function is adaptive), then the nucleotypic effects of genome size are predicted to influence environmental tolerances.

Smaller genomes enable greater phenotypic plasticity

One long-standing hypothesis is that higher ploidy is related to wider habitat breadth because polyploids can tolerate greater ecological stress. Higher ploidy is associated with greater heterozygosity (i.e. greater genetic diversity) and, frequently, higher genic content due to multiple genome copies, both of which are thought to promote plasticity and enable polyploids to withstand a greater range of environmental conditions than diploids. However, several studies testing this hypothesis have not observed polyploids to have greater habitat breadth (e.g. Stebbins 1985; Martin and Husband 2009; Johnson et al. 2014). Furthermore, these tests frequently find that diploids exhibit greater habitat breadth than polyploids (Petit and Thompson 1999; Hijmans et al. 2007; Brittingham et al. 2018). One reason is that traits of polyploids are not necessarily more variable than in diploids (Stebbins 1985; Wei et al. 2018).

We predict that one reason ploidy is not commonly found to correlate with ecological breadth is because genome size—rather than ploidy per se—drives variation in the absolute range of potential cell sizes and, by extension, phenotypic plasticity in rates of resource transport and metabolism. Thus, the phylogenetic scale-dependence of the relationship
between genome size and ploidy could lead to confounding patterns depending on the phylogenetic scale at which comparisons are made. For example, in the analysis of Rice et al. (2019), ploidy was determined relative to other closely related species, such that within genera or families ploidy and genome size are positively correlated, suggesting that the bias towards higher abundances of polyploids at higher latitudes may reflect nucelotypic effects of genome size on cell size and metabolism. The complex, fluctuating process of polyploidization and rediploidization, which can winnow the genome nonrandomly (Wendel 2015), would promote the proliferation of beneficial elements associated with genome duplications (e.g. more gene copies that can neofunctionalize) while reducing the size of the genome needed to maintain high rates of development and metabolism (Table 1).

We posit here that the nucleotypic effects of genome size, regardless of ploidy, may influence environmental tolerances. Because smaller genomes allow for greater variation in cell size (Figure 1), species with smaller genomes may be better able to fine tune their tissue structure to environmental conditions. This flexibility would allow species with smaller genomes to better optimize their metabolic rates in order to occupy a wider range of environmental conditions. Combined with the effects of genome size on rates of cell division (Van't Hof and Sparrow 1963; Hof 1965; Šimová and Herben 2012), the greater plasticity in cell size and higher rates of maximum photosynthesis available to species with small genomes may enable them to better colonize new habitats.

Greater variation in genome size in less productive habitats

In habitats that can support high rates of productivity and primary metabolism, species with small genomes are expected to predominate. At a broad scale, this prediction holds because angiosperms, which have, on average, smaller genomes than other vascular plants are dominant in most ecosystems, particularly those characterized by high productivity. However, high rates of metabolism, even among the angiosperms, are not always favorable. Two such habitats are those characterized by extreme water limitation and nutrient limitation, such as deserts and epiphytic habitats, and by extreme cold, such as high latitudes. Higher incidences of polyploids have been commonly reported in higher latitudes.
and among arctic floras (Brochmann et al. 2004; Rice et al. 2019), but arid habitats have received less attention.

Arid and epiphytic habitats are characterized by low productivity and may support species with large genomes. In these habitats, high rates of metabolism are not always favored, which may relax selection for small genomes. One strategy common in arid and epiphytic habitats is succulence, which is often associated with Crassulacean acid metabolism (CAM) photosynthesis. The CAM syndrome limits water loss by sequestering CO₂ uptake and water loss to nighttime which corresponds with periods of concomitant water loss to periods of high humidity and little energy available to drive water loss (i.e. at night). As a result, this strategy is characterized by low rates of resource transport. If metabolism is one axis of selection on genome size, then we would predict that CAM species may, therefore, have larger genomes than species employing C₃ metabolism. We tested this hypothesis using the taxonomic distributions of CAM photosynthesis from Smith and Winter (1996) and genome size data from the Kew Plant DNA C-Values Database (citation). For C₃, we used the broad distribution of angiosperms reported in Simonin and Roddy (2018), which are representative of extant angiosperm diversity. We scored as CAM the narrowest taxonomic level in the Kew DNA C-Values Database that was listed as containing CAM by Smith and Winter (1996). For example, if a genus were listed as containing any CAM species, all species in the genus were assumed to exhibit CAM photosynthesis. This strategy was biased against observing differences in genome size between C₃ and CAM species because it necessarily grouped some C₃ species as CAM. Nonetheless, genome sizes were significantly larger among CAM species than among C₃ species even after log-transforming the data to improve normality (t-test: t = 9.72, df = 272.6, P < 0.0001; Figure 6), consistent with the prediction that large genomes may evolve when selection for high rates of metabolism is weak.

Arid habitats are not exclusively composed of species with large genomes. Rather, they may harbor a diversity of strategies associated with divergent niches. While many C₃ species in deserts function during a relatively narrow window in the year, either because they are annuals or because physiological activity in perennial shrubs is restricted to a narrow range of environmental conditions, when water is available. Thus, because their carbon
gain is limited to such a short time period, they may have small genomes and cells that
enable high rates of metabolism. Desert shrubs have the highest rates of stem hydraulic
conductance measured in C3 plants (Mencuccini 2003), and even among species from
humid tropical forests, dry forest species have higher hydraulic conductance than wet
forest species (Brener-Arguedas et al. 2013). Thus, less productive habitats may allow for
multiple strategies that encompass a broader range of metabolic rates and, allowing for
greater variation in genome size at the community level than might be observed in
conditions more favorable to high rates of metabolism and productivity.

Smaller genomes increase the probability of invasiveness

The multifaceted effects of genome size on plant structure, function, and ecology (Table 1)
is particularly relevant to the study of invasive species. Identifying the traits that allow an
introduced species to establish, naturalize, and invade into a new environment is a central
aim of invasion biology (Simberloff 2011), with broader implications for plant
biogeographic patterns. Here we distinguish between nonnative species—those that survive
and reproduce in their introduced range—and nonnative invasive species—those that can
disperse and establish far from their original source of introduction (Richardson et al.
2011). This distinction is important because prior studies on the traits of ‘invaders’ focus
on these different subsets of species, which have slightly different, but overlapping, sets of
traits that determine whether they can survive and reproduce versus invade (Kleunen et al.
2015).

Early theory on the distinguishing traits of invasive plants postulated that “ideal weeds”
should grow rapidly, produce seed continuously and in high number throughout the
growing season, be tolerant to a wide range of environmental conditions, exhibit high trait
plasticity, and be able to reproduce vegetatively from fragments (Baker 1974). On average,
these predictions have been upheld, with nonnative invasive plants tending to exhibit traits
consistent with high fitness (e.g., number of flowers, fruits, or seed or germination rates),
high relative growth rates, high dispersal abilities (smaller seeds), and more efficient
carbon-capture strategies (e.g. high specific leaf area), relative to co-occurring native
species (Leishman et al. 2007; Kleunen et al. 2010; Ordonez et al. 2010; Kuester et al. 2014)
or naturalized but not invasive nonnative species (Rejmánek and Richardson 1996; Gallagher et al. 2014). Combined, these traits confer a growth advantage, such that plants with small seeds can disperse further distances, have shorter generation times, and higher relative growth rates, owing to the greater rates of cell division and higher metabolic rates provided by smaller genomes (Pandit et al. 2014; Suda et al. 2015). Indeed, even within species, populations with smaller genomes are more likely to successfully invade new habitats (Pysek et al. 2018).

Because many of the traits linked with invasiveness can be influenced by both ploidy and genome size, both have been implicated as underlying features driving invasion (Pandit et al. 2014; Suda et al. 2015). Because polyploids are thought to be better able to tolerate environmental fluctuations and to be better able to adapt to new environments, polyploids tend to be overrepresented among nonnative invasives compared to native angiosperms (Rejmánek and Richardson 1996; Prentis et al. 2008; Beest et al. 2011; Pandit et al. 2014). Similarly, nonnative invasive species tend to have smaller genomes than non-invasive plants (both native and non-native), which is thought to be due to the diverse effects of genome size on metabolism, rates of development and growth, and seed size (Rejmánek and Richardson 1996; Bennett et al. 1998; Kubešová et al. 2010; Pandit et al. 2014).

However, the complex, scale-dependent relationship between ploidy and genome size discussed previously complicate a clear understanding of the effects of ploidy versus genome size on invasiveness (Rejmánek and Richardson 1996; Pandit et al. 2014). Because angiosperms, which predominate among nonnative invasives, readily rediploidize and downsize their genomes subsequent to whole genome duplications (Leitch and Bennett 2004), assessing the relative effects of ploidy versus genome size on invasiveness can be difficult. For example, the likelihood of being invasive increases with chromosome number and ploidy but decreases with genome size (Rejmánek and Richardson 1996; Pandit et al. 2014). The multiple paths to polyploidization and the selective retention of only certain parts of the genome during subsequent genome downsizing (Wendel 2015) could explain how both higher ploidy and smaller genomes are correlated with invasiveness.
Implications for genome size variation on the Metabolic Theory of Ecology

As the major source of energy and matter for the biosphere, photosynthetic metabolism represents a first-order control over ecological processes globally. This fundamental link between metabolic and ecological processes has driven the development of the Metabolic Theory of Ecology (MTE) that provides a mechanistic framework for predicting variation in organismal life history attributes, population dynamics, and larger scale ecosystem processes from organismal-level traits related to supplying resources for metabolism (West et al. 1997; Enquist et al. 1998; West et al. 1999a; West et al. 1999b; West et al. 2002; Price et al. 2010). While appealing and seemingly endowed with incredible explanatory power, a number of criticisms of the theory and its assumptions have been consistently raised (Kozłowski and Konarzewski 2004; Kozłowski and Konarzewski 2005; Price et al. 2012). One primary assumption is that the sizes of terminal units in vascular networks (e.g. capillaries in circulatory systems or terminal veins in plant leaves) are invariant. The problems with this assumption have been thoroughly detailed for animal circulatory systems with the allometry of genome size and cell size emerging as a critical factor influencing how body size scales with metabolism (Kozłowski et al. 2003). Furthermore, the allometry of genome size and cell size (Figure 1) and the effects of genome size on maximum metabolic rate (Figure 3) presented here suggest that this assumption is violated in plants, as well. Modifications to the original model that relax some of its assumptions have improved model predictions for plants, particularly by allowing for variation in the packing of xylem conduits (Savage et al. 2010). However, the nucleotypic effects of genome size have yet to be incorporated, although they may further improve models and help to clarify the constraints and major innovations driving botanical form, function, and diversity.
Table 1. Brief summary of traits shown previously to correlate with genome size.

<table>
<thead>
<tr>
<th>Sizes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pollen volume</td>
<td>Bennett 1972</td>
</tr>
<tr>
<td>Cell mass</td>
<td>Martin 1966</td>
</tr>
<tr>
<td>Epidermal cell size</td>
<td>Beaulieu et al. 2008; Knight and Beaulieu 2008</td>
</tr>
<tr>
<td>Nuclear volume</td>
<td>Van’t Hof and Sparrow 1963; Baetcke et al. 1967</td>
</tr>
<tr>
<td>Nuclear dry mass</td>
<td>Bennett et al. 1983; White and Rees 1987</td>
</tr>
<tr>
<td>Seed mass</td>
<td>Grotkopp et al. 2004; Knight and Beaulieu 2008</td>
</tr>
<tr>
<td>Xylem vessel diameter</td>
<td>Maherali et al. 2009; Hao et al. 2013; De Baerdemaeker et al. 2018</td>
</tr>
<tr>
<td>Rates</td>
<td></td>
</tr>
<tr>
<td>Cell division rate, meiosis, mitosis</td>
<td>Van’t Hof and Sparrow 1963; Hof 1965; Bennett 1971</td>
</tr>
<tr>
<td>Minimum generation time</td>
<td>Bennett 1972</td>
</tr>
<tr>
<td>Leaf expansion rate</td>
<td>Grime et al. 1985</td>
</tr>
<tr>
<td>Phenology</td>
<td>Grime and Mowforth 1982</td>
</tr>
<tr>
<td>Frost tolerance</td>
<td>MacGillivray and Grime 1995</td>
</tr>
</tbody>
</table>
Figure 1. Genome size determines the minimum size of cells, and smaller genomes enable greater variation in final cell size. Data for meristematic cells (blue triangles) were taken from Šímová and Herben (2012), and the solid black line is the standard major axis regression through these points. Data for mature stomatal guard cells (circles and squares) for ferns (dark green), gymnosperms (pink), and angiosperms (light blue) were taken from Simonin and Roddy (2018), and data for mosses and hornwarts (light green) were taken from Field et al. (2015) and Renzaglia et al. (2017). The two regression lines through mature guard cell data represent quantile regressions through the 10th (lower) and 90th (upper) quantiles of the data.

Figure 2. Variation in the sizes and packing densities of stomatal guard cells and leaf veins with variation in genome size among Rhododendron sect. Schistanthe species. The dashed lines represent regressions through the 90th (upper) and 10th (lower) quantiles. Genome size limits the lower limit of cell size and the upper limit of cell packing densities, and there is greater variation in anatomical traits among species with smaller genomes.

Figure 3. Genome size limits the maximum rate of photosynthesis across C3 terrestrial plants. (a) Untransformed relationship and (b) log-transformed relationship. Dashed black lines are regressions through the upper 90th quantile. The solid black line is the standard major axis regression ($R^2 = 0.23, P < 0.0001$), and the solid red line is the phylogenetically corrected RMA regression.

Figure 4. Genome size limits the maximum rate of stomatal response (i.e. how fast stomata can open or close). Data taken from McAusland et al. (2016) and Kew Plant DNA C-values Database.

Figure 5. Relationship between genome size and ploidy for angiosperms. Each line represents the SMA regression within a genus. At narrow taxonomic scales, ploidy and genome size are correlated, but at broad taxonomic scales (i.e. among all angiosperms), there is no relationship between genome size and ploidy due to rediploidization.
Figure 6. Distributions of genome size for C3 and CAM species show CAM lineages have significantly larger genomes than C3 lineages. Lineages identified as CAM likely include many C3 species; see text for details on identification of photosynthetic pathways.
References


Beaulieu JM, IJ Leitch, S Patel, A Pendharker, CA Knight 2008 Genome size is a strong predictor of cell size and stomatal density in angiosperms. New Phytologist 179:975–986.


http://data.kew.org/cvalues/


Chamberlain S 2016 Branching: Fetch 'phylogenies' from many sources.


Mela ragno JE, B Mehrotra, AW Coleman 1993 Relationship between endopolyploidy and cell size in epidermal tissue of arabidopsis. The Plant Cell 5:1661–1668.


Savage VM, LP Bentley, BJ Enquist, JS Sperry, DD Smith, PB Reich, EI von Allmen 2010 Hydraulic trade-offs and space filling enable better predictions of vascular structure and function in plants. Proceedings of the National Academy of Sciences 107:22722–22727.


Šímová I, T Herben 2012 Geometrical constraints in the scaling relationships between genome size, cell size and cell cycle length in herbaceous plants. Proceedings of the Royal Society B: Biological Sciences 279:867-875.


The diagram shows a scatter plot with data points representing different plant species, categorized by their cell volume (µm³) and 2C genome size (pg). The x-axis represents the 2C genome size, while the y-axis represents the cell volume. Different categories of plants are distinguished by different markers on the graph:

- Meristematic cells: Blue triangles (△)
- Mosses, hornworts: Green squares (□)
- Ferns: Green circles (○)
- Gymnosperms: Red circles (○)
- Angiosperms: Blue circles (○)

The data points are distributed across a range of genome sizes and cell volumes, with a trend line indicating a general correlation between the two variables. The figure suggests that larger genome sizes tend to correlate with larger cell volumes.
Guard cell length (µm)

$R^2 = 0.36$

Stomatal response rate (mmol m$^{-2}$ s$^{-2}$)

2C Genome size (pg)