UC Davis

UC Davis Previously Published Works

Title

Genome size evolution is associated with climate seasonality and glucosinolates, but not life history, soil nutrients or range size, across a clade of mustards

Permalink

https://escholarship.org/uc/item/5dw382r8

Journal

Annals of Botany, 127(7)

ISSN

0305-7364

Authors

Cacho, N Ivalú
McIntyre, Patrick J
Kliebenstein, Daniel J
et al.

Publication Date

2021-06-24

DOI

10.1093/aob/mcab028

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Annals of Botany **127**: 887–902, 2021
doi: 10.1093/aob/mcab028, available online at www.academic.oup.com/aob

ANNALS OF BOTANY
Founded 1887

Genome size evolution is associated with climate seasonality and glucosinolates, but not life history, soil nutrients or range size, across a clade of mustards

N. Ivalú Cacho[1,2,3,*,†], Patrick J. McIntyre[2,4,†], Daniel J. Kliebenstein[5,6] and Sharon Y. Strauss[2]

[1]Instituto de Biología, Universidad Nacional Autónoma de México. Circuito Exterior, Ciudad Universitaria, Mexico City 04510, Mexico, [2]Center for Population Biology, University of California, One Shields Avenue, Davis, CA 95616, USA, [3]Department of Evolution of Ecology, University of California, One Shields Avenue, Davis, CA 95616, USA, [4]NatureServe, 1680 38th Street Suite 120, Boulder, CO 80301, USA, [5]Department of Plant Sciences, University of California, One Shields Avenue, Davis, CA 95616, USA and [6]DynaMo Centre of Excellence, University of Copenhagen, Thorvaldsensvej 40, DK-1871, Frederiksberg C, Denmark
[†]These authors contributed equally to this work.
** For correspondence. E-mail ivalu.cacho@gmail.com*

Received: 26 September 2020 Returned for revision: 15 February 2021 Editorial decision: 17 February 2021 Accepted: 21 February 2021
Electronically published: 1 March 2021

• **Background and Aims** We investigate patterns of evolution of genome size across a morphologically and ecologically diverse clade of Brassicaceae, in relation to ecological and life history traits. While numerous hypotheses have been put forward regarding autecological and environmental factors that could favour small vs. large genomes, a challenge in understanding genome size evolution in plants is that many hypothesized selective agents are intercorrelated.
• **Methods** We contribute genome size estimates for 47 species of *Streptanthus* Nutt. and close relatives, and take advantage of many data collections for this group to assemble data on climate, life history, soil affinity and composition, geographic range and plant secondary chemistry to identify simultaneous correlates of variation in genome size in an evolutionary framework. We assess models of evolution across clades and use phylogenetically informed analyses as well as model selection and information criteria approaches to identify variables that can best explain genome size variation in this clade.
• **Key Results** We find differences in genome size and heterogeneity in its rate of evolution across subclades of *Streptanthus* and close relatives. We show that clade-wide genome size is positively associated with climate seasonality and glucosinolate compounds. Model selection and information criteria approaches identify a best model that includes temperature seasonality and fraction of aliphatic glucosinolates, suggesting a possible role for genome size in climatic adaptation or a role for biotic interactions in shaping the evolution of genome size. We find no evidence supporting hypotheses of life history, range size or soil nutrients as forces shaping genome size in this system.
• **Conclusions** Our findings suggest climate seasonality and biotic interactions as potential forces shaping the evolution of genome size and highlight the importance of evaluating multiple factors in the context of phylogeny to understand the effect of possible selective agents on genome size.

Key words: Brassicaceae, climate, glucosinolates, plant defence, range size, seasonality, soil chemistry, *Streptanthus*.

INTRODUCTION

Angiosperm genomes range in size (C-value) over four orders of magnitude (approx. 2400-fold), from 1C = 0.061 pg (*Genlisea tuberosa* Rivadavia, Gonella & A.Fleischm., Lentibulariaceae; Fleischmann *et al.*, 2014) to 1C = 152.23 pg (*Paris japonica* Franch., Melianthaceae; Pellicer *et al.*, 2010, 2018). Although genome size is loosely correlated with ploidy and chromosome number, a general understanding of the factors associated with genome size variation is still limited. Ancestral angiosperms are thought to have had small genomes (2C approx.1.4 pg), and there have been multiple bouts of both increases and decreases in genome size across angiosperm evolution (reviewed in Soltis and Soltis, 2013).

One hypothesis of genome size evolution suggests inexorable genome bloat, due to an ever-increasing baseline after multiple rounds of polyploidization (Wendel, 2015), combined with gene duplication and replication of non-coding DNA (Soltis *et al.*, 2003; Leitch and Bennett, 2004). However, several mechanisms of genome downsizing act against a one-way increase in genome size (Leitch and Bennett, 2004; Wendel *et al.*, 2016; Zenil-Ferguson *et al.*, 2016; Simonin and Roddy, 2018). In rice and maize, close to half of the genes have been lost since the last round of polyploidization (Messing *et al.*, 2004; Wang *et al.*, 2005); numerous lineages whose evolutionary history is known to include multiple rounds of polyploidization and gene duplication have rather small genomes [e.g. *Arabidopsis thaliana* (L.) Heynh., 1C = 0.16 pg; Bennett *et al.*, 2003; Fawcett *et al.*, 2013]; similarly, plants that have among the largest genome sizes may have diploid karyotypes (e.g. *Fritillaria* L.; Kelly *et al.*, 2015). Thus, loss and acquisition of large chromosomal

© The Author(s) 2021. Published by Oxford University Press on behalf of the Annals of Botany Company.
All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

segments or smaller non-coding regions (e.g. transposable elements; Kalendar *et al.*, 2000; Piegu *et al.*, 2006; Wendel *et al.*, 2016), coupled with gene gain/loss and polyploidy can lead to contrasting trends in genome size over time, even within a clade (e.g. Wendel *et al.*, 2002; McIntyre, 2012; Kelly *et al.*, 2015; but see Clark *et al.*, 2016). Directional changes in genome size over time suggest that some factors are exerting selection directly or indirectly on genome size (Pellicer *et al.*, 2018; Simonin and Roddy, 2018; Leitch *et al.*, 2019).

Numerous hypotheses have been put forward suggesting autecological and environmental factors favouring small vs. large genomes (e.g. Grime and Mowforth, 1982; Macgillivray and Grime, 1995; Wendel *et al.*, 2002). A challenge in understanding forces shaping genome size evolution in plants is that many hypothesized selective agents are intercorrelated. For example, one classic hypothesis is that genome size is negatively related to overall plant growth rate, with large genomes limiting the ability of plants to grow rapidly (Mowforth and Grime, 1989; Lavergne *et al.*, 2010). Variation in growth rate may also reflect cell size (Niklas, 1994), seed size (Turnbull *et al.*, 2012), perenniality/annuality (Grime and Hunt, 1975; Garnier, 1992) and climate (for references, see Table 1), and each of these traits has separately been proposed as a correlate of genome size evolution (see Table 2 for a brief summary on potential drivers and mechanisms of genome size evolution). More specifically, across 23 *Crepis* L. species, the smallest genomes are found in annual species and in areas with less seasonal climates (Enke *et al.*, 2011); in *Artemisia* L., large genomes are found in woody perennials in arid environments (Torrell and Vallès, 2001), and in wild barley (*Hordeum spontaneum*), larger genomes are associated with drier and higher elevation environments, possibly mediated by transposons (Kalendar *et al.*, 2000). Thus, the intercorrelation of properties associated with growth rate and genome size clouds the role of each factor.

Other factors have also been considered drivers of genome size evolution. Soil nutrient availability, specifically of nitrogen (N) and phosphorus (P), and water availability (Castro-Jimenez *et al.*, 1989) have been proposed to restrict genome size through increased costs of nucleic acid synthesis (Hanson *et al.*, 2001; Leitch and Leitch, 2008; Guignard *et al.*, 2016). Geographic range and population size have also been proposed as factors related to genome size. Reduction in genome size through purging of deleterious and extraneous material from the genome is thought to be more efficient in populations with a large effective population size (N_e) (Lynch and Conery, 2003; but see Whitney *et al.*, 2010). Range size has also been shown to be positively related to large N_e; in 71 of 88 species, local abundance and geographic range were significantly positively correlated (Gaston, 1996). Thus, one would predict that smaller genomes should be found in species with larger ranges, where large N_e facilitates purging of extraneous material from the genome. Smaller genomes in species with larger ranges is supported by several studies that have indirectly addressed range size through invasive status and rarity (Vinogradov, 2003; Grotkopp *et al.*, 2004; Lavergne *et al.*, 2010; Pandit *et al.*, 2014).

The role of biotic interactions in genome size evolution has received much less attention than abiotic factors, despite the fact that interactions with pathogens, herbivores and mutualists can be crucially important to plant fitness and diversification rates (e.g. Ehrlich and Raven, 1964; Janz, 2011). A recent study showed that herbivores graze differentially on plants with larger genomes, establishing that genome size can play a role in influencing plant–herbivore interactions (Guignard *et al.*, 2019). A study on the evolution of glucosinolate synthesis (secondary compounds in Brassicales that can serve as a defence against herbivores) showed that the escalation of glucosinolate diversity occurred through both single-gene and whole-genome duplications, with the retention and neofunctionalization of core biosynthesis genes (Edger *et al.*, 2015). Thus, we hypothesize that there could be a positive relationship between genome size and the diversity or complexity of chemical defences against enemies, if gene and genome duplications are key in ongoing reciprocal arms race evolution. In support of this hypothesis, in tunicate bacterial symbionts, the biosynthetic pathways of toxic secondary metabolites are selectively retained, despite reductions in other genomic regions (Kwan *et al.*, 2012).

In sum, while significant progress has been made in identifying correlates of genome size and putative selective agents (see summary in Table 1), we are still far from a general consensus on factors correlated with genome size (Pellicer *et al.*, 2018). Here, we take advantage of our extensive knowledge of a clade of mustards (*Streptanthus* and close relatives) to identify trends and evaluate correlates of genome size evolution in this group. We contribute genome size estimates for 47 species of *Streptanthus* Nutt. and close relatives (Brassicaceae), and explore how genome size has evolved across the clade. We then relate changes in genome size across the clade to extensive datasets on species' life history traits, range size and to multivariate datasets of climate, soil characteristics and secondary chemistry for each species. We use data reduction and information criteria approaches to select the best models to explain genome size variation across the clade while taking into account intercorrelations among factors. We discuss our results in light of the *a priori* hypotheses above (Table 1).

MATERIALS AND METHODS

The system

Streptanthus Nutt. and close relatives in genera *Caulanthus* Watson and previously recognized *Guillenia* Greene have their centre of diversity in the California Floristic Province (CFP). They occupy a wide range of climates, habitats and soils, from the Mojave and Sonoran deserts in the south-west to wetter and high serpentine slopes in southern Oregon, from sea level near the California coast, to >3000 m in the Sierra Nevada, as well as continental habitats in New Mexico, Texas, Kansas, Nevada, Colorado and Louisiana (Al-Shehbaz, 2010; Baldwin *et al.*, 2012; Cacho *et al.*, 2014). There have been 4–5 independent origins of serpentine soil use, and about a third of the species show some degree of specialization to harsh serpentine soils (Cacho and Strauss, 2014). Species of *Streptanthus* are attacked by a diversity of herbivores, including vertebrates, aphids, beetles and specialist pierid butterflies (Shapiro, 1981; Strauss and Cacho, 2013), and, like other mustards, produce glucosinolates, which have been shown to play important roles in interactions with herbivores and pathogens, as well as in responses to

TABLE 1. *Brief overview of mechanisms, correlates and putative selective agents that have been hypothesized to relate to genome size (GS) in plants*

A. Correlate or putative selective agent on GS	Predicted relationship	Rationale	Relationship of factor in column A with genome size in tests of hypothesis
Mechanisms other than polyploidy that can drive genome size variation			
Variation in GC content	Quadratic, mixed	Tension between increases in GC content [possibly mediated by transposable element (TE) accumulation] and selection against high GC content in large genomes.	Parabola opening downward: Veselý *et al.* (2012); Šmarda *et al.* (2014). Parabola opening upward: Trávníček *et al.* (2019).
Endoreduplication	Unclear	Potential flexibility of developmental regulation of cellular DNA content allowing organisms to increase DNA content in somatic cells. More research needed in this area.	Positive: Trávníček *et al.* (2019).
Effective population size	Negative	Populations with large N_e are expected to be more efficient at purging extraneous DNA (deleterious mutations, duplication, haplotypes). Large N_e reduces genome size.	Negative: Lynch and Conery (2003). None: Whitney *et al.* (2010).
TEs or other repeats	Positive	The copy number of TEs and other repeats drives genome size.	Positive: Kalendar *et al.* (2000); Bilinski *et al.* (2018).
Developmental correlates			
Growth rate/cell production	Negative	DNA replication is costly and time-consuming, limiting growth rate.	Negative: Mowforth and Grime (1989); Lavergne *et al.* (2010); Bilinski *et al.* (2018); Simonin and Roddy (2018); Qiu *et al.* (2019). Positive: Müller *et al.* (2019).
Cell/seed size	Positive	Large cells are required to host large genomes; rapid cell expansion in polyploids.	Positive: Mowforth and Grime (1989); Thompson (1990); Wakamiya *et al.* (1993); Knight and Ackerly (2002); Grotkopp *et al.* (2004); Beaulieu *et al.* (2007); Knight and Beaulieu (2008); Bilinski *et al.* (2018); Simonin and Roddy (2018).
Life history (annuality)	Negative	Small genomes are expected in annuals (large GS would limit rapid growth).	No relationship: Grime and Mowforth (1982); Knight and Ackerly (2002). Negative: Price and Bachmann (1975); Watanabe *et al.* (1999); Enke *et al.* (2011); Qiu *et al.* (2019).
Ecological correlates			
Latitude	Mixed	Complex relationships through intercorrelated climatic variables, biotic interactions, life history or growth form. Largest genomes excluded from extreme latitudes.	Positive: Bottini *et al.* (2000); Souza *et al.* (2019). Negative: Grime and Mowforth (1982); Grotkopp *et al.* (2004). Mixed: Díez *et al.* (2013). Reviewed in: Knight and Ackerly (2002); Knight *et al.* (2005).
Elevation	Mixed, unimodal	Possibly driven by non-linear relationships with climatic variables such as temperature and precipitation, as well as biotic interactions and growth form, but generally the largest genomes are excluded from extreme elevations.	Negative: Bottini *et al.* (2000); Dušková *et al.* (2010); Díez *et al.* (2013); Bilinski *et al.* (2018). Positive: Caceres *et al.* (1998); Cerbah *et al.* (1999). Mixed evidence: Suda *et al.* (2003); reviewed in Knight *et al.* (2005).
Range size	Negative	Smaller genomes expected in plants with larger range sizes. Range size may reflect larger effective population size, leading to more efficient selection toward smaller genome sizes.	Positive (indirect): Vinogradov (2003) (threatened species smaller genomes). Tests are indirect, through conservation status or rarity (see below).
Colonization of new habitats or ability to expand range size	Negative	Larger genomes could present constraints on the ability to colonize novel areas, indirectly through other traits (e.g. greater growth rate or dispersal, shorter generation times).	Negative: Grotkopp *et al.* (2004); Lavergne *et al.* (2010); Pandit *et al.* (2014); Schmidt *et al.* (2017).
Secondary chemistry	Positive	Increased secondary chemistry diversity through gene or genome duplications (this study).	Positive: Kwan *et al.* (2012) (bacterial genome size); Edger *et al.* (2015).
Climatic temperature	Mixed	Smaller genomes allow for fast division and thus are expected in higher temperatures (assuming water availability). Relationship is complex because of correlations with life history, growth form and precipitation.	Negative: Thompson (1990); Wakamiya *et al.* (1993); McIntyre (2012); Souza *et al.* (2019). Positive: Suda *et al.* (2003). Mixed: Díez *et al.* (2013) Quadratic: Knight and Ackerly (2002).

TABLE 1. *Continued*

A. Correlate or putative selective agent on GS	Predicted relationship	Rationale	Relationship of factor in column A with genome size in tests of hypothesis
Climatic precipitation	Mixed	Relationship between genome size and precipitation is also correlated with life history, cell size, growth form and temperature.	Negative: Wakamiya *et al.* (1993); Bottini *et al.* (2000); Knight and Ackerly (2002); Suda *et al.* (2003); Grotkopp *et al.* (2004); Souza *et al.* (2019).
		Smaller cells (with less DNA) can better maintain turgor pressure at low water potentials; transposon copy number tends to be higher in dry habitats.	Positive: Price *et al.* (1981); Kalendar *et al.* (2000) (indirectly through aridity).
			Mixed: Díez *et al.* (2013).
			Quadratic: Knight and Ackerly (2002).
Climatic seasonality	Mixed	Related to growth strategy and coupling of temperature and precipitation. Seasonal environments where temperature and precipitation are decoupled can favour growth through cell expansion, and thus larger genomes.	Positive: Grime and Mowforth (1982) (larger genomes in seasonal environments); Enke *et al.* (2011) (smaller genomes in less seasonal climates); Díez *et al.* (2013).
			Negative: Qiu *et al.* (2019) (smaller genomes in more seasonal environments).
Soil P and N	Positive	DNA synthesis requires great amounts of N and P.	Positive: Hanson *et al.* (2001); Šmarda *et al.* (2013); Guignard *et al.* (2016).
Community composition	Mixed	GS affects plant–herbivore interactions which in turn affect community composition.	Guignard *et al.* (2019): GS as a functional trait that through interactions with herbivory and nutrient input explains differences in community composition.
Herbivory	Mixed	Herbivory may increase on plants with larger GS because at, the cellular level, they have higher N and P content and also relatively less cell wall material; also herbivory may decline on plants with larger genomes because of more complex or abundant secondary compounds.	Positive: Guignard *et al.* (2019).

The role of selective agents on GS can be complex, as the role of one factor may be impossible to distinguish from the influence of others that are correlated (e.g. life history or elevation and climate are intercorrelated and all are proposed to be related to genome size). Studies may address the relationship of multiple factors to GS, and can appear more than once in this table. Factors in bold are addressed in this study.

TABLE 2. *Genome size in relation to selected* a priori *autecological, microhabitat, climatic and defence traits, taking phylogenetic relationships into account*

Variable Category	Variable	Estimate	R^2	PGLS *P*-value*	lambda optimal	lambda lower
Autecological	Annuality	0.0074	–0.022	0.846	0.9998	0.8155
	Range size (log)	–0.0028	–0.021	0.792	0. 9999	0.8229
Harsh environment	Soil affinity (logit prop. serpentine)	–0.0051	–0.014	0.549	1.0000	0.8099
	Microhabitat bareness (asin)	–0.0937	–0.007	0.399	0.9999	0.7293
	Soils PC1	–0.0028	–0.026	0.757	1.0000	0.7525
	Soils PC2	0.0019	–0.027	0.823	1.0000	0.7730
	Soils PC3	–0.0004	–0.029	0.954	1.0000	0.7723
	Soil N (log N)	0.0203	–0.015	0.494	1.0000	0.7728
	Soil P (log Olsen P)	0.0009	–0.029	0.934	1.0000	0.7701
Climate	**Climate PC1**	**–0.0206**	**0.106**	**0.017***	**0.9995**	**0.7823**
	Climate PC2	0.0011	–0.022	0.868	0.9998	0.8200
	Climate PC3	–0.0036	–0.019	0.696	1.0000	0.8190
Investment in defence (glucosinolate production)	Total glucosinolate amount	0.0049	–0.009	0.424	0.9999	0.7816
	Glucosinolate richness	**0.0299**	**0.055**	**0.090•**	**0.9998**	**0.7913**
	Fraction of aliphatics	**–0.1181**	**0.067**	**0.069•**	**1.0000**	**0.7855**

Variables were selected based on hypotheses suggested in the literature (Table 1) and prior work. Phylogenetic generalized least squares analyses were performed with the function 'pgls' (R package 'Caper'; Orme *et al.*, 2013) across 1000 posterior trees (Cacho *et al.*, 2014). In bold are all significant relationships, marginal ones ($P_{PGLS} \leq 0.1$) are highlighted with a dot (•) and significance at ($P \leq 0.05$) is indicated by an asterisk (*). Bonferroni correction yields all *P*-values non-significant.

abiotic stress (Kliebenstein *et al.*, 2005; Hopkins *et al.*, 2009; del Carmen Martínez-Ballesta *et al.*, 2013).

Genome size estimates

We estimated 2C genome size for a total of 173 individuals (representing 114 populations) of 47 species (n = 1–4 populations per species, and an average of 3.5 individuals per species) of *Streptanthus* and close relatives (representing nearly 90 % of the species in the '*Streptanthus* complex' *sensu* Cacho *et al.*, 2014; see Fig. 1) using flow cytometry (FCM) as outlined by Dolezel *et al.* (2007) and following methods described by McIntyre (2012). To prepare nuclear suspensions for FCM, we ground 3–6 seeds from an individual together with seeds from one of several standards using two pieces of fine grit sandpaper

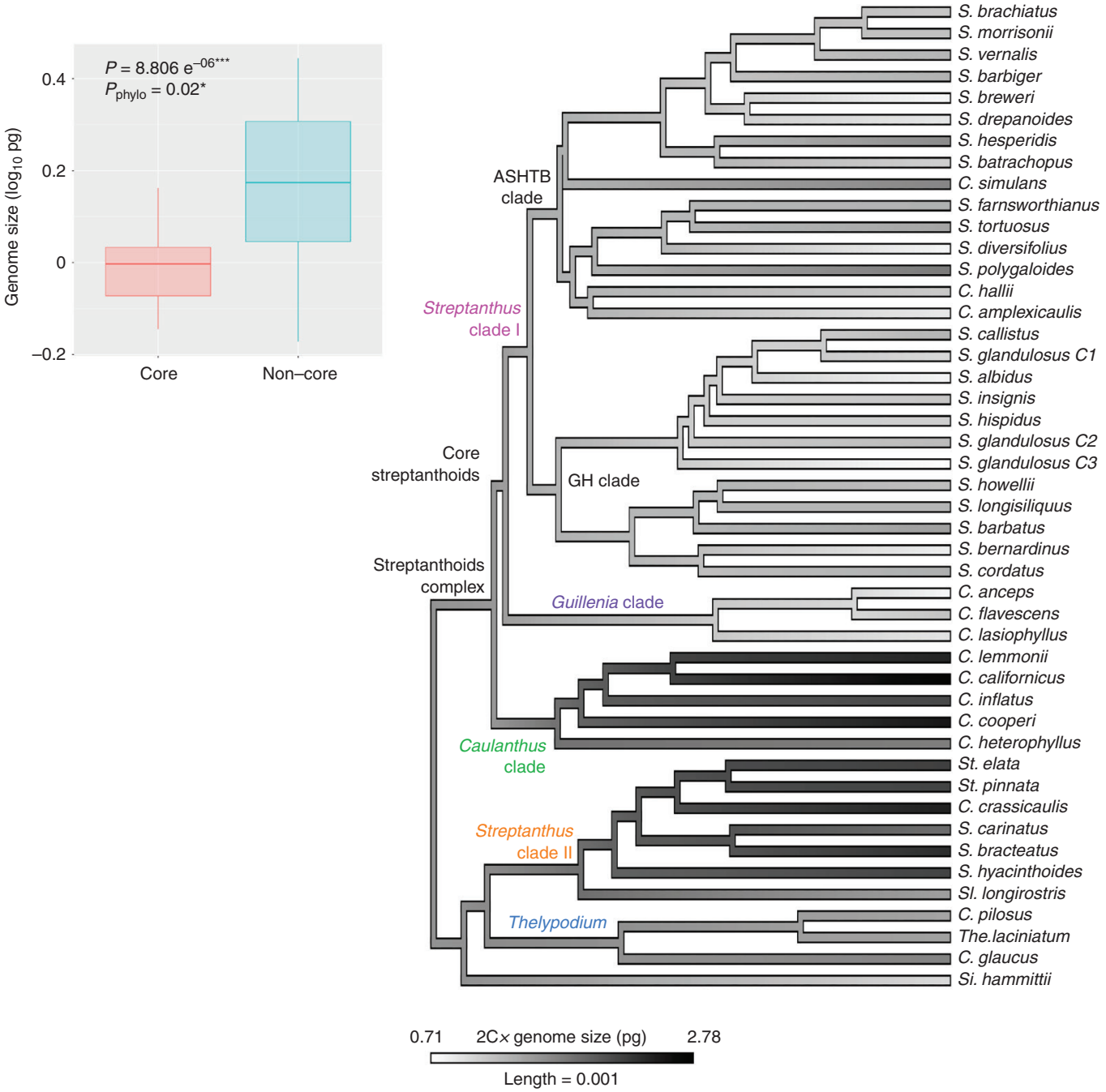

FIG. 1. Genome size in *Streptanthus* and close relatives, adjusted for recent polyploidy (log 2Cx values), mapped onto the maximum credibility tree of a 50 million generation Bayesian analysis (Cacho *et al.*, 2014) using function 'contmap' from R library 'phytools' (Revell, 2012). A phylogenetic ANOVA (inset) supports smaller genomes for core streptanthoids (*Streptanthus* clade I + *Guillenia* clade) compared with the rest of the species (function 'aov.phylo', 'geiger' package, Harmon *et al.*, 2008). For genome size estimates and chromosome counts, see Appendix 1. For phylogenetic signal estimates, see Supplementary data Table S1. Genera are abbreviated as follows: C = *Caulanthus*, S = *Streptanthus*, Si = *Sibaropsis*, Sl = *Streptanthella*, St = *Stanleya,* The = *Thyelypodium*.

(27.9 grit cm^{-2}) in a 60 mm diameter Petri plate. As a general standard, we used 20–30 seeds of diploid *Arabidopsis thaliana* Columbia (line 35 from Comai Lab stock http://comailab.genomecenter.ucdavis.edu; referred to here as Ath2C35). We also performed measurements using alternative standards, such as a diploid–tetraploid *A. thaliana* mixture (20–30 seeds), or *Solanum lycopersicum* L. cultivar Stupicke (2–3 seeds; 2C genome size = 1.96 pg) acquired from the Dolezel lab (Dolezel *et al.*, 2007). The Ath2C35 standard had a 2C genome size of 0.32 pg based on comparison with *S. lycopersicum*, which is within the typical reported genome size range for *A. thaliana* accessions (Bennett *et al.*, 2003; Schmuths *et al.*, 2004). The 2C genome size of *A. thaliana* from the 2*x*–4*x* mixture genome size was slightly larger than the Ath2C35 standard, at 0.352 pg (2*x*) and 0.704 pg (4*x*), again within the range of variation reported for *A. thaliana* accessions.

After co-grinding a given sample with the standard, we washed the sandpaper and plate with two 650 µL aliquots of Galbraith's buffer (Dolezel *et al.*, 2007), and filtered the suspension through two layers of Miracloth (CalBiochem, Pasadena, CA, USA). We then transferred 600 µL of the filtered solution to an ice-cold 5 mL falcon tube (BD Falcon, cat. no. 352008), stained with propidium iodide (50 µg mL^{-1}), treated with RNase (50 µg mL^{-1}) and kept on ice and protected from light until FCM analysis (30–60 min).

We carried out our FCM measurements with a Becton Dickinson FACScan flow cytometer equipped with a 488 nm argon laser, which is maintained at the UC Davis Flow Cytometry Shared Resource Laboratories (https://ccresources.ucdmc.ucdavis.edu/). We identified peak means with ModFit LT v.4.1 software (Verity Software House, Topsham, ME, USA). For analyses, we retained samples with a coefficient of variation <5 and at least 5000 nuclei (excluding potential debris).

We used seeds rather than fresh leaf tissues as sources for our measurements because preliminary exploration of genome size estimates in *Streptanthus* suggested that leaf tissue, which can be strongly pigmented in some species of this clade (Strauss and Cacho, 2013), resulted in peak location shifts relative to internal standards, a shift that was not observed in measurements derived from seeds. Others have also used seeds successfully as sources for genome size estimates (Sliwinska *et al.*, 2005; Jedrzejczyk and Sliwinska, 2010; McIntyre, 2012). Our genome size measurements based on seeds were comparable with available genome size estimates for closely related Brassicaceae, and were tightly correlated with independently reported genome size measurements for *Streptanthus* based on leaf tissue ($r = 0.99$; for references and details, see Supplementary Methods S1). No additional peaks that could be related to endopolyploidy or endosperm were observed in seed tissue. We compiled chromosome counts for all species, first from counts reported in the primary literature, and then relying on secondary sources including floras (Rollins, 1993; Al-Shehbaz, 2010; Baldwin *et al.*, 2012) and a chromosome count database for the Brassicaceae (Warwick and Al-Shehbaz, 2006). Members of *Streptanthus* have a base chromosome number of $x = 14$ (Al-Shehbaz, 2010; Baldwin *et al.*, 2012). Neopolyploidy is uncommon within this clade, with both diploid and tetraploid chromosome counts reported for only three of the 47 taxa in this study (Appendix 1). Of these three species, we treated our accessions of *Stanleya pinnata* Britton and *Streptanthus barbatus* S. Watson as tetraploids, as our estimates of 2C genome sizes for these two species were approximately twice as large as estimates for congeners. Also, our *S. pinnata* accession was from a region where tetraploids, but not diploids, are known (Cappa *et al.*, 2014). We treated our *S. cordatus* Nutt. accession as diploid, as the measured genome size was within the typical range of diploid *Streptanthus*.

We examined genome size evolution and its correlates in this clade using 2C*x* genome size estimates (Greilhuber *et al.*, 2005) in order to correct for recent polyploidy (Leitch and Bennet, 2004; Whitney *et al.*, 2010). This approach was chosen to avoid spurious correlations driven by one or a few points in the two species that seemed to correspond to tetraploid accessions (*Stanleya pinnata* and *Streptanthus barbatus*; see above). Species genome size estimates were log transformed, a practice recommended for studying the evolution of traits such as genome size (O'Meara *et al.*, 2006; Oliver *et al.*, 2007; Lysak *et al.*, 2009).

All analyses followed the most recent phylogeny of this group (Cacho *et al.*, 2014). Our dataset for genome size evolution in *Streptanthus* and close relatives consisted of 46 tips for which we had both genome size estimates and a phylogenetic framework (there is no molecular phylogeny that informs the relationships of *Streptanthus gracilis* Eastwood, *Streptanthus oblanceolatus* T.W.Nelson & J.P.Nelson and *Streptanthus oliganthus* Rollins).

Potential correlates of genome size evolution in streptanthoids: expectations and data collection

We focus on traits that have been previously suggested as potential drivers of genome size, or that we have found to play an important role in the evolutionary ecology of this clade.

Life history. If larger genomes interfere with rapid cell division and growth (Table 1), we expected small genome sizes to be favoured in streptanthoids with an annual life history, allowing them to capitalize on erratic rainfall events typical of the Mediterranean climate of the CFP. Life history (annual/non-annual) data were compiled from the Flora of North America (Al-Shehbaz, 2010), the Jepson Manual of California Plants (Baldwin *et al.*, 2012) and from our own observations for all species.

Range size. We predicted that species with larger ranges should have smaller genomes – if range size can be considered a proxy for N_e (Brown, 1984; Johnson, 1998; Borregaard and Rahbek, 2010; Table 1). For all species represented in our phylogeny (46 tips), we downloaded georeferenced occurrence records from the Global Biodiversity Information Facility (http://www.gbif.org) using the R package 'dismo' (Hijmans *et al.*, 2013) and screened records for errors (e.g. incorrect taxonomy, out of documented range, occurring in botanic gardens and truncated co-ordinates). The total number of curated unique locality records was 5970 (per species mean = 106.6, median = 49, min = 1 for two microendemics). To estimate the geographic range of each species, we placed a buffer radius of 10 km around each record location ('gBuffer' function in the 'Rgeos' library;

Bivand and Rundel, 2013) and then merged overlapping areas to avoid double counting ('joinPolys' function in the 'PBSmapping' R library). We used this buffer method rather than convex hulls bounding the full set of localities because it is less likely to include large amounts of unsuitable habitat (see also Nakazato *et al.*, 2010; Weber *et al.*, 2018).

Harsh edaphic environment. Because low nutrient availability might select against large genomes (see Table 1), we expected genome size to be smaller in streptanthoids occupying nutrient-poor (low N and P) habitats, such as serpentine soils characterized by low P (Cacho and Strauss 2014). We test this hypothesis using both a serpentine soil affinity metric and chemical analyses of soil samples from microsites occupied by plants in the wild. Serpentine affinity for each species was estimated as the proportion of collection records on serpentine soil (calculated by overlaying collection records onto a GIS layer of serpentine soil using R functions in the package 'dismo'). Field soil samples were collected from the root zone – the top 30 cm of soil – from randomly selected plants in the field ($n = 3$ plants per population), with replicate populations for most species. Samples were analysed at the University of California, Davis Analytical Laboratory (www.anlab.ucdavis.edu); for complete methods and data, see Cacho and Strauss (2014). Soil composition–genome size relationships were estimated for the sub-set of 36 species for which we had genome size, soil chemistry data and phylogeny.

Our previous work identified the amount of bare rock and soil surrounding plants (microhabitat 'bareness') as an ecologically important and phylogenetically conserved trait in this clade, one that integrates over a suite of selective pressures such as drought, UV exposure, nutrient availability, competition and vulnerability to herbivores (Strauss and Cacho, 2013; Cacho and Strauss, 2014). We estimated microhabitat bareness as the percentage of bare ground or rock in a 25 cm^2 quadrat centred on a *Streptanthus* plant in the field, for 5–15 individuals per population with multiple populations per species (for details on protocols and data, see Cacho and Strauss, 2014). We related genome size to bareness for 42 species for which we had these data.

Climate. The long-standing hypothesis of Grime and Mowforth (1982) of climate seasonality, favouring larger genomes mediated by the decoupling of cell expansion from cell division to achieve fast growth, has received mixed support across a range of plant groups (Castro-Jimenez *et al.*, 1989; Enke *et al.*, 2011; Pellicer *et al.*, 2018; Qiu *et al.*, 2019; see Table 1). We asked if genome size was positively related to climate seasonality by extracting the 19 standard Bioclim variables available in Worldclim v.2 (Fick and Hijmans 2017), which we downloaded for each of our unique georeferenced locality records at the 30 arc second scale (approx. 1 km at the equator) for our 47 focal species. We averaged each bioclimatic variable across localities for every species.

Plant secondary chemistry. Production of chemical defences such as glucosinolates often involves duplicated genes or is facilitated by genome duplication (Edger *et al.*, 2015), and thus chemical defence could be expected to be more complex in species with larger genomes. From an earlier study of this clade, the dominant glucosinolates were found to be those derived from methionine (aliphatic glucosinolates), and the fraction of total glucosinolates that are aliphatic has increased over evolutionary time across the clade (Cacho *et al.*, 2015). Also, glucosinolate richness (number of glucosinolate compounds produced by a species) correlates with soil specialization and microhabitat bareness, suggesting selection for increased defence in bare environments (Strauss and Cacho, 2013; Cacho *et al.*, 2015). Thus, we predicted that larger genomes would be associated with greater investment in defence in this clade of plants specialized in bare environments, where vulnerability to herbivores is increased (Strauss and Cacho, 2013; Strauss *et al.*, 2015).

Leaf samples for glucosinolate analyses were collected in the field from undamaged rosette or cauline leaves for 1–5 individuals per population and replicate populations per species when possible (mean number of populations per species = 1.7). For details on protocols and data, see Cacho *et al.* (2015). In brief, leaf tissue for 37 species was promptly placed in 90 % methanol to inactivate the myrosinase enzyme, and an equal amount of tissue from the same sample was dried for biomass quantification; samples were stored at –20 °C until extraction. Desulfoglucosinolates were analysed with high-performance liquid chromatography (HPLC) using a diode array detector as outlined elsewhere (Kliebenstein *et al.*, 2001, 2005). Identification of compounds was based on peak retention times and absorbance spectra, and concentrations were estimated from areas under the curves using published response factors (Reichelt *et al.*, 2002) and normalized by dry weight of the sample (see table S1 of Cacho *et al.*, 2015). Total glucosinolate production was calculated as the sum of the amounts produced of all glucosinolate compounds per plant; the fraction of aliphatic glucosinolates was calculated by dividing the sum of aliphatic glucosinolates per plant over total glucosinolate production; and glucosinolate richness was calculated as the number of unique compounds produced by a given species.

Multivariate datasets. Our climate, soils and glucosinolate chemistry data are multidimensional, with 19, 27 and 28 variables, respectively, many of which are intercorrelated. To reduce the dimensionality and complexity of these data sets, we used principal components analysis (PCA; 'princomp' function, option corr = TRUE) prior to phylogenetic informed analyses, an approach that has been used by others when evaluating correlates of genome size in a phylogenetic context (e.g. Trávníček *et al.*, 2019). For soils and climate data, where PCA was effective in reducing the dimensionality of our datasets, we used the first three PCs (that explained at least 60 % cumulative variation) to explore correlations with genome size using phylogenetic generalized least squares (PGLS; see below). When significant correlations ($P < 0.1$) were detected with any of the three PC axes, we identified relevant variables by examining PC loadings on those axes before proceeding with model selection analyses (see below).

Statistical analyses

Phylogenetic uncertainty. We integrated over uncertainty in phylogenetic inference by implementing analyses across

a random sample of trees from the posterior distribution of a 50 million generation Bayesian analysis (Cacho *et al.*, 2014); when a single tree was needed, we used the maximum clade credibility tree (MCCT). All analyses were done in R (R Development Core Team, 2018).

Evolution of genome size. We estimated phylogenetic signal in genome size using Blomberg's *K* (Blomberg *et al.*, 2003), a standard measure of phylogenetic signal for continuous characters. We tested significance in *K* being different from zero (when there is a lack of correspondence between a trait's evolution and the phylogeny) by comparing the empirical *K* values with values derived for null models that randomly shuffle taxa across 1000 trees randomly selected from the posterior with the function 'phylosignal' from the R package 'picante' v.1.6 (Kembel *et al.*, 2010; Cacho *et al.*, 2014).

To investigate potential divergent modes of evolution on genome size across clades of *Streptanthus* and relatives, we tested seven models of evolution that parameterize rate, optima and selection in relation to the phylogeny ('core' streptanthoids composed of species in *Streptanthus* clade I + *Guillenia* clade vs. the rest of the species; Fig. 1). The two simpler models we fit assume no selection on genome size (Brownian motion models): one assumes a single rate (BM1) across all species, and the other assumes a different rate in the 'core' streptanthoids as defined above (BMS) compared with the rest of the streptanthoids. The next two models in complexity invoke a selective regime for genome size evolution (Orstein–Uhlenbeck models) that consist of a single rate (σ) and strength (α) but vary in the number of optima they allow: either a single optimum (OU1) or multiple optima (OUM). The final three models, also OUM models, allow clade-dependent rates of evolution (OUMA), strengths of selection (OUMV) or both (OUMVA). To implement models, we used the function 'ouwie' (specifying the clade option, and root.station set to TRUE; Beaulieu *et al.*, 2012) across 500 trees from the posterior. We made sure the objective function was being estimated adequately by ensuring all values of the Hessian eigen decomposition were positive before proceeding to summarize the results (4.5 % of implemented models were removed). We used Akaike information criteria (AICs) to identify models that best explain genome evolution across the clade, while minimizing loss of information (Burnham and Anderson, 2002).

Phylogenetically informed approach to assess correlates of genome size. We used PGLS (Grafen, 1989) to evaluate the following 15 potential correlates of genome size: annuality (treated as a factor), range size, serpentine affinity, microhabitat bareness, soil N, soil P, soils PC1, PC2 and PC3, climate PC1, PC2 and PC3, total glucosinolate amount, glucosinolate richness and the fraction of glucosinolate that are aliphatic. When needed, variables were transformed to improve normality (see Table 2) based on Shapiro tests and visual inspection of the data. Also, we ensured that the PGLS residuals did not violate the assumption of normality for any of our models. The PGLS models were run using the R function 'pgls' (Orme *et al.*, 2013), using maximum likelihood to estimate lambda across a set of 1000 trees chosen at random from the posterior. Despite being widely viewed as too conservative, especially in light of *a priori* hypotheses in Table 1 (Streiner, 2015), we include a multiple comparisons Bonferroni correction.

AIC-based phylogenetic model selection approach. To identify variables that, while intercorrelated, could still account for variation in genome size, we adopted a modified information criteria model selection approach based on variables that showed correlations ($P < 0.1$) with genome size in our PGLS analyses. We included the following five variables in our initial full model: temperature seasonality (bioclim 4), temperature annual range (bioclim 7), precipitation seasonality (bioclim 15), glucosinolate richness and fraction of glucosinolates that are aliphatic. To identify the best models, we used the function 'phylostep' (option = lambda) from the R package 'phylolm' (Tung Ho and Ané, 2014) and the maximum clade credibility tree of Cacho *et al.* (2014).

RESULTS

Genome size in Streptanthus *and relatives*

Genome size (raw 2C values) in *Streptanthus* and relatives varied from 0.71 pg (*S. glandulosus*) to 3.89 pg (*Stanleya pinnata*; Fig. 1; Appendix1). The DNA content in *S. barbatus* (2.36 pg) and *St. pinnata* (3.89 pg) corresponds roughly to twice the amount of DNA present in their congeners and close relatives, and recent tetraploidy has been reported for these species based on chromosome counts (Baldwin *et al.*, 2012; Appendix 1).

Genome size evolution in Streptanthus *and relatives*

Genome size has a strong phylogenetic signal in *Streptanthus* and relatives ($K = 1.6$, $P = 0.001$); mean *K*-values and their corresponding *P*-values are reported in Supplementary data Table S1. The ancestral 2C*x* genome size for streptanthoids was estimated to be 1.28 pg, with variation that includes larger values in some subclades (*Streptanthus* clade II, 1.54 pg) and smaller values in others (*Guillenia* clade, 0.95 pg; Fig. 1; Supplementary data Table S2). A phylogenetic analysis of variance (ANOVA) supports core streptanthoids (*Streptanthus* clade I + Guillenia clade) having smaller genomes than the rest of the species (phylo ANOVA $P \leq 0.002$, standard ANOVA $P \leq 0.0001$; Fig. 1).

We evaluated seven models of genome size evolution, ranging from a Brownian motion model with a single rate (BM1) to an OU model allowing optima and strengths of selection to vary across clades (OUMVA). The model best fitting our data is a BMS model (AICc = –69.88), which parameterizes the rate of evolution of genome size in 'core' streptanthoids (*Streptanthus* clade I + *Guillenia* clades) as different from the rest of the clade but differences in rate estimates are marginal (Supplementary data Table S3). The BM1, OUM and OUMA models are within two AIC scores from the BMS model. Mean values for loglikelihood, AICc and relevant parameters for each model across 500 trees are presented in Supplementary data Table S3.

Correlates of genome size evolution in streptanthoids

Autecological traits. Genome size in *Streptanthus* and relatives is not correlated with annuality or range size (Table 2).

Soil environment. We found no significant relationships between genome size and soil PC axes reflecting overall soil chemistry (for PC loadings and variance explained, see Supplementary data Table S4), specific soil nutrients N and P, soil affinity (proportion of records on serpentine) or microhabitat bareness (Table 2).

Climate. Despite variation in genome size across the major subclades of *Streptanthus* and relatives, we find a clade-wide association between larger genomes and climate seasonality. Genome size is inversely related to climatic PC1 (estimate = –0.021, $R^2 = 0.11$, $P_{PGLS} = 0.017$; Table 2; Fig. 2), which primarily reflects temperature and precipitation seasonality. We thus examined the relationship of genome size to the three variables with highest loadings on climatic PC1 (Supplementary data Table S5): temperature seasonality, temperature annual range and precipitation seasonality (Table 3; Fig. 2). We consistently find that species with larger genome sizes occupy environments with greater temperature seasonality ($P_{PGLS} = 0.031$). Conversely, smaller genomes were associated with greater precipitation seasonality ($P_{PGLS} = 0.028$; Table 3; Fig. 3), which covaries negatively with temperature seasonality; temperature annual range is marginally positively correlated with genome size ($P_{PGLS} = 0.050$; Table 3; Fig. 2). These patterns also held when only considering the 41 taxa occurring within the CFP (2C*x* log genome size for CFP taxa: range = 0.72–2.78, mean = 1.19, s.d. = 0.5; Supplementary data Table S6).

Secondary chemistry. Species with larger genomes tended to produce a lower fraction of aliphatic glucosinolates ($P_{PGLS} = 0.067$; Table 2; Fig. 2). but more types of glucosinolate compounds overall ($P_{PGLS} = 0.09$; Table 2; Fig. 2), although both relationships are marginally significant. No significant relationship between genome size and total glucosinolate amount was observed (Table 2). PCA was not effective at reducing the dimensionality of this dataset (Supplementary data Table S7).

Model selection. Our AIC-based phylogenetic model selection approach identified a best model explaining genome size variation across the group that included temperature seasonality and fraction of aliphatic glucosinolates. Larger genomes were associated with greater temperature seasonality and a smaller fraction of aliphatic glucosinolates (Table 4; for R^2, coefficients and estimates, see Supplementary data Table S8). There were three other models within two AIC units of the best model, which, in addition to temperature seasonality and aliphatic glucosinolates, included either precipitation seasonality or temperature annual range, but not both, and one model included glucosinolate richness. In these models, smaller genomes are associated with greater precipitation seasonality, larger temperature annual range and a lower richness of glucosinolate compounds (Table 4; Supplementary data Table S8).

When applying a Bonferroni correction for multiple tests across climate, soils, defence and autecological traits (15 comparisons), none of the relationships remains significant; we note these corrections do not take into account the presence of *a priori* hypotheses (summarized in Table 1).

DISCUSSION

The evolution of genome size has been explored in a number of groups, and has been linked to a large number of abiotic and autecological traits, though not necessarily in a consistent manner. The variation in genome sizes we observe among *Streptanthus* and relatives (2C: 0.71–3.89 pg, Appendix 1) is encompassed within the variation observed in the family Brassicaceae (0.3–9.47 pg; Leitch *et al.*, 2019) with a similar distribution, centred on 2C = 1.0 pg (Supplementary data Fig. S1). Genome size in *Streptanthus* varies >5.4-fold (4.2-fold in diploids), compared with a 31.6-fold (17.26-fold in diploids) variation in Brassicaceae as a whole.

Our analyses of genome size in *Streptanthus* and allies support what is arguably one consistently strong common pattern in genome size evolution: the presence of a moderate to strong phylogenetic signal (Beaulieu *et al.*, 2007; Lysak *et al.*, 2009; Bainard *et al.*, 2012; Trávníček *et al.*, 2019; but see Barrett *et al.*, 2019; Müller *et al.*, 2019). A strong phylogenetic signal in genome size has been documented in other plant clades (e.g. Liliaceae; Beaulieu *et al.*, 2007), as well as across the angiosperms (Bainard *et al.*, 2012). Functional processes and genomic attributes related to genome size, such as DNA cytosine methylation (Alonso *et al.*, 2015), endoreduplication (Bainard *et al.*, 2012), GC content (Trávníček *et al.*, 2019) and genomic repeat abundances (Dodsworth *et al.*, 2015), also a exhibit phylogenetic signal in plants.

We find that genome size differs across subclades of *Streptanthus* and relatives (Fig. 1), which is consistent with previous studies that report differences in genome size across clades of Brassicaceae (Lysak *et al.* 2009). We add evidence of clades evolving genome sizes at slightly different rates (BMS model) with equivocal support for genome size evolving towards different optima or under different selective regimes in relation to these clades. Previously, Lysak *et al.* (2009) found no evidence of directional selection operating in genome size evolution across clades of Brassicaceae, and our analyses are consistent with those findings.

Numerous hypothesized selective agents on genome size have been proposed (Table 1). Using data on ecological and life history attributes across species of *Streptanthus* and relatives, we found signals of climate and secondary chemistry associated with genome size, after accounting for evolutionary relationships among species. Despite finding significant correlates of genome size, our best models only explained approx. 20 % of the variation in genome size.

Climate seasonality

Associations between genome size and environment have commonly been reported among plants, associated either

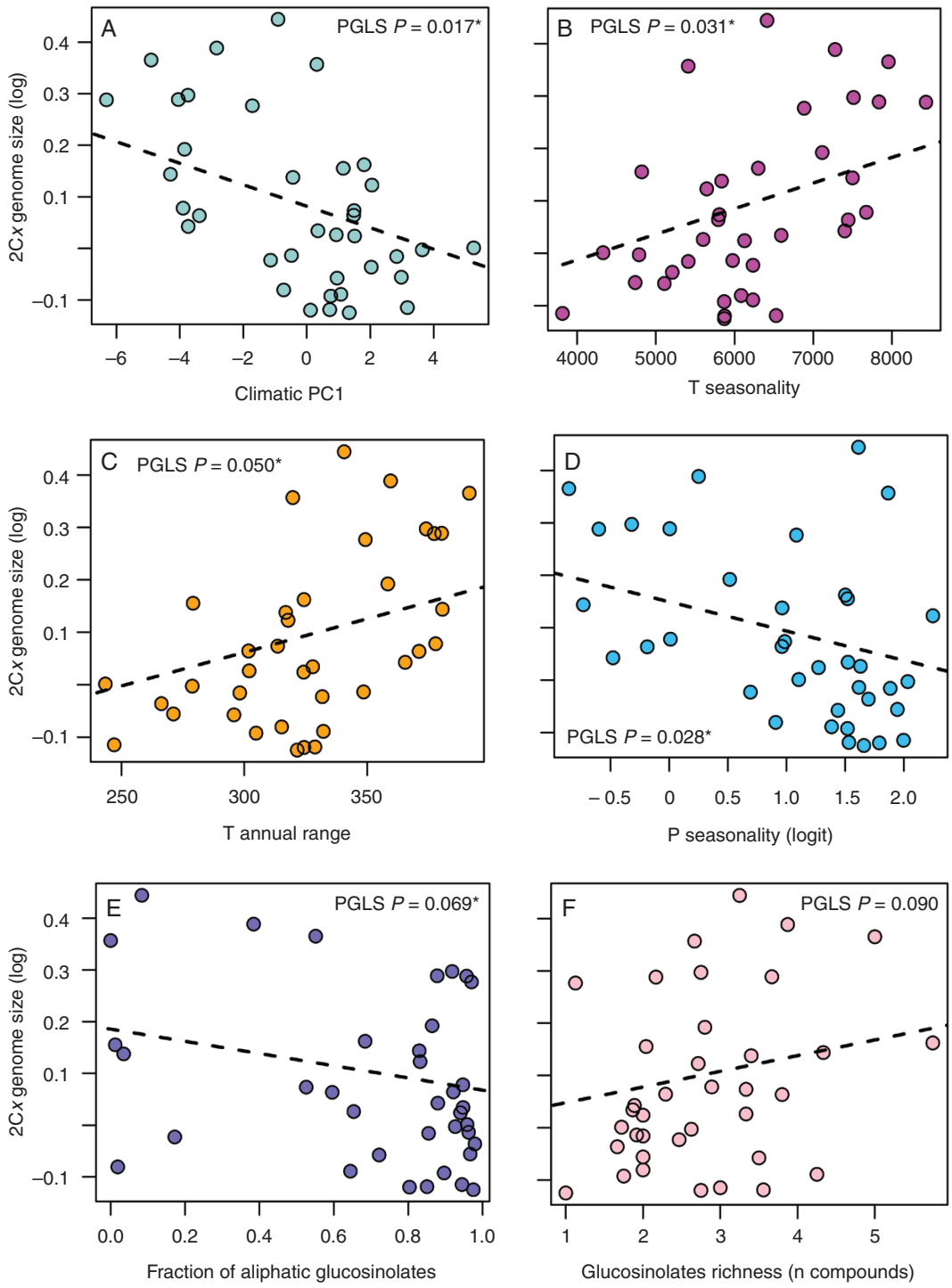

FIG. 2. Genome size in relation to investment in climatic and defence variables for *Streptanthus* and relatives. Taking phylogeny into account, genome size in this group correlates negatively with climate PC1 (A). Genome size has a positive association with temperature (T) seasonality (B) and temperature annual range (C), and a negative one with precipitation (P) seasonality (D). Smaller genomes are associated with a larger fraction of aliphatic glucosinolates (E) and also a smaller overall number of glucosinolate compounds (glucosinolate richness, F). Temperature seasonality and fraction of aliphatic glucosinolates are in all best models explaining genome size. Dotted lines represent phylogenetic generalized least squares (PGLS) model fit using the maximum clade credibility tree of a 50 million generation Bayesian analysis (Cacho *et al.*, 2014). For PGLS estimates across a sample of trees and best models, see Tables 2–4 and Supplementary data Table S8.

directly with temperature and rainfall (Grime and Mowforth, 1982; Bennett, 1987; Qiu *et al.*, 2019), or indirectly through altitude (Dušková *et al.*, 2010; Bilinski *et al.*, 2018) or latitude (Schmuths *et al.*, 2004; Knight *et al.*, 2005; Souza *et al.*, 2019). Grime and Mowforth (1982) suggested that highly seasonal environments, where conditions of temperature and

TABLE 3. *Phylogenetic generalized least squares models of genome size in relation to the three climatic variables with highest loadings on PC1 across 1000 posterior trees (Cacho et al., 2014) performed with the R function 'pgls' (package 'Caper'; Orme et al., 2013).*

Variable	estimate	R^2	*P*-val*	λ optimal	λ lower	*P*-val^B *
T seasonality (bio 4)	0.00004	0.082	**0.031***	0.9997	0.7947	**0.094 •**
T annual range (bio 7)	0.00092	0.065	**0.050***	0.9995	0.7924	0.150
P seasonality (bio 15, logit)	–0.06218	0.086	**0.028***	0.9998	0.8130	**0.084 •**

In bold are all significant relationships, marginal ones ($P_{PGLS} \leq 0.1$) are highlighted with a dot (•) and significance at ($P \leq 0.05$) is indicated by an asterisk (*). Bonferroni-corrected *P*-values are presented in the last column, *P*-val^B. PC loadings for climatic data are given in Supplementary data Table S5

precipitation relatively more favourable to cell division were limited to short windows, may generally select for rapid growth through cell expansion, temporally decoupled from cell division (Grime and Mowforth, 1982, p. 151). Their arguments built on the stronger inhibition of cell division but not cell expansion by lower temperatures, as well as on the positive association between genome size and the duration of the mitotic cell cycle (Van't Hof and Sparrow, 1963; Bennett, 1987; see also Müller *et al.*, 2019). Our results support this pattern: all of our best models include temperature seasonality, with larger genomes associated with more seasonal, continental environments, where the growing season would be shortened by sub-zero temperatures in winter and early spring. Smaller genomes are found in streptanthoid species that inhabit ocean-buffered less variable coastal climates (Fig. 3). The moderate Mediterranean climate of the CFP could represent favourable growing conditions during winter and spring, but we also find larger genomes associated with increased temperature seasonality when focusing only within CFP species. Thus, our findings are consistent with Grime and Mowforth's suggestions that larger genomes could be favoured in more temperature-seasonal environments.

Although water limitation may be the key factor for growth in Mediterranean habitats (see Table 1), such as the CFP (Baldwin 2014), precipitation seasonality was only included in our third best model explaining genome size. We found that greater precipitation seasonality was associated with smaller, not larger, genome sizes. This may be because moderate winter temperatures in coastal portions of the CFP allow for continued growth during the wet winters, such that precipitation and temperatures permitting growth are not decoupled in this region. Focus on joint patterns of seasonality in temperature and precipitation might better inform the association of genome size with climatic seasonality.

Secondary chemistry

Only recently has the importance of interactions shaping eco-evolutionary dynamics of genome size begun to be addressed (Guignard *et al.*, 2019). To our knowledge, our study is the first one explicitly looking at a relationship between genome size and plant secondary chemistry in a broad evolutionary context. We predicted that a greater diversity or amounts of defensive compounds being produced could be associated with larger genomes and, while our analyses reveal such a trend (genome size positively correlated with glucosinolate richness, $P_{PGLS} = 0.09$), this relationship was not significant when correcting for multiple comparisons. In contrast, we observed that plants with smaller genomes produced a larger fraction of aliphatic glucosinolates (derived from methionine). The fraction of aliphatic compounds produced by a species was consistently included in all our best models to explain genome size. Aliphatic glucosinolates are highly inducible (Textor and Gershenzon, 2009; Sato *et al.*, 2019), deter herbivores (Kliebenstein *et al.*, 2005; Hopkins *et al.*, 2009) and also affect tolerance of abiotic factors (Haugen *et al.*, 2008; del Carmen Martínez-Ballesta *et al.*, 2013; Salehin *et al.*, 2019). Adaptation to the abiotic environment may also indirectly reflect climatic preferences of herbivores (e.g. grasshoppers may prefer warm dry environments while slugs thrive in colder and wetter ones). In streptanthoids, the proportion of glucosinolates that is aliphatic increases over evolutionary time and is positively correlated with the occupation of nutrient-poor bare environments (Cacho *et al.*, 2015) where apparency to herbivores is higher (Strauss and Cacho, 2013). Experimental work will be required to tease apart the mechanisms underlying these correlations.

After considering many attributes hypothesized to correlate with genome size, and taking into account phylogenetic history, we found support only for climate seasonality and secondary chemistry as important correlates of genome size evolution in this group. Despite the 15 variables considered, we are still missing important factors, given that our best model explained only 19 % of the observed variation in genome size across *Streptanthus* and close relatives.

Other factors proposed to be associated with genome size, but lacking support in our study

We explored other environmental and autecological variables hypothesized to be related to genome size, but none was significant or included in our best models. In contrast to other studies (Watanabe *et al.*, 1999; Dušková *et al.*, 2010; Enke *et al.*, 2011), we did not find a correlation between genome size and life history, in particular annuality. Our analyses explicitly decoupled life history (annuality) from temperature and precipitation seasonality, and found that only climate seasonality predicted genome size. As a caveat, a majority of the species in *Streptanthus* and relatives are annuals (35/49 = 71 %), but biennial and perennial life histories have evolved independently

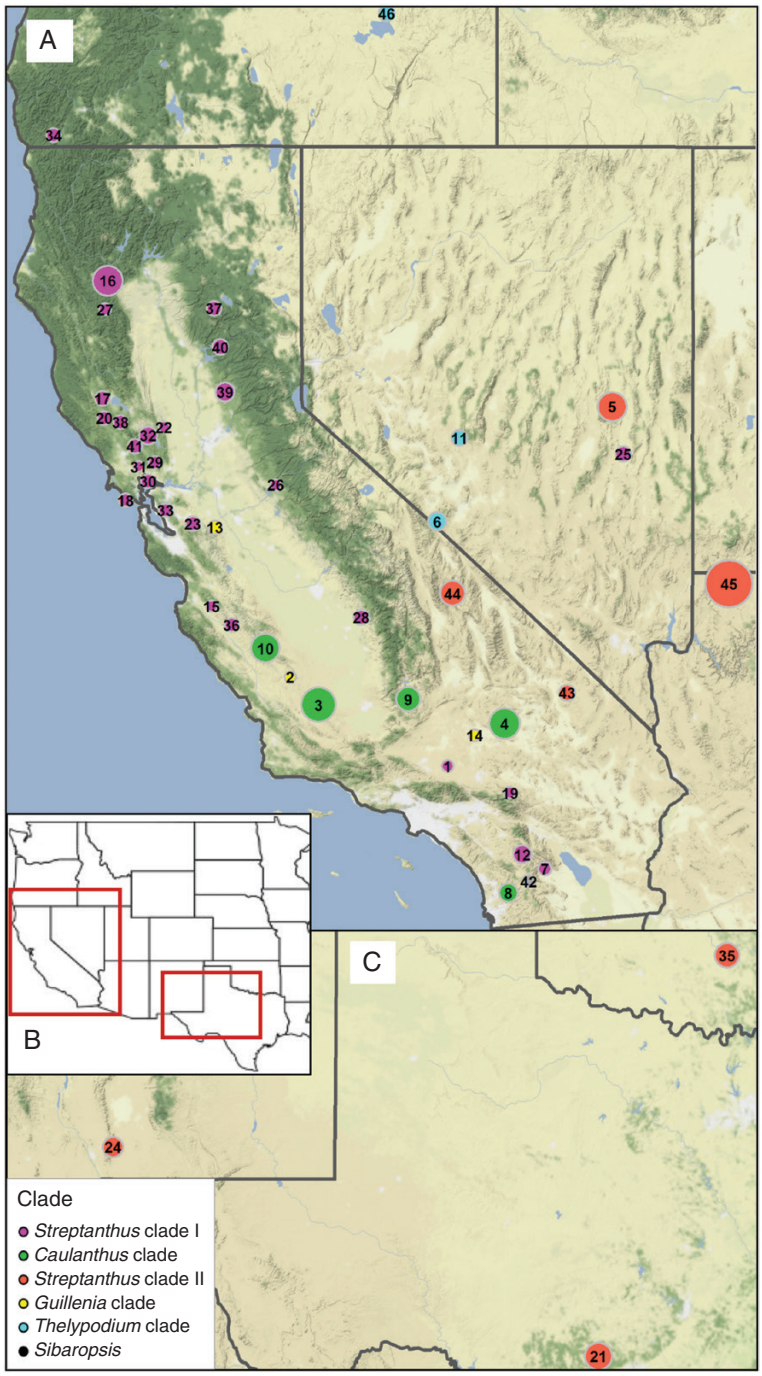

FIG. 3. Genome size in *Streptanthus* and close relatives tends to be smaller in environments with moderate temperature seasonality, such as those in the California Floristic Province, and larger in more seasonal continental areas when phylogeny is taken into account. In the figure, the top panel (A) illustrates species in California and adjacent Nevada, the line inset (B) illustrates the eastern area where species occur in New Mexico, Texas and Oklahoma (C). Each circle is plotted at the geographic centre of a species range, for lineages included in our analyses. Circle size on the map corresponds to standardized 2C*x* genome size of a species to better visualize differences on the landscape, and circle colours correspond to major clades (see key and Fig. 1). Species: (1) *C. amplexicaulis*, (2) *C. anceps*, (3) *C. californicus*, (4) *C. cooperi*, (5) *C. crassicaulis*, (6) *C. glaucus*, (7) *C. hallii*, (8) *C. heterophyllus*, (9) *C. inflatus*, (10) *C. lemmonii*, (11) *C. pilosus*, (12) *C. simulans*, (13) *C. flavescens*, (14) *C. lasiophyllus*, (15) *S. albidus*, (16) *S. barbatus*, (17) *S. barbiger*, (18) *S. batrachopus*, (19) *S. bernardinus*, (20) *S. brachiatus*, (21) *S. bracteatus*, (22) *S. breweri*, (23) *S. callistus*, (24) *S. carinatus*, (25) *S. cordatus*, (26) *S. diversifolius*, (27) *S. drepanoides*, (28) *S. farnsworthianus*, (29) *S. glandulosus_C1*, (30) *S. glandulosus_C2*, (31) *S. glandulosus_C3*, (32) *S. hesperidis,* (33) *S. hispidus*, (34) *S. howellii*, (35) *S. hyacinthoides*, (36) *S. insignis*, (37) *S. longisiliquus*, (38) *S. morrisonii*, (39) *S. polygaloides*, (40) *S. tortuosus*, (41) *S. vernalis*, (42) *Si. hammittii*, (43) *Sl. longirostris*, (44) *St. elata*, (45) *St. pinnata*, (46) *The. laciniatum*. Genera are abbreviated as follows: C = *Caulanthus*, S = *Streptanthus*, Si = *Sibaropsis*, Sl = *Streptanthella*, St = *Stanleya*, The = *Thelypodium*.

TABLE 4. *The best models to explain genome size from a model selection approach including significant variables from PGLS analyses*

Model	AIC (k = 2)	Δ AIC	Adj. R^2
T seasonality + Fr aliphatic	–53.896	0.00	0.186
T seasonality + Fr aliphatic + glucosinolate richness	–52.685	1.21	0.178
T seasonality + P seasonality + Fr aliphatic	–51.905	1.99	0.161
T seasonality + T annual range + Fr aliphatic	–51.897	2.00	0.161
T seasonality + glucosinolate richness	–51.876	2.02	–
T annual range + Fr aliphatic + glucosinolate richness	–51.684	2.21	–
T seasonality	–51.451	2.44	–
T seasonality + T annual range + Fr aliphatic + glucosinolate richness	–50.710	3.19	–
T seasonality + P seasonality + Fr aliphatic + glucosinolate richness	–50.696	3.20	–
T annual range + P seasonality + Fr aliphatic + glucosinolate richness	–50.002	3.89	–
T seasonality + T annual range + P seasonality + Fr aliphatics	–49.905	3.99	–
T seasonality + T annual range + glucosinolate richness	–49.876	4.02	–
Fr aliphatics	–49.815	4.08	–
Fr aliphatics + glucosinolate richness	–49.286	4.61	–
T seasonality + T annual range + P seasonality + Fr aliphatics + glucosinolate richness	–48.714	5.18	–
T seasonality + T annual range + P seasonality + glucosinolate richness	–48.031	5.86	–

T = temperature, P = precipitation, Fr = fraction. Variables were transformed as outlined in the Materials and Methods. Model selection was performed in a phylogenetic framework with the function 'phylostep' (option = lambda, 'phylolm' library; Tung Ho and Ané, 2014) and the MCCT from Cacho *et al.* (2014); adjusted R^2 values for best models were obtained with the function 'pgls' (option = lambda, 'caper' library; Orme *et al.*, 2013). For estimates of values for each of the best models see Supplementary data Table S8.

several times in this group. Annual streptanthoids tend to occur in the CFP, while perennials tend to have more eastern distributions, or occupy moister habitats or higher elevations within the CFP. Research using clades of widespread taxa with genome size and life history variation will be important in dissecting the correlations between life history and climate seasonality with respect to genome size variation.

Streptanthus and close relatives vary in range size >400-fold (298–129 948 km^2), but we find no evidence that species with larger ranges (and presumably larger N_e) had smaller genomes (Vinogradov, 2003; Lavergne *et al.*, 2010; Pandit *et al.*, 2014; Schmidt *et al.*, 2017), as have others (Whitney *et al.*, 2010). Our results thus offer no support for the hypothesis of Lynch and Conery (2003) of a negative association between population size and genome size, consistent with what others have found (Whitney *et al.*, 2010).

Smaller genomes have also been hypothesized to facilitate colonization of new areas and new habitats (Lavergne *et al.*, 2010; Schmidt *et al.*, 2017). In this context, smaller genome sizes in *Streptanthus* in the CFP could be associated with the recent radiation of the group onto harsh serpentine soils (Cacho and Strauss, 2014). While serpentine use has evolved 4–5 times in *Streptanthus*, we found no relationship between genome size and serpentine soil affinity. We also found no relationship to microhabitat bareness, a conserved trait across the clade, or with soil N or P concentrations (Šmarda *et al.*, 2013; Guignard *et al.*, 2016), which are hypothesized to be linked to costs of synthesis of DNA (Hanson *et al.*, 2001; Leitch and Leitch, 2008).

In summary, we find differences in genome size across subclades of *Streptanthus* and close relatives, consistent with a strong phylogenetic signal in genome size in this group of plants. We find evidence for rate variation in the evolution of genome size and no clear evidence that genome size evolves under different selective regimes among different clades of streptanthoids. After exploring a wide range of genome size correlates within *Streptanthus*, we find the most consistent relationships of genome size are with temperature seasonality (positively correlated) and with the fraction of methionine-based glucosinolates (aliphatic glucosinolates, negatively correlated) that a plant produces. Our results suggest that ecological interactions could also be important in shaping the evolutionary dynamics of genome size, consistent with recent findings (Guignard *et al.*, 2019), the classic idea of evolutionary escalation of plant defence akin to an arms race between plants and their herbivores (Ehrlich and Raven, 1964), and the origins of secondary compounds through gene duplication events (Edger *et al.*, 2015). Our study also illustrates the value of considering multiple intercorrelated factors simultaneously to evaluate hypotheses of genome size evolution in a phylogenetic context. Further investigation of these relationships, especially potential roles of species interactions in a phylogenetic context, may provide insights into the patterns of genome size variation in clades of angiosperms.

SUPPLEMENTARY DATA

Supplementary data are available online at https://academic.oup.com/aob and consist of the following. Methods S1: seed tissue as a source for genome size estimates in *Streptanthus* and close relatives. Table S1: phylogenetic signal values. Table S2: estimates of ancestral genome size. Table S3: parameters for BM-OU models of genome size evolution. Table S4: PCA loadings for soil multivariate data. Table S5: PCA loadings for climate multivariate data. Table S6: PGLS on genome size and climate variables for California Floristic Province taxa. Table S7: PCA loadings for glucosinolate multivariate data. Table S8: coefficients for best models from a model selection approach. Figure S1: genome size in streptanthoids in the context of genome sizes of Brassicaceae.

ACKNOWLEDGEMENTS

The authors thank B. McLaughlin and J. van Dyke from the UC Davis Flow Cytometry Shared Resource Laboratories for invaluable assistance; the Comai Lab for generously providing us with arabidopsis seed and lab space; B. Li and the Kliebenstein lab for technical assistance with glucosinolate data acquisition; B. Anacker for bioinformatic help to curate georeferenced records; E. Bergmann and C. Lee for assistance with ecological data collection. B. Drew, N. Fowler, R. O'Dell, D. O'Donnell, P. Schering, R. Raiche, J. Goddard, E. Gates and the University of California Natural Reserve System for valuable help in plant or soil collections or access to field sites. N.I.C., P.J.M. and S.Y.S. conceived the study and contributed to field data collections. N.I.C. and P.J.M. collected genome estimates data. N.I.C. and D.J.K. collected and interpreted glucosinolate data. N.I.C. analysed the data and wrote the first draft. P.J.M. contributed climate and spatial analyses. P.J.M. and S.Y.S. provided input for data analyses and subsequent drafts. All authors commented on and approved the final manuscript.

FUNDING

This work was supported by the National Science Foundation (DEB #0919559 to S.Y.S.), the Consejo Nacional de Ciencia y Tecnología (Conacyt #187083 to N.I.C.) and the program UNAM-DGAPA-PAPIIT (award IA201516 to N.I.C.).

LITERATURE CITED

Alonso C, Pérez 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* **5**: 1–9.

Al-Shehbaz IA. 2010. Brassicaceae. In: Flora of North America Editorial Committee, ed. *Flora of North America.* New York and Oxford: Oxford University Press, 224–746.

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.

Baldwin BG. 2014. Origins of plant diversity in the California Floristic Province. *Annual Review of Ecology, Evolution, and Systematics* **45**: 347–369.

Baldwin BG, Goldman DH, Keil DJ, Patterson R, Rosattii TJ, Wilken DH. 2012. *The Jepson manual, vascular plants of California*. Berkeley, CA: University of California Press.

Barrett CF, McKain MR, Sinn BT, *et al.* 2019. Ancient polyploidy and genome evolution in Palms. *Genome Biology and Evolution* **11**: 1501–1511.

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

Beaulieu JM, O'Meara BC, MacOS X. 2012. OUwie: analysis of evolutionary rates in an OU framework. https://rdrr.io/cran/OUwie/.

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

Bennett MD, Leitch IJ, Price HJ, Johnston JS. 2003. Comparisons with *Caenorhabditis* (approximately 100 Mb) and *Drosophila* (approximately 175 Mb) using flow cytometry show genome size in *Arabidopsis* to be approximately 157 Mb and thus approximately 25 % larger than the *Arabidopsis* genome initiative estimate of approximately 125 Mb. *Annals of Botany* **91**: 547–557.

Bilinski P, Albert PS, Berg JJ, *et al.* 2018. Parallel altitudinal clines reveal trends in adaptive evolution of genome size in *Zea mays*. *PLoS Genetics* **14**: 1–19.

Bivand R, Rundel C. 2013. rgeos: Interface to Geometry Engine - Open Source (GEOS). https://rdrr.io/cran/rgeos/.

Blomberg SP, Garland T Jr, Ives AR. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* **57**: 717–745.

Borregaard MK, Rahbek C. 2010. Causality of the relationship between geographic distribution and species abundance. *The Quarterly Review of Biology* **85**: 3–25.

Bottini MCJ, Greizerstein EJ, Aulicino MB, Poggio L. 2000. Relationships among genome size, environmental conditions and geographical distribution in natural populations of NW patagonian species of *Berberis* L. (Berberidaceae). *Annals of Botany* **86**: 565–573.

Brown JH. 1984. On the relationship between abundance and distribution of species. *The American Naturalist* **124**: 255.

Burnham KP, Anderson DR. 2002. *Model selection and multimodel inference: a practical information-theoretic approach.* New York: Springer-Verlag.

Caceres ME, De Pace C, Scarascia Mugnozza GT, Kotsonis P, Ceccarelli M, Cionini PG. 1998. Genome size variations within *Dasypyrum villosum*: correlations with chromosomal traits, environmental factors and plant phenotypic characteristics and behaviour in reproduction. *Theoretical and Applied Genetics* **96**: 559–567.

Cacho NI, Strauss SY. 2014. Occupation of bare habitats, an evolutionary precursor to soil specialization in plants. *Proceedings of the National Academy of Sciences, USA* **111**: 15132–15137.

Cacho NI, Burrell AM, Pepper AE, Strauss SY. 2014. Novel nuclear markers inform the systematics and the evolution of serpentine use in *Streptanthus* and allies (Thelypodieae, Brassicaceae). *Molecular Phylogenetics and Evolution* **72**: 71–81.

Cacho NI, Kliebenstein DJ, Strauss SY. 2015. Macroevolutionary patterns of glucosinolate defense and tests of defense-escalation and resource availability hypotheses. *New Phytologist* **208**: 915–927.

Cappa JJ, Cappa PJ, El Mehdawi AF, McAleer JM, Simmons MP, Pilon-Smits EA. 2014. Characterization of selenium and sulfur accumulation across the genus *Stanleya* (Brassicaceae): a field survey and common-garden experiment. *American Journal of Botany* **101**: 830–839.

del Carmen Martínez-Ballesta M, Moreno DA, Carvajal M. 2013. The physiological importance of glucosinolates on plant response to abiotic stress in Brassica. *International Journal of Molecular Sciences* **14**: 11607–11625.

Castro-Jimenez Y, Newton RJ, Price HJ, Halliwell RS. 1989. Drought stress responses of *Microseris* species differing in nuclear DNA content. *American Journal of Botany* **76**: 789–795.

Cerbah M, Coulaud J, Brown SC, Siljak-Yakovlev S. 1999. Evolutionary DNA variation in the genus *Hypochaeris*. *Heredity* **82**: 261–266.

Clark J, Hidalgo O, Pellicer J, *et al.* 2016. Genome evolution of ferns: evidence for relative stasis of genome size across the fern phylogeny. *New Phytologist* **210**: 1072–1082.

Díez CM, Gaut BS, Meca E, *et al.* 2013. Genome size variation in wild and cultivated maize along altitudinal gradients. *New Phytologist* **199**: 264–276.

Dodsworth S, Chase MW, Kelly LJ, *et al.* 2015. Genomic repeat abundances contain phylogenetic signal. *Systematic Biology* **64**: 112–126.

Dolezel J, Greilhuber J, Suda J. 2007. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* **2**: 2233–2244.

Dušková E, Kolář F, Sklenář P, *et al.* 2010. Genome size correlates with growth form, habitat and phylogeny in the Andean genus *Lasiocephalus* (Asteraceae). *Preslia* **82**: 127–148.

Edger PP, Heidel-Fischer HM, Bekaert M, *et al.* 2015. The butterfly plant arms-race escalated by gene and genome duplications. *Proceedings of the National Academy of Sciences, USA* **112**: 8362–8366.

Ehrlich PR, Raven PH. 1964. Butterflies and plants: a study in coevolution. *Evolution* **18**: 586–608.

Enke N, Fuchs J, Gemeinholzer B. 2011. Shrinking genomes? Evidence from genome size variation in *Crepis* (Compositae). *Plant Biology (Stuttgart, Germany)* **13**: 185–193.

Fawcett JA, Van de Peer Y, Maere S. 2013. Significance and biological consequences of polyploidization in land plant evolution. In: Leitch IJ, Greilhuber J, Dolezel J, Wendel JF, eds. *Plant genome diversity volume 2*. Wein: Springer, 277–293.

Fick SE, Hijmans RJ. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37**: 4302–4315.

Fleischmann A, Michael TP, Rivadavia F, *et al.* 2014. Evolution of genome size and chromosome number in the carnivorous plant genus *Genlisea* (Lentibulariaceae), with a new estimate of the minimum genome size in angiosperms. *Annals of Botany* **114**: 1651–1663.

Garnier E. 1992. Growth analysis of congeneric annual and perennial grass species. *Journal of Ecology* **80**: 665–675.

Gaston KJ. 1996. The multiple forms of the interspecific abundance–distribution relationship. *Oikos* **76**: 211–220.

Grafen A. 1989. The phylogenetic regression. *Philosophical Transactions of the Royal Society B: Biological Sciences* **326**: 119–157.

Greilhuber J, Dolezel J, Lysák MA, Bennett MD. 2005. The origin, evolution and proposed stabilization of the terms 'genome size' and 'C-value' to describe nuclear DNA contents. *Annals of Botany* **95**: 255–260.

Grime JP, Hunt R. 1975. Relative growth-rate: its range and adaptive significance in a local flora. *Journal of Ecology* **63**: 393–422.

Grime JP, Mowforth MA. 1982. Variation in genome size – an ecological interpretation. *Nature* **299**: 151–153.

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

Guignard MS, Nichols RA, Knell RJ, *et al.* 2016. Genome size and ploidy influence angiosperm species' biomass under nitrogen and phosphorus limitation. *New Phytologist* **210**: 1195–1206.

Guignard MS, Crawley MJ, Kovalenko D, *et al.* 2019. Interactions between plant genome size, nutrients and herbivory by rabbits, molluscs and insects on a temperate grassland. *Proceedings of the Royal Society B: Biological Sciences* **286**: 20182619.

Hanson L, McMahon KA, Johnson MAT, Bennett MD. 2001. First nuclear DNA C-values for 25 angiosperm families. *Annals of Botany* **87**: 251–258.

Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W. 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics* **24**: 129–131.

Haugen R, Steffes L, Wolf J, Brown P, Matzner S, Siemens DH. 2008. Evolution of drought tolerance and defense: dependence of tradeoffs on mechanism, environment and defense switching. *Oikos* **117**: 231–244.

Hijmans RJ, Phillips S, Leathwick J, Elith J. 2013. dismo: species distribution modeling. http://CRANR-project.org/package=dismo.

Hopkins RJ, van Dam NM, van Loon JJ. 2009. Role of glucosinolates in insect–plant relationships and multitrophic interactions. *Annual Review of Entomology* **54**: 57–83.

Janz N. 2011. Ehrlich and Raven revisited: mechanisms underlying codiversification of plants and enemies. *Annual Review of Ecology, Evolution, and Systematics* **42**: 71–89.

Jedrzejczyk I, Sliwinska E. 2010. Leaves and seeds as materials for flow cytometric estimation of the genome size of 11 Rosaceae woody species containing DNA-staining inhibitors. *Journal of Botany* **2010**: 930895.

Johnson CN. 1998. Species extinction and the relationship between distribution and abundance. *Nature* **394**: 272.

Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH. 2000. Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. *Proceedings of the National Academy of Sciences, USA* **97**: 6603–6607.

Kelly LJ, Renny-Byfield S, Pellicer J, *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.

Kembel SW, Cowan PD, Helmus MR, *et al.* 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**: 1463–1464.

Kliebenstein DJ, Kroymann J, Mitchell-Olds T. 2005. The glucosinolate–myrosinase system in an ecological and evolutionary context. *Current Opinion in Plant Biology* **8**: 264–271.

Kliebenstein DJ, Lambrix VM, Reichelt M, Gershenzon J, Mitchell-Olds T. 2001. Gene duplication in the diversification of secondary metabolism: tandem 2-oxoglutarate-dependent dioxygenases control glucosinolate biosynthesis in *Arabidopsis*. *The Plant Cell* **13**: 681–693.

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.

Kwan JC, Donia MS, Han AW, Hirose E, Haygood MG, Schmidt EW. 2012. Genome streamlining and chemical defense in a coral reef symbiosis. *Proceedings of the National Academy of Sciences, USA* **109**: 20655–20660.

Lavergne S, Muenke NJ, Molofsky J. 2010. Genome size reduction can trigger rapid phenotypic evolution in invasive plants. *Annals of Botany* **105**: 109–116.

Leitch AR, Leitch IJ. 2008. Genomic plasticity and the diversity of polyploid plants. *Science* **320**: 481–483.

Leitch IJ, Bennett MD. 2004. Genome downsizing in polyploid plants. *Biological Journal of the Linnean Society* **82**: 651–663.

Leitch I, Johnston E, Pellicer J, Hidalgo O, Bennett MD. 2019. Angiosperm DNA C-values database (release 9.0). https://cvalues.science.kew.org/.

Lynch M, Conery JS. 2003. The origins of genome complexity. **302**: 1401–1404.

Lysak MA, Koch MA, Beaulieu JM, Meister A, Leitch IJ. 2009. The dynamic ups and downs of genome size evolution in Brassicaceae. *Molecular Biology and Evolution* **26**: 85–98.

Macgillivray C, Grime J. 1995. Genome size predicts frost-resistance in British herbaceous plants – implications for rates of vegetation response to global warming. *Functional Ecology* **9**: 320–325.

McIntyre PJ. 2012. Cytogeography and genome size variation in the *Claytonia perfoliata* (Portulacaceae) polyploid complex. *Annals of Botany* **110**: 1195–1203.

Messing J, Bharti AK, Karlowski WM, *et al.* 2004. Sequence composition and genome organization of maize. *Proceedings of the National Academy of Sciences, USA* **101**: 14349–14354.

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.

Müller LLB, Zotz G, Albach DC. 2019. Bromeliaceae subfamilies show divergent trends of genome size evolution. *Scientific Reports* **9**: 1–12.

Nakazato T, Warren DL, Moyle LC. 2010. Ecological and geographic modes of species divergence in wild tomatoes. *American Journal of Botany* **97**: 680–693.

Niklas KJ. 1994. *Plant allometry: the scaling of form and process.* Chicago: University of Chicago Press.

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

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.

Orme CDL, Freckleton RP, Thomas GH, Petzoldt T, Fritz SA. 2013. *Caper: comparative analyses of phylogenetics and evolution in R.* https://cran.r-project.org/web/packages/caper/vignettes/caper.pdf

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

Pellicer J, Fay MF, Leitch IJ. 2010. The largest eukaryotic genome of them all? *Botanical Journal of the Linnean Society* **164**: 10–15.

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

Piegu B, Guyot R, Picault N, *et al.* 2006. Doubling genome size without polyploidization: dynamics of retrotransposition-driven genomic expansions in *Oryza australiensis*, a wild relative of rice. *Genome Research* **16**: 1262–1269.

Price HJ, Bachmann K. 1975. DNA content and evolution in the Microseridinae. *American Journal of Botany* **62**: 262–267.

Price HJ, Chambers KL, Bachmann K. 1981. Geographic and ecological distribution of genomic DNA content variation in *Microseris douglasii* (Asteraceae). *Botanical Gazette* **142**: 415–426.

Qiu F, Baack EJ, Whitney KD, *et al.* 2019. Phylogenetic trends and environmental correlates of nuclear genome size variation in *Helianthus* sunflowers. *New Phytologist* **221**: 1609–1618.

R Development Core Team. 2018. *R: a language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Reichelt M, Brown PD, Schneider B, *et al.* 2002. Benzoic acid glucosinolate esters and other glucosinolates from *Arabidopsis thaliana*. *Phytochemistry* **59**: 663–671.

Revell LJ. 2012. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**: 217–223.

Rollins RC. 1993. *The Cruciferae of continental North America : systematics of the mustard family from the Arctic to Panama.* Stanford: Stanford University Press.

Salehin M, Li B, Tang M, *et al.* 2019. Auxin-sensitive Aux/IAA proteins mediate drought tolerance in *Arabidopsis* by regulating glucosinolate levels. *Nature Communications* **10**: 4021.

Sato Y, Tezuka A, Kashima M, *et al.* 2019. Transcriptional variation in glucosinolate biosynthetic genes and inducible responses to aphid herbivory on field-grown *Arabidopsis thaliana*. *Frontiers in Genetics* **10**: 787.

Schmidt JP, Drake JM, Stephens P. 2017. Residence time, native range size, and genome size predict naturalization among angiosperms introduced to Australia. *Ecology and Evolution* **7**: 10289–10300.

Schmuths H, Meister A, Horres R, Bachmann K. 2004. Genome size variation among accessions of *Arabidopsis thaliana*. *Annals of Botany* **93**: 317–321.

Shapiro AM. 1981. Egg-mimics of *Streptanthus* (Cruciferae) deter oviposition by *Pieris sisymbrii* (Lepidoptera: Pieridae). *Oecologia* **48**: 142–143.

Simonin KA, Roddy AB. 2018. Genome downsizing, physiological novelty, and the global dominance of flowering plants. *PLoS Biology* **16**: 1–15.

Sliwinska E, Zielinska E, Jedrzejczyk I. 2005. Are seeds suitable for flow cytometric estimation of plant genome size? *Cytometry. Part A* **64**: 72–79.

Šmarda P, Hejcman M, Březinová A, *et al.* 2013. Effect of phosphorus availability on the selection of species with different ploidy levels and genome sizes in a long-term grassland fertilization experiment. *New Phytologist* **200**: 911–921.

Šmarda P, Bureš P, Horová L, *et al.* 2014. Ecological and evolutionary significance of genomic GC content diversity in monocots. *Proceedings of the National Academy of Sciences, USA* **11**: E4096–E4102.

Soltis DE, Soltis PS, Bennett MD, Leitch IJ. 2003. Evolution of genome size in the angiosperms. *American Journal of Botany* **90**: 1596–1603.

Soltis PS, Soltis DE. 2013. Angiosperm phylogeny: a framework for studies of genome evolution. In: Leitch IJ, Greilhuber J, Dolezel J, Wendel JF, eds. *Plant genome diversity volume 2: physical structure, behaviour and evolution of plant genomes.* Wien: Springer, 1–11.

Souza G, Costa L, Guignard MS, *et al.* 2019. Do tropical plants have smaller genomes? Correlation between genome size and climatic variables in the Caesalpinia Group (Caesalpinioideae, Leguminosae). *Perspectives in Plant Ecology, Evolution and Systematics* **38**: 13–23.

Strauss SY, Cacho NI. 2013. Nowhere to run, nowhere to hide: the importance of enemies and apparency in adaptation to harsh soil environments. *The American Naturalist* **182**: E1–14.

Strauss SY, Cacho NI, Schwartz MW, Schwartz AC, Burns KC. 2015. Apparency Revisited. *Entomologia Experimentalis et Applicata* **157**: 74–85.

Streiner DL. 2015. Best (but oft-forgotten) practices: the multiple problems of multiplicity – whether and how to correct for many statistical tests. *American Journal of Clinical Nutrition* **102**: 721–728.

Suda J, Kyncl T, Freiova R. 2003. Nuclear DNA amounts in Macaronesian angiosperms. *Annals of Botany* **92**: 153–164.

Textor S, Gershenzon J. 2009. Herbivore induction of the glucosinolate–myrosinase defense system: major trends, biochemical bases and ecological significance. *Phytochemistry Reviews* **8**: 149–170.

Thompson K. 1990. Genome size, seed size and germination temperature in. *Evolutionary Trends in Plants* **4**: 113–116.

Torrell M, Vallès J. 2001. Genome size in 21 *Artemisia* L. species (Asteraceae, Anthemideae): systematic, evolutionary, and ecological implications. *Genome* **44**: 231–238.

Trávníček P, Čertner M, Ponert J, Chumová Z, Jersáková J, Suda J. 2019. Diversity in genome size and GC content shows adaptive potential in orchids and is closely linked to partial endoreplication, plant life-history traits and climatic conditions. *New Phytologist* **224**: 1642–1656.

Tung Ho LS, Ané C. 2014. A linear-time algorithm for gaussian and non-gaussian trait evolution models. *Systematic Biology* **63**: 397–408.

Turnbull LA, Philipson CD, Purves DW, *et al.* 2012. Plant growth rates and seed size: a re-evaluation. *Ecology* **93**: 1283–1289.

Van't Hof J, Sparrow AH. 1963. A relationship between DNA content, nuclear volume, and minimum mitotic cycle time. *Proceedings of the National Academy of Sciences, USA* **49**: 897–902.

Veselý P, Bureš P, Šmarda P, Pavlíček T. 2012. Genome size and DNA base composition of geophytes: the mirror of phenology and ecology? *Annals of Botany* **109**: 65–75.

Vinogradov AE. 2003. Selfish DNA is maladaptive: evidence from the plant Red List. *Trends in Genetics* **19**: 609–614.

Wakamiya I, Newton RJ, Johnston JS, Price HJ. 1993. Genome size and environmental factors in the genus *Pinus*. *American Journal of Botany* **80**: 1235–1241.

Wang X, Shi X, Hao B, Ge S, Luo J. 2005. Duplication and DNA segmental lss in the rice genome: implications for diploidization. *New Phytologist* **165**: 937–946.

Warwick SI, Al-Shehbaz IA. 2006. Brassicaceae: chromosome number index and database on CD-Rom. *Plant Systematics and Evolution* **259**: 237–248.

Watanabe K, Yahara T, Denda T, Kosuge K. 1999. Chromosomal evolution in the genus *Brachyscome* (Asteraceae, Astereae): statistical tests regarding correlation between changes in karyotype and habit using phylogenetic information. *Journal of Plant Research* **112**: 145–161.

Weber MG, Cacho NI, Phan MJQ, Disbrow C, Ramírez SR, Strauss SY. 2018. The evolution of floral signals in relation to range overlap in a clade of California Jewelflowers (*Streptanthus* s.l.). *Evolution* **72**: 798–807.

Wendel JF. 2015. The wondrous cycles of polyploidy in plants. *American Journal of Botany* **102**: 1753–1756.

Wendel JF, Cronn RC, Alvarez I, Liu B, Small RL, Senchina DS. 2002. Intron size and genome size in plants. *Molecular Biology and Evolution* **19**: 2346–2352.

Wendel JF, Jackson SA, Meyers BC, Wing RA. 2016. Evolution of plant genome architecture. *Genome Biology* **2016**: 17–37.

Whitney KD, Baack EJ, Hamrick JL, *et al.* 2010. A role for nonadaptive processes in plant genome size evolution? *Evolution* **64**: 2097–2109.

Zenil-Ferguson R, Ponciano JM, Burleigh JG. 2016. Evaluating the role of genome downsizing and size thresholds from genome size distributions in angiosperms. *American Journal of Botany* **103**: 1175–1186.