New Phytologist

Research

Phylogenetic trends and environmental correlates of nuclear genome size variation in *Helianthus* sunflowers

Fan Qiu[1], Eric J. Baack[2], Kenneth D. Whitney[3], Dan G. Bock[4], Hannah M. Tetreault[1], Loren H. Rieseberg[4] and Mark C. Ungerer[1]

[1]Division of Biology, Kansas State University, Manhattan, KS 66506, USA; [2]Department of Biology, Luther College, Decorah, IA 52101, USA; [3]Department of Biology, University of New Mexico, Albuquerque, NM 87131, USA; [4]Department of Botany, University of British Columbia, Vancouver, British Columbia, Canada

Author for correspondence:
Fan Qiu
Tel: +1 7855325574
Email: fqiu@ksu.edu

Received: *3 July 2018*
Accepted: *29 August 2018*

New Phytologist (2018)
doi: 10.1111/nph.15465

Key words: genome size (GS) evolution, growing season, *Helianthus*, phylogenetic comparative analysis, temperature seasonality.

Summary

- Flowering plants serve as a powerful model for studying the evolution of nuclear genome size (GS) given the tremendous GS variation that exists both within and across angiosperm lineages.
- *Helianthus* sunflowers consist of *c*. 50 species native to North America that occupy diverse habitats and vary in ploidy level. In the current study, we generated a comprehensive GS database for 49 *Helianthus* species using flow cytometric approaches. We examined variability across the genus and present a comparative phylogenetic analysis of GS evolution in diploid *Helianthus* species.
- Results demonstrated that different clades of diploid *Helianthus* species showed evolutionary patterns of GS contraction, expansion and relative stasis, with annual diploid species evolving smaller GS with the highest rate of evolution. Phylogenetic comparative analyses of diploids revealed significant negative associations of GS with temperature seasonality and cell production rate, indicating that the evolution of larger GS in *Helianthus* diploids may be more permissible in habitats with longer growing seasons where selection for more rapid growth may be relaxed.
- The *Helianthus* GS database presented here and corresponding analyses of environmental and phenotypic correlates will facilitate ongoing and future research on the ultimate drivers of GS evolution in this well-studied North American plant genus.

Introduction

Nuclear genome size (GS) varies dramatically in angiosperms, spanning greater than 2000-fold across flowering plant species (Leitch & Leitch, 2013; Pellicer *et al.*, 2018). The distribution of angiosperm GS variation is markedly non-normal, however, being strongly right-skewed and with only a small number of lineages represented by extremely large GSs (Kelly & Leitch, 2011; Leitch & Leitch, 2013; Kelly *et al.*, 2015). Considerable variation nonetheless can be found both within and across major flowering plant groups, with reports of 63-fold variation found within genera (Grover *et al.*, 2008) and reports of significant GS variation found even within species (Diez *et al.*, 2013; Li *et al.*, 2017; but see Smarda & Bures, 2010). Flowering plants thus represent a powerful system for studying both the proximate mechanisms that drive GS changes and the ultimate forces that may promote or constrain these changes.

GS can evolve bidirectionally. Genome expansion in plants is attributable most notably to polyploidization events (Wood *et al.*, 2009) and amplification of transposable elements (TEs), especially long terminal repeat (LTR) retrotransposons (Kumar & Bennetzen, 1999; Hawkins *et al.*, 2008). These elements typically account for the majority of nuclear DNA in plant species with large genomes and can undergo episodic bursts of activity leading to dramatic increases in GS over both short- and long-term evolutionary time scales (Park *et al.*, 2012; Estep *et al.*, 2013; Bennetzen & Wang, 2014). By contrast, evolution toward smaller GS is facilitated by recombinational processes such as intrastrand homologous recombination and illegitimate recombination, and mechanisms of deletion-biased DNA double strand break repair (Devos *et al.*, 2002; Hawkins *et al.*, 2009; Schubert & Vu, 2016).

Comparative approaches have increasingly been used to investigate both phenotypic and environmental correlates of GS variability. Some of the strongest phenotypic relationships include positive associations of GS with seed mass (Bennett, 1972; Thompson, 1990; Knight *et al.*, 2005), cell size (Knight & Beaulieu, 2008; Vesely *et al.*, 2012) and cell cycle time (Bennett, 1972; Ivanov, 1978; Francis *et al.*, 2008). The strong positive relationship between GS and phenotypes at the cellular level has spurred investigations at higher phenotypic scales (Knight & Beaulieu, 2008) and whether GS variation ultimately is associated with plant life history strategies (Bennett, 1987; Knight *et al.*, 2005; Greilhuber & Leitch, 2013). Accordingly,

correlations have been detected between GS and phenotypes such as relative growth rate, photosynthetic rate and leaf anatomical traits (Knight *et al.*, 2005; Beaulieu *et al.*, 2007a; Gallagher *et al.*, 2011; Kang *et al.*, 2014). Typically, larger genomes are associated with slower growth rates (thought to reflect a cost of the synthesis and maintenance of extra DNA), lower photosynthetic rates and lower specific leaf area (SLA). Unlike studies examining correlations with cellular-level traits, however, studies at higher phenotypic scales have demonstrated inconsistent and even opposing results across different plant groups (Knight *et al.*, 2005). Nonetheless, and consistent with the hypothesis that life history strategies may exert strong selection on GS, species with larger genomes are less likely to be categorized as weeds or invasive species (Kubesova *et al.*, 2010; Pandit *et al.*, 2014; Suda *et al.*, 2015) and more likely to have a perennial than annual life cycle (Bennett, 1972).

That life history strategies may exert selection on GS is further suggested by correlations of GS with environmental parameters such as altitude, latitude, temperature and precipitation (Knight & Ackerly, 2002; Grotkopp *et al.*, 2004; Knight *et al.*, 2005; Kang *et al.*, 2014; Du *et al.*, 2017; Bilinski *et al.*, 2018; Lyu *et al.*, 2018). Some of the most intriguing environmental correlates indicate that species with larger genomes tend to be excluded from more climatically extreme and/or temporally variable environments where growing seasons are shorter (Knight & Ackerly, 2002; Knight *et al.*, 2005). Species that occupy such environments are expected to experience natural selection for more rapid growth and accelerated reproduction. Because of negative correlations between GS and aspects of plant growth (Bennett, 1987; Mowforth & Grime, 1989; Grotkopp *et al.*, 2004; Francis *et al.*, 2008; Bilinski *et al.*, 2018), such species also may experience natural selection for smaller GS. This selection could be relaxed, however, in environments where growing seasons are more prolonged.

Helianthus sunflowers are native to North America and consist of *c.* 50 species varying in ploidy, life history, ecology and geographic distribution (Heiser *et al.*, 1969; Kane *et al.*, 2013). GS estimates have been reported for *Helianthus* species previously (Sims & Price, 1985), although these earlier estimates were limited to 19 diploid species and were based on older Feulgen staining approaches (Sims & Price, 1985). In the current study, we provide GS estimates for 49 *Helianthus* species based on flow cytometric approaches. Our database consists of 39 diploid, six tetraploid and six hexaploid GS estimates, including 43 novel GS estimates from 41 species. Estimates are based on multiple populations per species and biological replication within populations. Flow cytometric estimates for eight species were published previously (Baack *et al.*, 2005; Qiu & Ungerer, 2018), but are included here for completeness and comparative purposes. We evaluate variation in GS across the genus as a whole and separately in diploid species in the context of a well-resolved phylogeny (Stephens *et al.*, 2015). We examine correlations between diploid GS and environmental parameters defining the diverse natural geographic ranges in which *Helianthus* species are found, and specifically test whether *Helianthus* species with smaller (larger) GS are associated with environmental parameters that define shorter (longer) growing seasons. We demonstrate strong phylogenetic trends both within and between major *Helianthus* lineages and show that, as predicted, the evolution of GS in diploid *Helianthus* species is associated with temperature seasonality and corresponding growing season variability. Finally, we demonstrate a strong negative correlation between GS and cell production rate, a surrogate measure of cell cycle time, and suggest that growth rate variability among species may provide a mechanistic basis for these environmental correlations.

Materials and Methods

Plant materials and genome size estimates

Seeds of the sunflower species utilized in this study were acquired from the United States Department of Agriculture (USDA) National Plant Germplasm System and/or were collected in the field (Supporting Information Table S1). Estimates of nuclear GS were determined by laser-based flow cytometry assaying 5000–10 000 events per sample. We followed a true internal standard protocol (*sensu* Dolezel & Bartos, 2005), which included co-chopping of fresh leaf material of sample and standard before stain application; the methods are outlined in Qiu & Ungerer (2018). Isolated nuclei were stained with propidium iodide solution (BioSure, Grass Valley, CA, USA) and sample G1 peak sizes were estimated by comparison to G1 peak sizes of a common set of internal standards (Table S1). The database of 483 individual GS measurements (Table S1) and 51 species-level mean GS estimates derived from them (Table S2) is based on data generated at multiple institutions and on multiple flow cytometer instruments. All flow cytometers used to generate data for the current study were laser-based and a common set of internal standards was used (Table S1). These shared features of data generation protocols, together with the standardized use of the intercalating dye propidium iodide, have been demonstrated to provide reliable GS estimates for plants, with only negligible differences in interlaboratory comparisons of the same samples (Dolezel *et al.*, 1998).

Ancestral state estimates and rates of genome size evolution

A maximum likelihood (ML) phylogeny of 32 diploid sunflower species (Stephens *et al.*, 2015) was used to evaluate patterns of GS evolution within a phylogenetic framework and to reconstruct an ancestral GS estimate. Stephens *et al.* (2015) recognized three main clades in this phylogeny: an annual clade of eight species, a southeastern (SE) perennial clade of eight species and a perennial clade of large-statured species (large perennial) consisting of 11 species, with five additional species located outside these clades. Molecular branch lengths of this ML tree were transformed to make them ultrametric using the penalized likelihood method implemented in r8s v.1.8 (Sanderson, 2003). The ultrametric conversion was performed with a smoothing parameter of 32, which was selected by cross validation based on two calibration dates from Strasburg *et al.* (2009): a divergence time of

1.1 million yr ago (Ma) between *H. annuus* and *H. argophyllus*, and a divergence time of 1.8 Ma between *H. annuus* and *H. petiolaris*. The ultrametric tree was then used as the reference phylogeny in comparative tests to estimate rates of GS evolution and to test environmental and phenotypic correlates of GS in diploid sunflowers.

We used the software StableTraits (Elliot & Mooers, 2014) to reconstruct ancestral GS estimates and model rates of GS evolution through time. This program samples rates from a heavy-tailed distribution, a generalization of a Brownian motion (BM) model, which allows for modeling traits evolving along the phylogeny under selection. StableTraits was run for 10 million generations, sampling every 1000 generations with two independent chains. Effective sample sizes (ESSs) were accessed using Tracer (Rambaut *et al.*, 2018) and were > 200 for all parameters estimated.

To determine whether rates of GS evolution differ among the three main clades of the diploid *Helianthus* phylogeny (i.e. annual clade, SE perennial clade and large perennial clade), we compared the fit of two BM and five Ornstein–Uhlenbeck (OU) models. Both BM and OU models estimate the rate of stochastic motion (σ^2). The major difference is that the OU process allows the trait to fluctuate around an optimum value (θ) in parameter space with a strength of attraction (α) towards that optimum, while BM allows the trait to move equally to any parameter space. BM1 and BMS assign single and multiple rates (σ^2) of random drift. OU1 and OUM model single and multiple optima (θ) for different clades with a single α and σ^2. The remaining models assume either multiple σ^2 (OUMV), multiple α (OUMA), or multiple α and σ^2 (OUMVA) among clades. For the OU process, θ_0 is dropped from the model as it is assumed that the starting value is distributed according to the stationary distribution. The best fit model was selected based on the sample size-corrected Akaike information criterion (AIC_c). Confidence intervals for parameter estimates were obtained from 100 parametric bootstraps implemented in the package OUwie (Beaulieu *et al.*, 2012) in R (R Development Core Team, 2016) based on the best fit model.

Environmental data

Geographic information of each species' range was obtained from the Global Biodiversity Information Facility website (http://www.gbif.org/). We then filtered all coordinates based on criteria that samples (1) are within North America, and (2) have valid specimen records in herbaria (i.e. excluding human observations as well as other unknown bases for records). Environmental variables of filtered geographic coordinates were obtained from the WorldClim data set (http://www.worldclim.org; Hijmans *et al.*, 2005), and include: annual mean temperature, temperature seasonality (the amount of temperature variation over a 12-month period based on the standard deviation of monthly temperature averages), annual precipitation, precipitation seasonality (the amount of precipitation variation over a 12-month period measured as a ratio of the standard deviation of the monthly total precipitation to the mean monthly total precipitation) and latitude. Species-level mean values for each climatic variable were then obtained by averaging across values from all collection locations.

Cell production rate estimates

Six seeds for each of 22 diploid *Helianthus* species (Table S2) were sterilized in a 10% bleach solution for 5 min, then rinsed five times with sterilized distilled water. Seeds were imbibed on moist filter paper overnight in Petri dishes, seed coats removed the following morning, and embryos placed on Petri dishes containing 0.8% micropropagation agar Type-II (Caisson Laboratories Inc., Smithfield, UT, USA), 100 μg ml^{-1} ampicillin and 25 μg ml^{-1} gentamicin (Invitrogen). Each replicate, six in total, consisted of a block containing three Petri dishes. For each block, seeds were randomly assigned to three Petri dishes, with two dishes containing eight seeds and one dish containing six seeds. Seeds were oriented along a line in Petri dishes such that radicle emergence was toward the center of the dish, and plates were stored near vertical in a dark cabinet at 23°C with seeds oriented horizontally at the top of the plates. Seed and root tip positions were marked daily for 5 d and root growth rate (GRr) was estimated as mm d^{-1}. On day 5, roots were harvested from plates and fixed with a 10% formaldehyde in 1× PBS solution for 3 h, rinsed and stored in 1× PBS.

Cell production rate measures the rate of increase for a given population of cells, in this case root cells (Baskin, 2000), and is significantly correlated with cell cycle duration (Beemster *et al.*, 2002). Root cell walls were stained with an orange fluorescent dye, lipophilic carbocyanine 1 mg ml^{-1} SP-DiIC18 (Invitrogen) dissolved in 100% ethanol, for 10 min and mounted on a microscope slide. A rectangle of fingernail polish was used as a spacer to reduce pressure from the coverslip. A Zeiss LSM 5 PASCAL laser-scanning confocal microscope equipped with a Zeiss Axiocam HR digital camera was used to image roots. Fluorescence emission of DiIC18 was accomplished using the 543 nm line with the 20×/0.5 objective. Using the Pascal imaging software, images were captured along the entire root beginning at the root meristem and moving proximally at a field of 450 μm. Mature cells were identified as those consistently producing root hairs (Foreman & Dolan, 2001). Cell length (L) measurements were carried out using ImageJ software (Schneider *et al.*, 2012) with *c.* 20 mature cells measured per sample. Cell production rate (P) was calculated according to P = GRr/L (Baskin, 2013), where GRr = root growth rate and L = average mature cell length. Four to six individuals per species were assayed by these methods and means were determined for a species-level cell production rate estimate.

Phylogenetic regression

Relationships between GS and climatic/phenotypic variables were tested using phylogenetic generalized least-squares (PGLS) in R packages ape and caper (Paradis *et al.*, 2004; Orme, 2013). Regressions were fit by ML and using Pagel's λ (Pagel 1999; Freckleton *et al.*, 2002) as a measure of phylogenetic signal. Pagel's λ ranges from 0 to 1, with 0 indicating no phylogenetic

signal and 1 indicating strong phylogenetic signal (equal to the Brownian expectation). Both single and multiple PGLS regressions were performed and all variables were log-transformed to ensure the data conformed to BM evolution (O'Meara *et al.*, 2006; Oliver *et al.*, 2007).

Results

Genome size variation in *Helianthus*

Fig. 1 depicts GS variability for 49 *Helianthus* species, including 39 diploid species, six tetraploid species and six hexaploid species. Estimates for *H. annuus*, *H. petiolaris*, *H. anomalus*, *H. deserticola* and *H. paradoxus* were reported previously by Baack *et al.* (2005) and estimates for *H. agrestis*, *H. carnosus* and *H. porteri* were reported previously by Qiu & Ungerer (2018). All other GS estimates were newly generated for the current study, although independent estimates for some *Helianthus* species have been reported elsewhere (Sims & Price, 1985; Tetreault & Ungerer, 2016; Mascagni *et al.*, 2017). Newly reported estimates in the current study are based on an average of 3.57 accessions per species, with between one and 10 biological replicates assayed per accession (Table S1). A total of 175 accessions of *Helianthus* were utilized for the current database.

GS estimates varied 3.59-fold for diploid species, ranging from 6.38 pg in *H. neglectus* to 22.93 pg in *H. agrestis*, with a mean of 10.13 pg. Estimates varied 1.56-fold in tetraploids, ranging from 15.68 pg in *H. hirsutus* to 24.41 pg in *H. smithii*, with a mean of 20.18 pg. In hexaploids, estimates varied 1.19-fold, ranging from 19.69 pg in *H. californicus* to 23.41 pg in *H. tuberosus*, with a mean of 22.23 pg (Fig. 1; Table S2). Average GS for hexaploids was only slightly larger than that for tetraploids, and the largest GS estimates genus-wide were unexpectedly those of two tetraploid species, *H. schweinitzii* and *H. smithii* (Fig. 1).

Patterns of genome size evolution in diploid species

A divergence time estimate of the extant diploid species is 2.394 Ma. This estimate is within the range of previous estimates, 1.7–8.2 Ma (Schilling, 1997; Kantar *et al.*, 2014; Mason, 2018). The reconstructed ancestral GS (2C) estimate for the species depicted in Fig. 2 is 8.82 pg (95% confidence interval, 7.85–10.01).

Clade-specific patterns include GS increases, decreases and relative stasis. Species in the SE perennial clade have GS estimates 1.12–1.44-fold higher than the ancestrally reconstructed estimate. By contrast, the annual clade shows a general pattern of genomic downsizing, although two annual species (*H. argophyllus* and *H. niveus*) have nuclear GS values very near to the ancestral estimate, and a third, *H. bolanderi*, has a GS estimate slightly higher than the reconstructed estimate. The large perennial clade, consisting of 11 species, shows the least amount of GS variability and evolutionary change, with all estimates between 0.91- and 1.07-fold that of the ancestrally reconstructed estimate.

The most pronounced pattern of GS evolution for an individual species is large-scale GS expansion for the annual diploid species *H. agrestis*. GS for this species is 2.60-fold higher than the ancestrally reconstructed estimate and 1.59- to 3.59-fold higher than for any other *Helianthus* diploid species. Although an annual, this species is not part of the annual clade, but phylogenetically positioned basally to diploid perennial *Helianthus* species (Fig. 2).

For rate test comparisons among the three clades, the best fit model was an OUMV model (Table S3), which infers a different rate parameter (σ^2) and a different optimum value (θ) for each clade. The large perennial clade was inferred to have the lowest rate of GS evolution (Table S4; Fig. S1), estimated to be 3.7 times and 10 times lower than the SE perennial and annual clades, respectively. This result is consistent with the StableTrait estimate, which shows the large perennial clade to have the shortest branch lengths (Fig. 2). The optimum GS was higher in the SE perennial clade (back-transformed 2C = 11.68) than in the large perennial clade (back-transformed 2C = 8.75) and annual clade (back-transformed 2C = 7.96). Taken together, these results suggest that *Helianthus* species in the annual clade have evolved towards a smaller GS with higher rates of evolution compared to species in the SE perennial and large perennial clades.

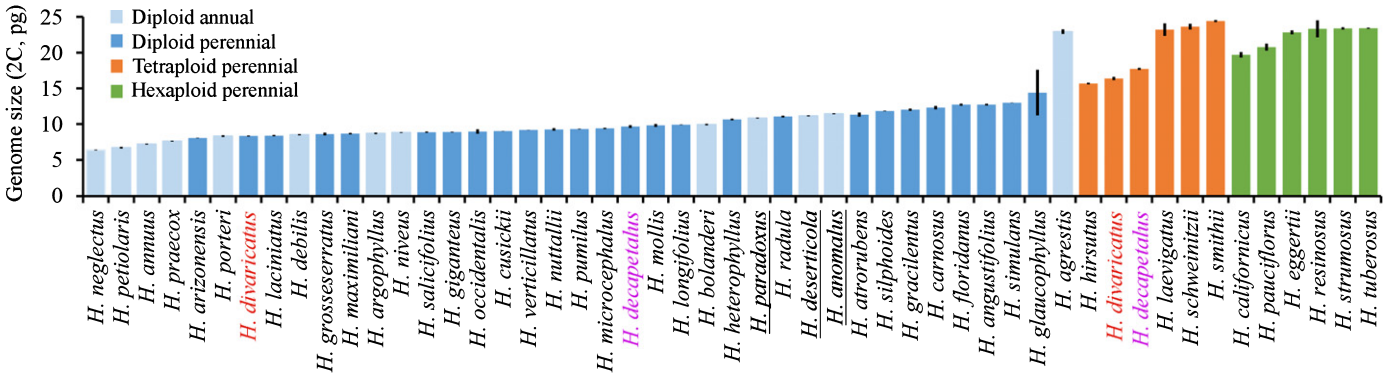

Fig. 1 Genome size (GS) estimates for 49 *Helianthus* species. Histogram bars represent averages of population means, with an average of 3.57 populations assayed per species (Supporting Information Table S2). Error bars indicate ± 1 SE and represent the standard error of the population means. Species exhibiting cytotypic variation (both diploid and tetraploid populations) are indicated with colored text and species of diploid hybrid origin are underlined.

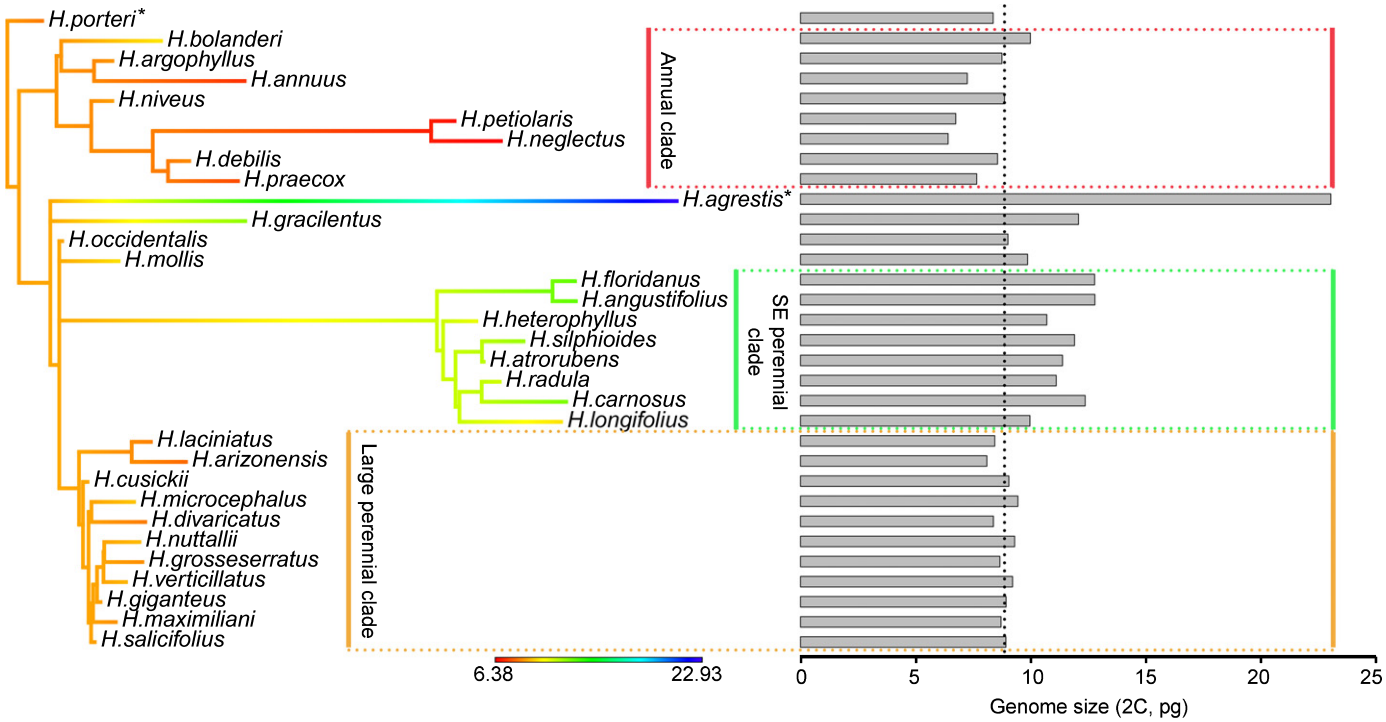

Fig. 2 Phylogenetic relationships and nuclear genome size (GS; 2C) estimates (gray bars) for 32 diploid *Helianthus* species. Tree topology is based on the phylogeny presented by Stephens *et al.* (2015). Branch lengths depict rates of GS evolution and the vertical dotted black line represents the ancestral GS estimate (8.82 pg; 95% confidence interval 7.85–10.01). Asterisks represent two annual *Helianthus* species phylogenetically positioned outside of the annual clade. SE, southeastern.

GS and environmental correlates in diploid *Helianthus* species

GS and two environmental parameters examined in the study, temperature seasonality and precipitation seasonality, showed phylogenetic signal with λ differing significantly from zero (Table S5). ML estimation of λ ranged from 0.798 to 1.000 for the other three environmental parameters, indicating that environmental characteristics experienced by closely related species are more similar than expected by chance. In both single and multiple PGLS regressions, GS was significantly and negatively associated with temperature seasonality (Fig. 3; Tables 1, 2).

GS and cell production rate are negatively correlated in diploid *Helianthus* species

A significant negative correlation was observed between GS and cell production rate based on a subset of 22 diploid *Helianthus* species (Fig. 4; $r = -0.678$; $P < 0.001$). This negative relationship remained significant when excluding the three annual hybrid species (i.e. $n = 19$) and controlling for phylogenetic relationships via PGLS ($r = -0.179$; $P = 0.001$).

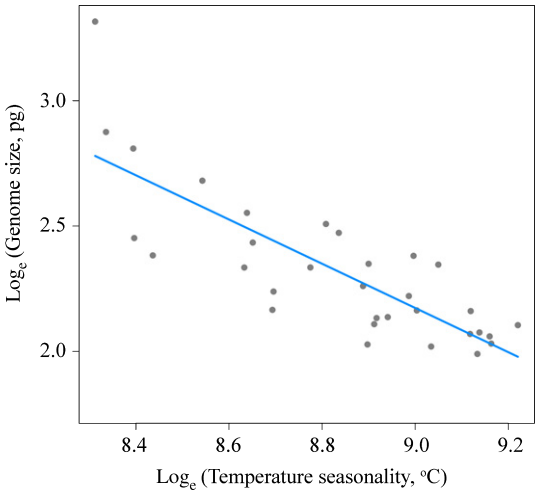

Fig. 3 Partial regression leverage plot showing the negative relationship between temperature seasonality and genome size (GS). Data points represent the residual values after accounting for effects of other environmental factors.

Discussion

The vast GS variability that exists among plant species has intrigued biologists for decades and stimulated efforts both to catalogue variation across diverse plant groups and to understand the mechanisms through which this variation arose and ultimately is controlled (Knight *et al.*, 2005; Hawkins *et al.*, 2008; Whitney *et al.*, 2010; Leitch & Leitch, 2013; Pellicer *et al.*, 2018). Phylogenetic comparative analyses provide the most powerful approach for elucidating patterns of GS expansion and contraction in light of evolutionary relationships and for exploring the potential causes and consequences of such changes through correlational analyses of GS with environmental and biological

Table 1 Single phylogenetic generalized least-squares regressions between genome size and environmental variables based on 32 diploid species.

	Coefficient	*t*-value	*P*-value
Latitude (°)	−0.315	−1.524	0.138
Annual mean temperature (°C)	0.176	1.605	0.119
Temperature seasonality (°C)	−0.335	−2.977	**0.006**
Annual precipitation (mm)	0.149	2.570	0.015
Precipitation seasonality	−0.070	−0.961	0.344

Bold type indicates a significant correlation after Bonferroni correction ($\alpha = 0.01$).

Table 2 Multiple phylogenetic generalized least-squares regressions between genome size and environmental variables based on 32 diploid species.

	Coefficient	*t*-value	*P*-value
Latitude (°)	1.677	2.718	0.012
Annual mean temperature (°C)	0.473	1.823	0.080
Temperature seasonality (°C)	−0.823	−4.033	**0.000**
Annual precipitation (mm)	0.090	1.402	0.173
Precipitation seasonality	0.013	0.188	0.852

Bold type indicates a significant correlation after Bonferroni correction ($\alpha = 0.01$).

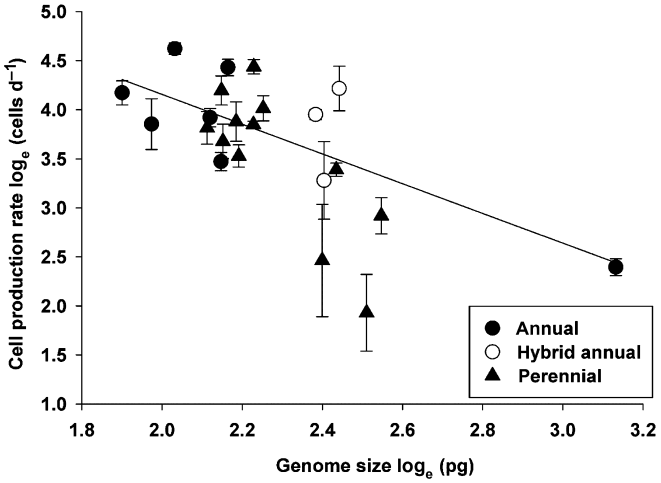

Fig. 4 The negative relationship between genome size (GS) and cell production rate for 22 diploid *Helianthus* species (uncorrected analysis: $r = -0.678$, $P < 0.001$). This relationship remains significant when the three species of diploid hybrid origin (white circles) are excluded and phylogenetic relationships are accounted for (PGLS analysis: $r = -0.179$, $P = 0.001$). Error bars indicate ± 1 SE of the mean.

variables. Such studies have been performed at multiple phylogenetic levels, from comparisons of diverse land plants (Beaulieu *et al.*, 2007a,b; Knight & Beaulieu, 2008; Whitney *et al.*, 2010; Bainard *et al.*, 2012; Vesely *et al.*, 2012; Lomax *et al.*, 2014; Alonso *et al.*, 2015; Bromham *et al.*, 2015) to taxonomically more restricted analyses within plant families (Veleba *et al.*, 2014; Carta & Peruzzi, 2016; Ng *et al.*, 2016), and genera (Albach & Greilhuber, 2004; Grotkopp *et al.*, 2004; Jakob *et al.*, 2004; Gallagher *et al.*, 2011; Kang *et al.*, 2014; Meudt *et al.*, 2015; Baniaga *et al.*, 2016; Mandak *et al.*, 2016; Du *et al.*, 2017). Because of the highly dynamic nature of GS evolution and potential problems of representative sampling within groups for broader taxonomic comparisons, studies at lower taxonomic ranks – i.e. genera – may provide the highest explanatory resolution for understanding the forces governing the evolution of GS (Hawkins *et al.*, 2008).

In the current study, we provide a comprehensive GS database for species in the sunflower genus *Helianthus* and present an initial analysis of phylogenetic trends and environmental/phenotypic correlates of GS variability in diploid *Helianthus* species utilizing a well-resolved and robust phylogeny (Stephens *et al.*, 2015). Flow cytometry estimates presented here based on multiple populations per species and biological replication within populations exhibit variable correspondence with some of the earliest estimates for *Helianthus* species generated by Feulgen staining. For example, GS estimates for 19 *Helianthus* species reported by Sims & Price (1985) differ, on average, by *c.* 6% from corresponding flow cytometry estimates reported herein, with estimates for eight species differing by < 3.5% and estimates for three species differing by > 10%.

While considerable GS variability is found within both diploid and tetraploid *Helianthus* species, a comparison of mean GS estimates across ploidy levels closely matches a predicted 2-fold increase in tetraploids (20.18 pg) vs diploids (10.13 pg). This observation does not necessarily support the notion that genomic downsizing has been absent in *Helianthus* tetraploids, however, as GS comparisons among direct progenitor–derivative species would provide a better assessment of the prevalence of genomic downsizing. Indeed, for three tetraploid species that arose via autopolyploidy (*H. decepetalus*, *H. divaricatus* and *H. hirsutus*) and for which progenitors are known, evidence of genomic downsizing is apparent in all three cases (Table S6). For the single hexaploid species for which ancestry has been determined – *H. tuberosus* (23.41 pg) derived from *H. grosseserratus* (8.62 pg) and *H. hirsutus* (15.68 pg) (Bock *et al.*, 2014) – there again is evidence of modest genomic downsizing (i.e. observed and expected = 23.41 and 24.30 pg, respectively). It is noteworthy that all six hexaploid species possess GS values that are lower than those estimated for the two largest tetraploid species (i.e. *H. schweinitzii* and *H. smithii*). The underlying causes of these observations currently are unknown, but include moderate to larger-scale genomic downsizing in hexaploids and/or GS expansion for other tetraploid *Helianthus* species such as *H. laevigatus*, *H. schweinitzii* and *H. smithii* (Fig. 1). GS evolutionary dynamics in polyploid *Helianthus* species will be the focus of future investigations.

Within diploid *Helianthus* species, perennial species typically have larger genomes than annual species (Fig. 1). This observation is generally consistent with patterns observed across most plant groups assayed to date (Bennett, 1972; Albach & Greilhuber, 2004). A noteworthy exception is the annual species *H. agrestis*, with the largest GS by far for any diploid. GS in

H. agrestis more closely resembles estimates for tetraploids and hexaploids than other diploids. *H. agrestis* previously was demonstrated to have experienced proliferation events of LTR retrotransposons (Tetreault & Ungerer, 2016). Amplification of these sequences represents a known mechanism of rapid and potentially large-scale genome expansion in plants (Park *et al.*, 2012; Estep *et al.*, 2013; Bennetzen & Wang, 2014). Proliferation events of LTR retrotransposons also have been implicated in genome expansion in three additional annual *Helianthus* diploid species of ancient hybrid origins, *H. anomalus*, *H. deserticola* and *H. paradoxus* (Ungerer *et al.*, 2006). While GS estimates for the three ancient hybrid species are considerably lower than for *H. agrestis*, and lower than those reported for several diploid perennial *Helianthus* species (Fig. 1), the three hybrid species represent the next-largest annual diploid GS estimates after *H. agrestis* (Fig. 1).

Phylogenetic trends of GS evolution in *Helianthus* diploid species

Different directions and rates of GS evolution are apparent among the three well-supported *Helianthus* clades. Species in the annual clade have, on average, experienced genomic contraction compared to the ancestrally reconstructed estimate (Fig. 2). This pattern is consistent with smaller genomes favored in annual species due to selection for more accelerated growth and faster reproduction associated with an annual life history (Bennett, 1972). Driving this pattern may be a negative correlation between GS and cell production rate (Fig. 4), as nuclear DNA must be fully copied when cells divide mitotically. GS may thus place constraints on cell division rates and ultimately whole-plant relative growth rates in *Helianthus*, as has been described for angiosperms more generally (Knight *et al.*, 2005; Greilhuber & Leitch, 2013). The annual clade also exhibits the highest normalized rate of GS evolution standardized against the large perennial clade (Fig. S1).

By contrast, all species in the SE perennial clade have experienced genomic expansion relative to the ancestral reconstructed estimate, albeit at a lower normalized rate of evolution (Fig. S1). It is interesting to note that, as the clade name implies, species within this monophyletic group are generally restricted geographically to the SE region of the United States where climate conditions accommodate longer growing seasons. Selection for more accelerated growth and faster reproduction in advance of late-season adverse weather conditions might be relaxed in such regions with longer growing seasons. Under such circumstances, genome expansion via proliferation of repetitive sequences such as LTR retrotransposons could be more permissible. The relationship between GS and environmental variables associated with growing season length are explored in further detail in the Environmental and biological correlates section below.

The third main clade of *Helianthus* diploids, the large-statured perennial clade, displays far less variability among species, with all GS estimates notably similar to the reconstructed ancestral estimate. Species within this clade are found throughout most of the continental United States and most have wide geographic ranges, although some (e.g. *H. arizonensis*, *H. laciniatus* and *H. verticillatus*) are more geographically restricted (Heiser *et al.*, 1969; Stephens *et al.*, 2015). The observed variability among these three *Helianthus* clades is interesting and provides additional evidence that evolutionary trends of GS variation are not uniformly in the direction of GS expansion (Bennetzen & Kellogg, 1997; Hawkins *et al.*, 2009).

Environmental and biological correlates

Relationships between GS and environmental and/or climatic variables have been explored in several plant groups with comparisons drawn at various taxonomic levels (Knight & Ackerly, 2002; Grotkopp *et al.*, 2004; Vesely *et al.*, 2012; Diez *et al.*, 2013; Kang *et al.*, 2014; Carta & Peruzzi, 2016; Du *et al.*, 2017; Li *et al.*, 2017; Bilinski *et al.*, 2018; Lyu *et al.*, 2018). While findings have not been universally consistent (Knight *et al.*, 2005), multiple studies have suggested that environmental conditions ultimately may place constraints on the evolution of GS.

In the current study, both single and multiple PGLS regressions identified a significant correlation between GS and temperature seasonality in the diploid dataset (Fig. 3). Temperature seasonality measures changes in temperature over a 12-month period; in the native range of wild *Helianthus* species, geographic regions with lower (higher) temperature seasonality experience longer (shorter) growing seasons. These observations suggest that the evolution of larger GS in diploid *Helianthus* species may be more permissible in habitats with longer growing seasons and constrained in regions where growing seasons are shorter. Qualitatively similar findings have been reported in other plant GS variation studies at the generic rank, at least for some environmental variables (Grotkopp *et al.*, 2004; Carta & Peruzzi, 2016), although such results are not universally observed (Kang *et al.*, 2014).

While the causes of these patterns currently are unknown, an ultimate explanation may relate to effects of GS variation on cell production rates (Fig. 4), and potential cascading effects on whole-plant relative growth rates. In angiosperms, strong negative correlations are found between nuclear GS and cell cycle rates (Francis *et al.*, 2008), as nuclear DNA must be fully copied each time a cell divides. As such, species that experience natural selection for more rapid growth and earlier reproduction may experience simultaneous selection for smaller GS (Bennett, 1972). In regions with longer growing seasons, however, this selection may be relaxed. Consistent with this interpretation, cell production rate for the 19 non-hybrid diploid *Helianthus* species (see Fig. 4) is positively correlated with temperature seasonality of the species' natural ranges (PGLS: $r = 1.723$; $P = 0.003$). A negative correlation between GS and root meristem growth rate (RMGR) was demonstrated previously (Gruner *et al.*, 2010), with implications of reduced RMGR on root function and nutrient uptake discussed in light of the evolution of GS. It is perhaps also noteworthy that the *Helianthus* diploid species with by far the largest nuclear genome, *H. agrestis*, although an annual, has a geographically restricted range limited largely to central and southern Florida, a natural habitat with lower temperature seasonality and a longer growing season. Future work in the

Helianthus system will investigate whether diploid species differ in predictable physiological and growth phenotypes related to the general predictions described above.

Conclusions

GS in *Helianthus* sunflowers varies considerably among species. This variability is influenced by ploidy variation and also is likely to be impacted by differential activity and proliferation of transposable elements such as LTR retrotransposons. Tetraploid and hexaploid species displayed unexpectedly similar GS estimates, probably as a result of genome contraction in hexaploids and/or GS expansion in some tetraploids. Phylogenetic comparative analyses of diploid *Helianthus* species demonstrate clade-specific patterns of genomic contraction, genomic expansion and relative stasis. PGLS analyses of GS with multiple environmental variables suggest that GS evolution also may be influenced by ecological factors such as growing season length. The database presented here and corresponding analyses will facilitate ongoing and future research on the mechanisms of GS increases and decreases in this important group of native North American plants.

Acknowledgements

We thank Anastasia Weston, David Carius, Isabel De Ayala, Abigail Mark, Tobias Baskin, Dan Boyle and Megan Martinez for technical assistance. This work was supported by NSF DEB 1342787 to MCU, Jesse Nippert and Lynn Clark, Kansas State University and NSF DEB 1257965 to KDW and LHR.

Author contributions

FQ and MCU planned and designed the research. FQ, EJB, KDW, DGB, HMT, LHR and MCU performed experiments, FQ analyzed the data, FQ and MCU wrote the manuscript, and EJB, KDW, DGB, HMT and LHR provided comments on drafts of the manuscript. All authors read and approved the final manuscript.

ORCID

Eric J. Baack http://orcid.org/0000-0002-2492-8546
Dan G. Bock http://orcid.org/0000-0001-7788-1705
Loren H. Rieseberg http://orcid.org/0000-0002-2712-2417
Hannah M. Tetreault http://orcid.org/0000-0001-5911-7316
Kenneth D. Whitney http://orcid.org/0000-0002-2523-5469

References

Albach DC, Greilhuber J. 2004. Genome size variation and evolution in *Veronica*. *Annals of Botany* **94**: 897–911.

Alonso C, Perez R, Bazaga P, Herrera CM. 2015. Global DNA cytosine methylation as an evolving trait: phylogenetic signal and correlated evolution with genome size in angiosperms. *Frontiers in Genetics* **6**: 4.

Baack EJ, Whitney KD, Rieseberg LH. 2005. Hybridization and genome size evolution: timing and magnitude of nuclear DNA content increases in *Helianthus* homoploid hybrid species. *New Phytologist* **167**: 623–630.

Bainard JD, Bainard LD, Henry TA, Fazekas AJ, Newmaster SG. 2012. A multivariate analysis of variation in genome size and endoreduplication in angiosperms reveals strong phylogenetic signal and association with phenotypic traits. *New Phytologist* **196**: 1240–1250.

Baniaga AE, Arrigo N, Barker MS. 2016. The small nuclear genomes of *Selaginella* are associated with a low rate of genome size evolution. *Genome Biology and Evolution* **8**: 1516–1525.

Baskin TI. 2000. On the constancy of cell division rate in the root meristem. *Plant Molecular Biology* **43**: 545–554.

Baskin TI. 2013. Patterns of root growth acclimation: constant processes, changing boundaries. *Wiley Interdisciplinary Reviews-Developmental Biology* **2**: 65–73.

Beaulieu JM, Jhwueng DC, Boettiger C, O'Meara BC. 2012. Modeling stabilizing selection: expanding the Ornstein-Uhlenbeck model of adaptive evolution. *Evolution* **66**: 2369–2383.

Beaulieu JM, Leitch IJ, Knight CA. 2007a. Genome size evolution in relation to leaf strategy and metabolic rates revisited. *Annals of Botany* **99**: 495–505.

Beaulieu JM, Moles AT, Leitch IJ, Bennett MD, Dickie JB, Knight CA. 2007b. Correlated evolution of genome size and seed mass. *New Phytologist* **173**: 422–437.

Beemster GTS, De Vusser K, De Tavernier E, De Bock K, Inze D. 2002. Variation in growth rate between *Arabidopsis* ecotypes is correlated with cell division and A-type cyclin-dependent kinase activity. *Plant Physiology* **129**: 854–864.

Bennett MD. 1972. Nuclear DNA content and minimum generation time in herbaceous plants. *Proceedings of the Royal Society of London B: Biological Sciences* **181**: 109–135.

Bennett MD. 1987. Variation in genomic form in plants and its ecological implications. *New Phytologist* **106**: 177–200.

Bennetzen JL, Kellogg EA. 1997. Do plants have a one-way ticket to genomic obesity? *Plant Cell* **9**: 1509–1514.

Bennetzen JL, Wang H. 2014. The contributions of transposable elements to the structure, function, and evolution of plant genomes. *Annual Review of Plant Biology* **65**: 505–530.

Bilinski P, Albert PS, Berg JJ, Birchler JA, Grote MN, Lorant A, Quezada J, Swarts K, Yang J, Ross-Ibarra J. 2018. Parallel altitudinal clines reveal trends in adaptive evolution of genome size in *Zea mays*. *PLoS Genetics* **14**: e1007162.

Bock DG, Kane NC, Ebert DP, Rieseberg LH. 2014. Genome skimming reveals the origin of the Jerusalem artichoke tuber crop species: neither from Jerusalem nor an artichoke. *New Phytologist* **201**: 1021–1030.

Bromham L, Hua X, Lanfear R, Cowman PF. 2015. Exploring the relationships between mutation rates, life history, genome size, environment, and species richness in flowering plants. *American Naturalist* **185**: 507–524.

Carta A, Peruzzi L. 2016. Testing the large genome constraint hypothesis: plant traits, habitat and climate seasonality in Liliaceae. *New Phytologist* **210**: 709–716.

Devos KM, Brown JK, Bennetzen JL. 2002. Genome size reduction through illegitimate recombination counteracts genome expansion in Arabidopsis. *Genome Research* **12**: 1075–1079.

Diez CM, Gaut BS, Meca E, Scheinvar E, Montes-Hernandez S, Eguiarte LE, Tenaillon MI. 2013. Genome size variation in wild and cultivated maize along altitudinal gradients. *New Phytologist* **199**: 264–276.

Dolezel J, Bartos J. 2005. Plant DNA flow cytometry and estimation of nuclear genome size. *Annals of Botany* **95**: 99–110.

Dolezel J, Greilhuber J, Lucretti S, Meister A, Lysak MA, Nardi L, Obermayer R. 1998. Plant genome size estimation by flow cytometry: inter-laboratory comparison. *Annals of Botany* **82**: 17–26.

Du YP, Bi Y, Zhang MF, Yang FP, Jia GX, Zhang XH. 2017. Genome size diversity in *Lilium* (Liliaceae) is correlated with karyotype and environmental traits. *Frontiers in Plant Science* **8**: 1303.

Elliot MG, Mooers AO. 2014. Inferring ancestral states without assuming neutrality or gradualism using a stable model of continuous character evolution. *BMC Evolutionary Biology* **14**: 226.

Estep MC, DeBarry JD, Bennetzen JL. 2013. The dynamics of LTR retrotransposon accumulation across 25 million years of panicoid grass evolution. *Heredity* **110**: 194–204.

Foreman J, Dolan L. 2001. Root hairs as a model system for studying plant cell growth. *Annals of Botany* **88**: 1–7.

Francis D, Davies MS, Barlow PW. 2008. A strong nucleotypic effect on the cell cycle regardless of ploidy level. *Annals of Botany* **101**: 747–757.

Freckleton RP, Harvey PH, Pagel M. 2002. Phylogenetic analysis and comparative data: a test and review of evidence. *American Naturalist* **160**: 712–726.

Gallagher RV, Leishman MR, Miller JT, Hui C, Richardson DM, Suda J, Travnicek P. 2011. Invasiveness in introduced Australian acacias: the role of species traits and genome size. *Diversity and Distributions* **17**: 884–897.

Greilhuber J, Leitch IJ. 2013. Genome size and the phenotype. In: Leitch IJ, ed. *Plant genome diversity, vol. 2*. Vienna, Austria: Springer, 323–344.

Grotkopp E, Rejmanek M, Sanderson MJ, Rost TL. 2004. Evolution of genome size in pines (*Pinus*) and its life-history correlates: supertree analyses. *Evolution* **58**: 1705–1729.

Grover CE, Yu Y, Wing RA, Paterson AH, Wendel JF. 2008. A phylogenetic analysis of indel dynamics in the cotton genus. *Molecular Biology and Evolution* **25**: 1415–1428.

Gruner A, Hoverter N, Smith T, Knight CA. 2010. Genome size is a strong predictor of root meristem growth rate. *Journal of Botany* **2010**: 390414.

Hawkins JS, Grover CE, Wendel JF. 2008. Repeated big bangs and the expanding universe: directionality in plant genome size evolution. *Plant Science* **174**: 557–562.

Hawkins JS, Proulx SR, Rapp RA, Wendel JF. 2009. Rapid DNA loss as a counterbalance to genome expansion through retrotransposon proliferation in plants. *Proceedings of the National Academy of Sciences, USA* **106**: 17811–17816.

Heiser CB, Smith DM, Clevenger SB, Martin WC. 1969. The North American sunflowers (*Helianthus*). *Memoirs of the Torrey Botanical Club* **22**: 1–218.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**: 1965–1978.

Ivanov VB. 1978. DNA content in the nucleus and rate of development in plants. *Soviet Journal of Developmental Biology* **9**: 39–53.

Jakob SS, Meister A, Blattner FR. 2004. Considerable genome size variation of *Hordeum* species (Poaceae) is linked to phylogeny, life form, ecology, and speciation rates. *Molecular Biology and Evolution* **21**: 860–869.

Kane NC, Burke JM, Marek L, Seiler G, Vear F, Baute G, Knapp SJ, Vincourt P, Rieseberg LH. 2013. Sunflower genetic, genomic and ecological resources. *Molecular Ecology Resources* **13**: 10–20.

Kang M, Tao JJ, Wang J, Ren C, Qi QW, Xiang QY, Huang HW. 2014. Adaptive and nonadaptive genome size evolution in karst endemic flora of China. *New Phytologist* **202**: 1371–1381.

Kantar MB, Baute GJ, Bock DG, Rieseberg LH. 2014. Genomic variation in *Helianthus*: learning from the past and looking to the future. *Briefings in Functional Genomics* **13**: 328–340.

Kelly LJ, Leitch IJ. 2011. Exploring giant plant genomes with next-generation sequencing technology. *Chromosome Research* **19**: 939–953.

Kelly LJ, Renny-Byfield S, Pellicer J, Macas J, Novak P, Neumann P, Lysak MA, Day PD, Berger M, Fay MF, *et al.* 2015. Analysis of the giant genomes of *Fritillaria* (Liliaceae) indicates that a lack of DNA removal characterizes extreme expansions in genome size. *New Phytologist* **208**: 596–607.

Knight CA, Ackerly DD. 2002. Variation in nuclear DNA content across environmental gradients: a quantile regression analysis. *Ecology Letters* **5**: 66–76.

Knight CA, Beaulieu JM. 2008. Genome size scaling through phenotype space. *Annals of Botany* **101**: 759–766.

Knight CA, Molinari NA, Petrov DA. 2005. The large genome constraint hypothesis: evolution, ecology and phenotype. *Annals of Botany* **95**: 177–190.

Kubesova M, Moravcova L, Suda J, Jarosik V, Pysek P. 2010. Naturalized plants have smaller genomes than their non-invading relatives: a flow cytometric analysis of the Czech alien flora. *Preslia* **82**: 81–96.

Kumar A, Bennetzen JL. 1999. Plant retrotransposons. *Annual Review of Genetics* **33**: 479–532.

Leitch IJ, Leitch AR. 2013. Genome size diversity and evolution in land plants. In: Leitch IJ, ed. *Plant genome diversity, vol. 2*. Vienna, Austria: Springer, 307–322.

Li DD, Lu YL, Guo SL, Yin LP, Zhou P, Lou YX. 2017. Nuclear DNA contents of *Echinchloa crus-galli* and its Gaussian relationships with environments. *Acta Oecologica-International Journal of Ecology* **79**: 36–47.

Lomax BH, Hilton J, Bateman RM, Upchurch GR, Lake JA, Leitch IJ, Cromwell A, Knight CA. 2014. Reconstructing relative genome size of vascular plants through geological time. *New Phytologist* **201**: 636–644.

Lyu HM, He ZW, Wu CI, Shi SH. 2018. Convergent adaptive evolution in marginal environments: unloading transposable elements as a common strategy among mangrove genomes. *New Phytologist* **217**: 428–438.

Mandak B, Krak K, Vit P, Pavlikova Z, Lomonosova MN, Habibi F, Wang L, Jellen EN, Douda J. 2016. How genome size variation is linked with evolution within *Chenopodium* sensu lato. *Perspectives in Plant Ecology Evolution and Systematics* **23**: 18–32.

Mascagni F, Giordani T, Ceccarelli M, Cavallini A, Natali L. 2017. Genome-wide analysis of LTR-retrotransposon diversity and its impact on the evolution of the genus *Helianthus* (L.). *BMC Genomics* **18**: 634.

Mason CM. 2018. How old are sunflowers? A molecular clock analysis of key divergences in the origin and diversification of *Helianthus* (Asteraceae). *International Journal of Plant Sciences* **179**: 182–191.

Meudt HM, Rojas-Andres BM, Prebble JM, Low E, Garnock-Jones PJ, Albach DC. 2015. Is genome downsizing associated with diversification in polyploid lineages of *Veronica*? *Botanical Journal of the Linnean Society* **178**: 243–266.

Mowforth MA, Grime JP. 1989. Intra-population variation in nuclear-DNA amount, cell-size and growth-rate in *Poa annua* L. *Functional Ecology* **3**: 289–295.

Ng CH, Lee SL, Tnah LH, Ng KKS, Lee CT, Madon M. 2016. Genome size variation and evolution in Dipterocarpaceae. *Plant Ecology Diversity* **9**: 437–446.

Oliver MJ, Petrov D, Ackerly D, Falkowski P, Schofield OM. 2007. The mode and tempo of genome size evolution in eukaryotes. *Genome Research* **17**: 594–601.

O'Meara BC, Ane C, Sanderson MJ, Wainwright PC. 2006. Testing for different rates of continuous trait evolution using likelihood. *Evolution* **60**: 922–933.

Orme D. 2013. The caper package: comparative analysis of phylogenetics and evolution in R. *R package version* v. *0.5.2*. [WWW document] URL https://cran.mtu.edu/web/packages/caper/index.html.

Pandit MK, White SM, Pocock MJO. 2014. The contrasting effects of genome size, chromosome number and ploidy level on plant invasiveness: a global analysis. *New Phytologist* **203**: 697–703.

Pagel M. 1999. Inferring the historical patterns of biological evolution. *Nature* **401**: 877–884.

Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290.

Park M, Park J, Kim S, Kwon JK, Park HM, Bae IH, Yang TJ, Lee YH, Kang BC, Choi D. 2012. Evolution of the large genome in *Capsicum annuum* occurred through accumulation of single-type long terminal repeat retrotransposons and their derivatives. *Plant Journal* **69**: 1018–1029.

Pellicer J, Hidalgo O, Dodsworth S, Leitch IJ. 2018. Genome size diversity and its impact on the evolution of land plants. *Genes* **9**: 88.

Qiu F, Ungerer MC. 2018. Genomic abundance and transcriptional activity of diverse gypsy and copia long terminal repeat retrotransposons in three wild sunflower species. *BMC Plant Biology* **18**: 6.

R Development Core Team. 2016. *R: a language and environment for statistical computing. R v. 3.2.1*. Vienna, Austria: R Foundation for Statistical Computing.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67**: 901–904.

Sanderson MJ. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* **19**: 301–302.

Schilling EE. 1997. Phylogenetic analysis of *Helianthus* (Asteraceae) based on chloroplast DNA restriction site data. *Theoretical and Applied Genetics* **94**: 925–933.

Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH image to ImageJ: 25 years of image analysis. *Nature Methods* **9**: 671–675.

Schubert I, Vu GTH. 2016. Genome stability and evolution: attempting a holistic view. *Trends in Plant Science* **21**: 749–757.

Sims LE, Price HJ. 1985. Nuclear DNA content variation in *Helianthus* (Asteraceae). *American Journal of Botany* **72**: 1213–1219.
Smarda P, Bures P. 2010. Understanding intraspecific variation in genome size in plants. *Preslia* **82**: 41–61.
Stephens JD, Rogers WL, Mason CM, Donovan LA, Malmberg RL. 2015. Species tree estimation of diploid *Helianthus* (Asteraceae) using target enrichment. *American Journal of Botany* **102**: 910–920.
Strasburg JL, Scotti-Saintagne C, Scotti I, Lai Z, Rieseberg LH. 2009. Genomic patterns of adaptive divergence between chromosomally differentiated sunflower species. *Molecular Biology and Evolution* **26**: 1341–1355.
Suda J, Meyerson LA, Leitch IJ, Pysek P. 2015. The hidden side of plant invasions: the role of genome size. *New Phytologist* **205**: 994–1007.
Tetreault HM, Ungerer MC. 2016. Long terminal repeat retrotransposon content in eight diploid sunflower species inferred from next-generation sequence data. *G3: Genes – Genomes – Genetics* **6**: 2299–2308.
Thompson K. 1990. Genome size, seed size and germination temperature in herbaceous Angiosperms. *Evolutionary Trends in Plants* **4**: 113–116.
Ungerer MC, Strakosh SC, Zhen Y. 2006. Genome expansion in three hybrid sunflower species is associated with retrotransposon proliferation. *Current Biology* **16**: R872–R873.
Veleba A, Bures P, Adamec L, Smarda P, Lipnerova I, Horova L. 2014. Genome size and genomic GC content evolution in the miniature genome-sized family Lentibulariaceae. *New Phytologist* **203**: 22–28.
Vesely P, Bures P, Smarda P, Pavlicek T. 2012. Genome size and DNA base composition of geophytes: the mirror of phenology and ecology? *Annals of Botany* **109**: 65–75.
Whitney KD, Baack EJ, Hamrick JL, Godt MJW, Barringer BC, Bennett MD, Eckert CG, Goodwillie C, Kalisz S, Leitch IJ, *et al.* 2010. A role for nonadaptive processes in plant genome size evolution? *Evolution* **64**: 2097–2109.
Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences, USA* **106**: 13875–13879.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article:

Fig. S1 Rate comparisons across three clades with normalized rates of genome size evolution.

Table S1 Individual GS estimates for 49 *Helianthus* species.

Table S2 Fifty-one flow cytometry estimates of 2C nuclear genome size for 49 species in *Helianthus.*

Table S3 Results of genome size evolution models based on 32 diploid *Helianthus* species.

Table S4 Estimated parameters of genome size evolution for the three main *Helianthus* clades.

Table S5 Phylogenetic signal tests.

Table S6 Observed and expected (based on additivity) genome size for three tetraploid *Helianthus* species and one hexaploid *Helianthus* species.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.