Genome Size and DNA base composition of Geophytes: the Mirror of Phenology and Ecology?

BIOLOGICAL SCIENCES: Evolution

Pavel Vesely¹*, Petr Bures¹, Petr Smarda¹, Tomas Pavlivek²

¹ Department of Botany and Zoology, Masaryk University, Kotlarska 2, CZ-61137, Brno, Czech Republic
² Institute of Evolution, University of Haifa, Mt. Carmel, 31905, Haifa, Israel.
* corresponding author: Pavel Vesely: pvesely@sci.muni.cz
Abstract

Genome size is known to affect various plant traits such as stomatal size, seed mass, and flower or shoot phenology. However, these associations are still not well understood for the species with giant genomes, mostly represented with geophytic plants. No detailed associations are known between DNA base composition and genome size or species ecology. Genome sizes and GC contents were measured in 219 geophytes together with tentative morpho-anatomical and ecological traits.

Increase in genome size was associated with earliness of flowering, tendency to grow in humid conditions, and tightly positively correlated with increase in stomatal size, namely in extremely large genomes. Seed mass of geophytes was closely related to their ecology, but not to genomic parameters. Genomic GC content showed unimodal relationship with genome size but no relation with species ecology.

Evolution of genome size in geophytes is closely related with their ecology and phenology and also associates with remarkable changes in DNA base composition. While geophytism together with producing larger cells seems advantageous strategy for fast development of an organism in seasonal habitats, this is perhaps the drought sensitivity of large stomata which restrict the occurrence of geophytes with giant genomes to water non-stressed conditions.

Key words: Life-form, geophytes, genome size evolution, GC content, phenology, stomatal length, seed mass, ecology
Introduction

Genome size and its consequences

Genome size (i.e., the DNA content of the unreplicated nucleus [2C]) (1) varies considerably among eukaryotic organisms, with a minimum (1C = 2.25 Mbp) reported in Encephalitozoon intestinalis (Zygomyctota; 2) and a maximum (1C = 1 369 200 Mbp) disclosed in Chaos chaos (Amoebozoa; 3), representing a difference of more than 600 000-fold. In tracheophytes, the difference is smaller but still remarkable: a minimum of 1C = 63.57 Mbp is present in Genlisea margaretae (Lentibulariaceae; 4) and, for a long time, a known maximum of 1C = 124 597.2 Mbp has been related to a geophyte Fritillaria assyriaca (Liliaceae; 5). However, during the last year, two larger genomes were reported in another geophytic species, Trillium hagae (Melanthiaceae) with 1C = 129 536.1 Mbp (6) and Paris japonica (Melanthiaceae) with 1C = 148 880.9 Mbp (7). Thus, the divergence in tracheophytes exceeds 2 300-fold.

Regardless of the causes of the differences in genome size (e.g., polyploidy or the amount of repetitive and non-coding DNA; 8), the actual amount of nuclear DNA may limit some plant traits, such as maximum height, growth rate, generation time, and presence in certain types of niches (9, 10). For species with larger genomes, the mitotic cell cycle takes additional time (11, 10), and replicating more DNA is perhaps more energetically demanding (12). Cell division and, thus, the growth rates of plants with large genomes should be slower than those with smaller genomes in the same ecological conditions, as has been supported experimentally on root cells (13). Moreover, mitosis and hence plant growth are inhibited by the low temperatures (14) that occur frequently during the spring in the temperate zone. This finding may imply that the very early spring ecological and temporal niche, and thus the rapid growth rate, during this period may appear to be inaccessible for herbaceous plant species with large genomes. However, among early spring flowering species, some display rapid growth and development in spite of their unusually large genomes (5). Interestingly, most of these plants are geophytes, indicating that in addition to ecology-driven forces, the evolution of genome size is also closely related with specific life-forms.

Genome size and life-forms

The standardized terminology of plant life history distinguishes the following terrestrial life-forms: phanerophyte, chamaephyte, hemicryptophyte, geophyte, and therophyte (15). According to the triangular scheme by Leitch & Bennett (16), for every herbaceous species, there exists a minimum generation time that is determined by its genome size and having a large genome clearly does not allow a plant species to adopt certain life strategies. A therophyte (an ephemeral or annual species), therefore, must have a small genome to be able to complete its life cycle before the end of the growing season. Phanerophytes (trees and shrubs) are not temporally limited like therophytes, however, some structural restriction may exist on the cell size (correlated with genome size) composing their bodies (17, 18). Therefore, only several woody angiosperm plants are known to possess larger genomes (cf. 5). The remaining categories (monocarpic or perennial hemicryptophytes and geophytes) are also thought to show some kind of genome size associated ecological constraints (9), but show no clear reason to be strongly limited with genome size (16). Nevertheless, the extremely large genome sizes are not found randomly in all categories but clearly predominate in species with geophytic life-form (e.g., the genera Paris, Trillium, Fritillaria, Erythronium and Leucojum; 5). Until recently, however, genome size in geophytes has not been studied systematically, disallowing to make serious conclusions on the expected role of this life-form in the evolution of plant genome gigantism. The published data on genome size of geophytic species indicates
that genome gigantism perhaps is not allowed in all phylogenetic lineages and is associated with some specific ecological conditions, such as spring growth (19). However, no detailed study of geophytes’ ecology in relation to their genome size has been made, to date, across a wider phylogenetic spectrum.

**Geophytic life-form**

In this study, we focused on a certain type of geophytes, seasonal plants with a storage organ (i.e., bulb, tuber, fusiform roots or thick rhizome), that is usually subterranean but may also be partially above ground level (e.g., *Urginea maritima*). This definition partly differs from the strict definition of geophytes by Raunkiaer (15). When we consider geophytes, we describe the seasonal type only, not those that have neither a storage organ nor seasonal behaviour (e.g., *Elytrigia repens*). Geophytism in this form should be considered an adaptation to the cyclicity of changes in environmental conditions, such as alternation of short light phase by a long period of shading or appearance of long periodic droughts, which are inaccessible for most species with other life strategies (20, 21, 22). Geophytes overcome the unfavourable growing conditions by persisting in the form of a subterranean organ that can accumulate sufficient amount of nutrients available for the fast development of the aboveground body during annually cycling periods of suitable conditions (23, 24). The stored reserves allow to geophytes successful development even at availability of a very low initial external energetic supply. This trait provide them with a favourable strategy (i) in steppe vegetation in seasonally arid Mediterranean climate or (ii) in temperate deciduous forests in which the light phase is restricted to cold seasons characterized by a low initial level of solar insolation – a niche generally avoided by therophytes. Unlike the therophytes, in which the plant body development depends on the speed of cell division, the fast development of a geophytic body is enabled by pumping the water to the cells, which are pre-divided underground in the storage organ during the photosynthetically inactive dormant period (19, 25, 26, 27, 28). Because their cells are full of water, the plant body is frequently fleshy and to protect themselves from herbivore damage, they usually contain various toxic or repellent substances (e.g., the genera *Cyclamen, Colchicum, Urginea* and *Allium*; 29, 30, 22).

**Genome size and geophytes’ phenology**

Geophytes are often regarded as spring-flowering plants. However, in temperate zones, there are geophytes that flower not only in the spring (March–May, e.g., *Gallanthus nivalis, Eranthis hyemalis*), but also in the summer (June–August, e.g., *Allium flavum, Urginea maritima*), autumn (September–November, e.g., *Colchicum autumnale, Sternbergia colchiciflora*) and winter (December–February, e.g., *Helleborus niger*). Since the beginning of more intensive study of genome sizes, a negative correlation has been reported between genome size and the timing of shoot expansion (19, 31). As the timing of shoot expansion is related to the timing of flowering, a relationship may also be present between genome size and the timing of flowering. This relationship has been already observed in geophytic genus *Allium* (32, but see also 33), however, it is not clear to what extent these observations may be generalized for geophytes at all.

**Genome size and stomatal length**

Stomata play a key role in the regulation of gas exchange and overall plant photosynthesis (balancing carbon assimilation and plant water status), and are considered one of the key elements in the evolution of advanced terrestrial plants (34, 35). The regulation of stomata and carbon dioxide uptake is related to stomatal length because the large stomata are known to increase sensitivity of plants to drought (36, 37, 38).
In higher plants, stomatal guard cell length is known to be very closely correlated with genome size (39, 11, 40, 41, 42) implying that the evolution of genome size would have also necessarily important physiological consequences. This correlation partly results from the constraint on cell functioning given by the existence of an optimum ratio between nuclear (≈ DNA content) and cytoplasm volumes, for more than hundred of years known as “karyoplasmic ratio” (43, 12, 44). However, the relationship between genome size and stomatal length (cell size) was usually studied on a limited scale in species with relatively small genomes, rarely larger than $2C = 60,000$ Mbp (41). Thus, the question remains whether the observed relationship is universal across whole spectrum of genome and cell sizes and what consequences extremely large stomata may have for the ecology of large genomed species.

**Genome size and seed mass**

Earlier studies reported positive correlations between genome size and seed mass (45). If there is a causal relationship between cell size and genome size, some dependence of seed mass on genome size could be also expected, based on the fact that seeds are also composed of a ± definite number of cells (but see also 46, 47). However, in some cases (11), a triangular type of relationship was revealed: species with small genomes can have either small or large seeds, but species with larger genomes have only larger seeds. A recent study (48) that was based on a survey of 1222 species showed a generally positive correlation between genome size and seed size. Here, we were interested to see whether a similar relationship might be found also in geophytes and whether this relationship might anyway differ for extremely large genomes.

**GC content variation**

The genomic percentage of guanine + cytosine bases (GC content) is highly variable among prokaryotes and other unicellular organisms (49). In tracheophytes, GC content is generally narrower and ranges from 35% to 40% (50, 49). Only grasses (family Poaceae) are known to date to significantly exceed this range and regularly contain species with GC contents higher than 40%. Since the early studies on DNA base composition, a question has arisen to what extent changing of GC content is a passive consequence of molecular mechanisms of genome size change and to what extent it may represent a selective advantage for species with certain ecology. While the relation of GC content with genome size used to be addressed in studies of genome composition from prokaryotes to plants, the knowledge of possible ecological consequences of GC content change are completely unknown in plants.

A positive relationship between genome size and GC content has been documented in prokaryotes (51) and vertebrates (52). Some reports have shown that the relationship might also exist in plants (53, 54), where positive correlations were reported in groups of phylogenetically related taxa (e.g., species within a particular genus or family). A negative correlation is known only for teleost fishes (52). In plants the most detailed study across 54 species from diverse angiosperms and gymnosperm families did not show any significant trend (50). Until recently, the correlation of genome size and GC content has been tested only with classical statistical tests, which do not consider the phylogenetic dependence of compared species. In addition to the very poor knowledge on GC content in plants, this practice may result in biased estimates on the relationship. Therefore, a phylogenetically corrected analysis is still desirable. Geophytes may be a very suitable group for such an analysis, as they occur in a wide spectrum of phylogenetic lineages of tracheophytes and cover almost the entire scale of plant genome sizes, providing enough robust data for hypothesis testing. They may be also useful for the initial testing of possible consequences of
GC content on plant ecology, although the restriction of the study on the geophytic species necessarily limits the scale of testable conditions and associations.

Results and discussion

Genome size
The genome sizes measured in 219 geophytic species in this study ranged 220-fold with minimum $2C = 582.18$ Mbp detected in *Aristolochia rotunda* (Aristolochiaceae), and the maximum $2C = 128 273.07$ Mbp found in *Sprekelia formosissima* (Amaryllidaceae; Table S2).

Genome sizes exceeding $2C = 40 000$ Mbp were found in 33 out of all 219 geophytic species analyzed. They are represented namely with species from orders **Liliales** and **Asparagales**. Some large genomed geophytes may be found also in the families *Ranunculaceae*, *Paeoniaceae*, *Adoxaceae* and *Araceae*. In contrast to these, geophytes with small genomes are found regularly across the whole tracheophyte phylogeny.

A simple inspection of the $C$-value database (5) indicates that genome size in geophytes is perhaps larger compared to the remaining plant life-forms, although the exact effect of geophytic life-form on genome evolution remains to be exactly tested with phylogenetically independent methods. The geophytic life-form is only rarely found in species with lower genome sizes and, despite the clear overrepresentation of large genomes in geophytes (406-fold; min: 733.5 Mbp; max: 297 761.88 Mbp; 53, 7), they show a very similar range of genome size to the non-geophytic plants (667-fold; 4, 5).

Genome size and stomatal length
Logarithmed genome size and stomatal length data were tightly linearly correlated (Table 1) with the parameters of the regression line very close to that observed across whole angiosperms by Beaulieu *et al.* (41) and Hodgson *et al.* (42). This correlation was significant also when calculating with PICs (Table 1). In addition to the former analyses with limited amount of species with extremely large genomes, our analysis suggests that this relationship may be extremely tight in the case of giant genomes that never show stomatal size less than a certain threshold (Fig. 1). A different slopes of regression lines may be therefore observed in the separate datasets with species with low-to-high genome sizes ($2C < 40 000$ Mbp; $r = 0.448$; $p < 0.001$; Fig. 1), and extremely high genome sizes ($2C > 40 000$ Mbp; $r = 0.545$; $p = 0.003$; Fig. 1). The steepness of the regression line in species with extremely large genomes was not simply an artifact of data division but seems due to the surprising lack of small-sized stomata in extremely large genomes. This trend was documented also with separate analyses of PICs for the two subsets of data where correlation of PICs for genome size and stomatal size was much tighter in the case that nodal genome size was calculated as $> 40 000$ Mbp ($2C < 40 000$ Mbp: PIC: $r = 0.579$; $p < 0.001$; $2C > 40 000$ Mbp: PIC: $r = 0.789$; $p < 0.001$). To verify that the absence of smaller than expected stomatal sizes in extremely large genomes (apparent namely in genome sizes $> 60 000$ Mbp; Fig. 1) is not the artifact of the life-form limited scope of our study, we supply also a measurement of stomatal sizes for known non-geophytic species with extremely large genome (*Tradescantia virginiana*) as well as further also the measurements from the plants with the absolutely highest known genome sizes (nine species from genus *Fritillaria*). Including data from these species did not change the trend that was showing some strong threshold for minimum stomatal size in species with $2C > 60 000$ Mbp (Fig. 1).

Assuming that the size of a cell may be determined jointly with ecologically driven constraints on leaf morphology and plant physiology (37, 42) and genome size, our data suggests that extremely large genomes do not allow stomatal size to fall below certain limits
that is perhaps necessary for correct cell functioning. At the same time, however, producing of extremely large stomata could be disadvantageous for some reasons because they are only rarely observed in the case of giant genomes. With decreasing genome size, the association of genome size with minimum cell size became perhaps less constraining, allowing the stomatal size to be driven to an increasing intensity with other specific eco-morphological and eco-physiological constraints. As a consequence, the stomatal sizes in small genomed species show a relatively high variation and close correlation with specific eco-physiological adaptations rather than to be in strict correspondence with genome size (cf. 42).

**Genome size and flowering phenology**

A close negative relationship was observed between genome size and end of flowering (Fig. 2; Table 1). The relationship was verified also using analysis of PICs (Table 1). There was observed also a weak association of large genome sizes with onset of flowering, however, this relationship was below conventional significance levels by means of classical statistics and also by analyzing PICs (Table 1).

As mentioned in the introduction, low temperatures inhibit mitosis and, therefore, growth of the entire plant. However, in temperate zones, low temperatures frequently occur in early spring for extended periods of time, so a herbaceous spring plant cannot produce a large body composed of many small cells. To overcome this problem, two basic strategies have evolved among spring herbaceous plants. The first is to be an ephemeral therophyte with a small body that may develop fast in response to increasing temperatures. The second is to generate a larger body from fewer but larger cells (Fig. 3), as typical of many spring (vernal) geophytes (25) often called as ephemeroids. The cells and whole organs of ephemeroid plants are pre-formed during the photosynthetically inactive period in preceding season to avoid the mitotic inhibition induced by low temperatures in spring (19, 25, 26, 27, 28). After sprouting in the spring, ephemeroids grow mostly by expansion of cells rather than by cell division. As the growing season of these plants is not longer than three months when enough water is available, they do not invest much energy in building an organ skeleton: instead of producing thick cell walls, they maintain sufficient turgor pressure in their cells to preserve their body forms. In late spring, when water becomes less available, it may become difficult for plants to maintain enough turgor pressure, their leaves wither, and they finish growing season with bearing fruits (27).

The fact that geophytic species with smaller genome sizes could flower at any time during the growing season (Fig. 2), while species with increased genome size are conspicuously concentrated only in the spring-month categories of flowering, clearly indicates that this is not geophytism itself which is responsible for the increase of genome size but rather that the evolution of giant genomes in plants is associated with certain ecological conditions related to water availability. From the point of view of giant genomes, the major advantage of geophytism is the ability to replicate DNA and to pre-divide cells during the dormant period that helps to overcome the difficulties with slow DNA replication associated with large genome size. Moreover, such produced cells are larger because of existence of some functional constraint on the minimum cells size at a given DNA content (11, 40). To have large cells which are only pumped with water in spring may even be advantageous and positively selected in order to quickly develop a large functional body irrespectively of the initial nutrition, insolation and temperature conditions. This is a favourable strategy in many temporally limited habitats that would allow principally large genomes to evolve elsewhere in environments with enough long favorable conditions allowing cell division during the dormant and photosynthetically inactive period (i.e., avoid the occurrence of large genomes in alpine geophytes, such as *Crocus* or *Veratrum*, because of long-lasting frosts). Beyond the positive selection on cells size, it cannot be, however, excluded that genome gigantism in
certain geophytes is the outcome of a neutral process and a passive consequence of some common molecular force (such as unconstrained retrotransposon proliferation) that is passively tolerated until reaching some metabolic limits of a plant or resource capacity of the environment. This view would be well in accordance with increase of giant genomes observed with evolution of parasitism in some plants (e.g., *Viscum*) enabling them to become relatively independent on the resource limitation of the external environment.

As our data showed, the reason why large genome sizes are not more widespread in geophytes is anyway related to water availability and perhaps with water sensitivity associated with extremely large genome size. Water status of a plant is not only related to the water availability in the environment, but also critically depends on the preventing its lost thorough stomata during gas exchange and carbon dioxide uptake for photosynthetic processes (34, 55), that may be very critical for the fleshy design of plant bodies in many geophytes. This process is sensitive to the regulability and fast response of stomata to changing leaf water status that might be simply mechanistically related to the stomatal size and design (37, 56, 38). Experiments in deciduous trees show that larger stomata are slower to close and have a potential to cause hydraulic dysfunction under drought conditions (57, 37). A negative role of stomatal size may be seen also in our data indicating that the length of growing season shortens with increasing stomatal size in 68 analyzed geophytes (Table 1). In addition to the slower response a further disadvantage of large stomata is apparently in the less effective CO₂ uptake given mechanistically because of larger distance that a molecule have to diffuse thorough stomatal pore being larger and deeper in large sized stomata (56, 38). To achieve a sufficient CO₂ influx at a given total stomatal pore area, plants with larger stomata have to retain their stomata open for longer time which necessarily increases the sensitivity of such plants to periods of droughts forcing stomata closure. Hence, species with larger stomata may be expected generally to be less tolerant to water stress and selection of smaller stomata improving leaf water use efficiency is also commonly observed as a response to drought treatments (58, 59) or treatment with abscisic acids, a plant hormone released under water stress (60). Some useful adaptation to overcome the difficulties with CO₂ uptake and water loss may be prostrate leaf growth that developed in several lineages of South African geophytes (61). This morphology supports favorable moisture conditions underneath the leaves and these plants may even utilize CO₂ produced by soil microorganisms prospering in such conditions (62).

Because of the photorespiration and loss of CO₂ fraction are increased at higher temperatures, the activity of geophytic species is concentrated to cooler seasons and not to warm and moist periods of the year (63). With the tendency of decrease of CO₂ uptake effectivity species with large genomes and stomata may be therefore expected to direct their activity towards the cooler seasons and to finish their growing season before high temperature periods. Indeed, we observed a significant negative correlation between the length of growing season and stomatal size in 68 analyzed geophytes (Table 1), indicating that despite of ± equal start of active growth in all geophytes enabled with minimum temperatures, these are large genomed species which are forced to end growth earlier before incidence of high temperature periods.

Both slower stomata closure and decreased efficiency of CO₂ uptake generally limits the success of actively growing plants with extremely large genomes and stomata in arid or short-term volatile humid climate. Noting the unexpectedly tight correlation of stomatal and genome size observed in giant genomed plants, the drought sensitivity associated with presence of extremely large stomata may be therefore viewed as the most limiting factor in the evolution of extreme genome sizes and might set up the upper limits for maximum genome size for a plant in a given environment. This is also perhaps a reason why extremely large genomes are more frequently found in geophytes from habitats where cooler climate
coincides with long stable humid periods, such as occurs in temperate forests or Mediterranean mountains.

**Genome size and GC content**

For all 219 analyzed species GC content ranged from 35.75 % in *Allium ursinum* (Alliaceae) to 49.73 % in *Streptopus amplexifolius* (Liliaceae), on the scale 2C ranging from 582.18 Mbp to 128 273.07 Mbp (Table S2). Spearman’s Rho test of the data showed no significant relationship between both genomic measures (Table 1) but a significant negative correlation was shown to exist between genome size and GC content by using analysis of PICs (Table 1). The more detailed analysis of the data set suggests that the fail of the conventional statistics to reveal significant correlation might result from the existing unimodal relationship between genome size and GC content with peak approximately about 18 400 Mbp (Fig. 4). For species below this genome size threshold (126 species), a positive trend between genome size and GC content is found (Spearman’s Rho: $r_s = 0.334; p < 0.001$), although not well supported by the analysis of PICs ($r = 0.165; p = 0.134$). In species with genomes larger than 18 400 Mbp (93 species), the negative correlation retained (Spearman’s Rho: $r_s = −0.219; p = 0.035$; PIC: $r = −0.431; p < 0.001$). In general, the genomic GC content showed no clear association with any of the tested ecological variables (Table 1) but some trends were observable between GC and soil moisture, GC and soil nitrogen by using non-parametric correlation (Table 1). However, these correlations were mediated only with the correlation of GC content with genome size and completely disappeared when analyzed with PICs (Table 1).

Our findings contradict the reported absence of association between genome size and GC content by Barow & Meister (50) conducted on a limited sample of angiosperms. The positive correlation between genome size and GC content is generally not surprising, as the similar trends has already been reported in bacteria (51), vertebrates (64), and in some small genomed genera of *Poales* (53, 54). However, the significant trend of decreasing GC with increasing genome size is unusual, so far being known only in teleost fishes (52). This negative trend between genome size and GC content in large genomes might be explained by three hypotheses. (i) This might result from constraints on chromatin condensation which tends to be generally more condensed in larger genomes (65). This condensation might be perhaps facilitated in CG-poor DNA known to have higher curvature and ability to form non-linear structure over larger DNA regions (66, 67). (ii) The synthesis of GC base pairs is economically more expensive compared to synthesis of AT base pairs (68) so that extremely large genomes might be selected to be formed from AT-rich bases not to spare with cell energetic resources. (iii) Due to the stronger stacking interactions and triple bounding GC base pairs are much stable compared to AT base pair (69). Therefore, species growing later in the season and striving from higher temperature and UV stress might favor GC-rich over AT-rich structure of their genomes. The third hypothesis well agree with our observation that genomes of early flowering geophytes are larger and generally also much AT-rich compared to those small genomed and GC-rich flowering later in the season. Nevertheless, the causality between the expected DNA stability and plants ecology, as well as the two remaining hypotheses, needs to be exactly tested.

**Seed mass, genome size and ecology**

Seed mass was measured for 57 plant species. The comparison of their genome sizes showed no evident trend. This result is similar to analysis by Beaulieu et al. (48) showing only very weak relationships between both parameters on a much robust dataset. Compared to stomatal size, seed mass seems to be less constrained by genome size and much tightly driven by ecological and function constraints. This conclusion can be supported by significant correlations present between seed mass and ecological indicator values (EIV) for light
(negative; Table 1) and continentality (positive; Table 1). In both cases, ecological indicator values for the studied species had no phylogenetic signal, hence only results of conventional statistics were discussed. Nevertheless, the detected correlations may have a reasonable ecological explanation: germinating (non-parasitic and non-mycotrophic) plants can draw energy only from reserves stored in the seeds. This could be a favourable strategy for successful seedling development in extreme environments with limited or unpredictable supplies of energy and nutrients before the seedling become able to utilize energy and nutrients completely from external environment (i.e., develops enough leaf and root system). This may be the case of shady (low light input) and continental biotopes (low minimum temperatures and low precipitations), where species are known to have generally heavier seeds (70, 71, 72). Analogous results were also reported by Alexander et al. (73), who found a positive relationship between seed mass and altitude.

Although we did not find significant correlation between seed mass and genome size among studied species, a positive relationship between both these variables was documented in certain groups of related taxa (e.g., a genus or family; e.g. Allium, Crepis, Pinus; cf. 48: Table 1, 18: Fig. 5A), where the seeds share the same general design and similar ecological conditions shape their seeds size. In such a case, there exist numerous examples where seed size of closely related species differing in genome size has been used as a species discriminatory character in plant taxonomy (e.g., in polyploid relatives such as between diploid Stellaria pallida and tetraploid S. media, or between Ornithogalum umbellatum and O. divergens; 74).

Conclusions

Presented data confirmed that the evolution of extreme genome sizes in geophytes is closely mirrored with their phenology and ecology. Although large plant genomes are present in species with geophytic life-form more frequently than in species of any other life-form, this phenomenon does not perhaps implicate that the geophytic life-form itself is primarily responsible for the increase in genome size. It seems, however, that the necessary preclude behind origin of large genomes is providing enough reserves and allowing enough time for cell division by its placing to a period of unfavorable climatic conditions for plant growth. By other words, the increase of genome size and corresponding presence of large cells could be an advantageous evolutionary strategy to quickly exploit and successfully compete in seasonal environments by means of rapid development of plant body by expansion of predetermined cells. At the same time, however, there seems to be a trade-off between genome size and stomatal size that increases sensitivity to droughts and limits the appearance and activity of geophytic species with extremely large genomes to temporary stable humid periods and environments. The genome size and phenology of geophytes also associates, surprisingly, with remarkable changes in the overall DNA base composition. While GC content seems not to be associated directly with plant ecology, the unimodal relationship between genomic GC content and genome size indicates that some universal constraints may operate on the DNA composition of large genomes. The identification of these constrains and evaluation their role in evolution of genome size makes a challenge for future genomic research.

Material and methods

Sampling and cultivation

In total, 219 species were sampled (Table S2). They were collected mostly from wild populations in the Czech Republic, Slovakia, Italy, France, Portugal, Romania, Bulgaria, Turkey, and Israel. Some samples were obtained also from collections in the larger Czech
Flow cytometry
Fresh leaf material was taken from plants in cultivation, and the genome sizes and GC contents were estimated for each species using flow cytometry with internal standardization (ML, Partec GmbH). Intercalating propidium iodide or AT-selective DAPI dyes were used in a two-step procedure with Otto I and Otto II buffers (75). Simultaneous measurements with intercalating and AT-selective fluorochromes were used to estimate the AT and GC contents, according to the methodology by Barow & Meister (50). Detailed sample preparation and dye concentrations follow Šmarda et al. (54). As primary standards we used the ones recommended by Doležel & Greilhuber (76). In contrast to the original values derived from the comparison with human sequence (based on human 2C = 7 pg), their genome sizes and GC contents were derived from the comparison with the fully sequenced cultivar of *Oryza sativa* 'Nipponbare' (2C = 777.64 Mbp, GC = 43.6 %, genome cover 95 %), a gold standard with the most complete genomic sequence of an angiosperm species to date. Genomic parameters of primary standards were calculated directly from measurements with rice (for *Solanum lycopersicum*). For primary standards with higher genome sizes they were derived from the standard sample ratios with respective dyes given by Doležel et al. (77). Using this way human cells shows a genome size of 6 055.03 Mbp (6.19 pg), that is very close to the DNA content of human expected by Human Genome sequence Consortium (78; 6 153.33 Mbp) and the estimates of human genome size with alternative biochemical methods (cf. 76). We avoided to use human genome as a gold standard because of the still uncertain completeness of its repetitive portion in the genome sequence and because of possible effect of isochore structure and unusual GC content patterning of its genome to the binding ability of used dyes, that might differ from that in plant genomes. In the case of peak overlaps between standard and sample a series of secondary standards chosen from available samples was also established (Table S1). In the supplementary data, we provide rates of the internal standard of measurements with propidium iodide and DAPI. The base contents were calculated in a spreadsheet by Šmarda et al. (54) available at http://sci.muni.cz/botany/systemgr/download/Festuca/ATGCFlow.xls.

Phenology of flowering and growth
In total, the data on flowering phenology were analyzed in 164 species of geophytes. For 134 species they were excerpted from the compendium Flora d’Italia (79) comprising regionally standardized data to majority of analyzed species (134, i.e., 82 %). For the remaining 30 studied species not included in the Pignatti’s Flora, but cultivated or native to the Czech Republic (74), comparable data on flowering phenology were extrapolated from a comparison of data on species common to both regions using a logarithmic regression. The phenology of the species in the Czech Republic was either observed in the Botanical Garden of the Faculty of Science, Masaryk University, Brno, or excerpted from the Czech flora (74). The extrapolated data were rounded to be consistent with those in the Pignatti’s Flora. The duration of the growing season (i.e., the time since sprouting to the senescence of the photosynthetic apparatus) for 93 European geophytic species was observed in the field in the surroundings of Brno and in the Botanical Garden of the Faculty of Science, Masaryk University, Brno.
Stomatal length measurement
Stomatal size is a species-specific trait with a limited variation that could be induced by environmental conditions (80) or endopolyploidy, as common in cells of other plant tissues (81, 82). For the majority of samples, we measured stomatal size from surface impressions by means of a microlrelief method: a transparent nail polish was applied to the leaf surface in a thin layer, and when the polish dried, the layer was removed with adhesive tape and placed on a microscope slide. Slides were observed with the Olympus BX-51 microscope, under 400–1000× magnification. Digitally documented slides were analyzed manually in the Olympus Cell^F program. The average stomatal length for each species was based on at least 30 measurements from 1–2 leaves.

Seed mass measurement
In species where fruit were available, the seed masses were determined. If possible, at least 30 dry seeds were weighed per species (Kern 770 analytical balances; Kern & Sohn GmbH) to estimate the average seed mass. Seeds from the latest harvest were preferred in the measurements. Dry single-seed fruits (achene) were weighed whole (i.e., with the pericarp), as the pericarp weight was negligible relative to the seed weight.

Ecological characteristics
To characterize species ecology we used Pignatti’s indicator values (83), describing the ecological requirements of species in 12-degree ordinal scale (light, temperature, continentality, soil moisture, soil reaction, and soil nitrogen). These ordinal data are a reliable approximation of in situ measured ecological parameters of species habitats (84) and are frequently used for approximation of species ecology in European ecology research. Pignatti’s values were preferred over the analogous Ellenberg’s (85) values, namely, because of their wide availability for majority of studied species.

Statistical and phylogenetic analyses
Data were analyzed by both classic and phylogenetic statistical methods using the programs Statistica 9 (Starsoft Inc., Spearman’s Rho correlation test, Linear regression) and Phylocom 4.1 (analysis of traits module, Phylogenetic independent contrast analysis [PIC]) (86). For the phylogenetic analyses of samples in Phylocom, an evolutionary tree was compiled for tested species (see supplementary data). The phylogeny used followed the APG 3 system (87, 88), and the end-taxa were arranged according to previous studies when available (87, 89). Clades without a clearly resolved phylogenetic history or unclear phylogenetic positions were treated as polytomies. In those cases, putative trees were drawn, based on a conventional hierarchy of taxonomical categories (e.g., species clustered in sections and sections in genera). Values of statistical tests were rounded to three decimal places.

Acknowledgements
We would like to thank Jana Božková, Patricia Cardet, Petr Halas, Martin Jiroušek, Helena Prokešová, Martin Smetana and Libor Ulrych for their help with the field collections, Vladimír Řehořek for help with species identifications and Daniela Bártová and Ivana Hralová for constructive discussions. We would like to thank the Botanical Garden of the Faculty of Science, Masaryk University, Brno, the Botanical Garden of the Faculty of Science, Charles University, Prague and the Prague Botanic Garden, Prague, for providing plant material. This study was supported by the Ministry of Education, Youth and Sports of
the Czech Republic (Projects no. MSM0021622416 and LC06073) and Czech Science Foundation (Projects no. P506/11/0890 and GACR206/08/P222).

References


Figure and Table legends
Fig. 1. The relationship between logarithmed genome size and stomatal length in geophytes.
Regression lines equations (lines not in the graph): \( y = 0.1761 \times + 0.8492; \ r = 0.580; \ p = 0.000 \), small to large genomes \( (2C < 40,000\text{Mbp}; \ y = 0.1434 \times + 0.971; \ r = 0.448; \ p = 0.000) \) and extremely large genomes \( (2C > 40,000\text{Mbp}; \ y = 0.5342 \times - 0.8289; \ r = 0.545; \ p = 0.003) \).

Fig. 2. The relationship between the end of the flowering period and genome size.
Kruskal-Wallis H-test \( (9; 164) = 21.985; \ p = 0.009 \). Median of genome size = \( 24826.817 - 2185.274 \times \) end of flowering.

Fig. 3. The contrasting stomatal size of the spring geophyte *Gagea lutea* (a) and summer geophyte *Anthericum ramosum* (b).

Fig. 4. The relationship between genome size and genomic guanine + cytosine percentage in geophytes. (Genome size is logarithmed.)

Table 1. Results of statistical tests for analyzed traits. Statistically significant (at 0.05 level) results are in bold. \( r_s \) – Spearman’s Rho correlation coefficient; \( a \) – linear regression with logarithmed values; \( b \) – trait had no phylogenetic signal.