Genome sizes and genomic guanine+cytosine (GC) contents of the Czech vascular flora with new estimates for 1700 species

Velikost genomu a genomický obsah guaninu a cytosinu (GC) českých cévnatých rostlin s novými měřeními pro 1700 druhů

Petr Šmarda[1], Ondřej Knápek[1], Alexandra Březinová[1], Lucie Horová[1], Vít Grulich[1], Jiří Danihelka[1,2], Pavel Veselý[1], Jakub Šmerda[1], Olga Rotreklová[1] & Petr Bureš[1]

[1]*Department of Botany and Zoology, Masaryk University, Kotlářská 2, CZ-61137 Brno, Czech Republic, e-mail: smardap@sci.muni.cz, ondrej.knapek@gmail.com, sasa.brezinova @seznam.cz, lhorova@email.cz, grulich@sci.muni.cz, danihel@sci.muni.cz, pvesely@ sci.muni.cz, jsmerda@mail.muni.cz, orotrekl@sci.muni.cz, bures@sci.muni.cz;* [2]*Institute of Botany, The Czech Academy of Sciences, Zámek 1, CZ-252 43 Průhonice, Czech Republic*

Šmarda P., Knápek O., Březinová A., Horová L., Grulich V., Danihelka J., Veselý P., Šmerda J., Rotreklová O. & Bureš P. (2019) Genome sizes and genomic guanine+cytosine (GC) contents of the Czech vascular flora with new estimates for 1700 species. – Preslia 91: 117–142.

The content of DNA in the somatic, unreplicated cell nuclei (genome size) and DNA base composition (GC content) are the basic genomic parameters that can be measured by flow cytometry. Genome size, or ploidy level, can affect many plant properties and are therefore included as important features in modern biological floras and plant trait databases. However, genomic data are still largely measured mainly for taxonomic and biosystematic purposes, and despite the popularity of flow cytometry in the Czech Republic, this information is still missing for many of the vascular plants in the Czech flora. The biological significance of the GC content is less clear compared to the genome size, which, along with the greater difficulty in measuring it, results in the absence of such information for the vast majority of species. Here, we measure these two genomic parameters for 1908 samples of 1700 species in the Czech vascular flora. Here for the first time are reported the genome sizes of more than 1000 species and GC contents of more than 1500 species, which more than doubles the amount of information on the GC content of vascular plants. Together with the published data obtained in our laboratory using the same methods and flow cytometers, this information is now available for 1910 species that occur in the Czech Republic (~83% of this country's permanent flora, excluding apomicts). They are summarized in a table, accompanied by information about assumed chromosome number, ploidy level and an estimate of the monoploid genome size. We further provide a descriptive analysis of this dataset, list extreme values and comment on some cytogeographically interesting findings. This dataset is the largest and most comprehensive set of genomic data covering almost the entire flora of a country. It will serve as the basis of the karyological traits section of the Czech plant trait database Pladias (www.pladias.cz) and for testing of hypotheses about the evolution and biological relevance of these genomic parameters.

K e y w o r d s: chromosome number, Czech Republic, DNA base composition, DNA content, flow cytometry, GC content, genome size, life form, polyploidy, vascular plants

Introduction

The somatic, unreplicated nuclear DNA content (2C-DNA content, 2C genome size) is one of the basic genome characteristics that strongly varies in vascular plants (Leitch et al. 2005, Leitch & Leitch 2013). Not only does DNA conserve information about genes and their expression, the DNA amount itself has several passive effects on cellular form and function. The more DNA a cell has, the more space is required for its storage and transcription, and the more time is needed for its replication. Genome size, therefore, predetermines whether a plant will manage all the necessary DNA replications (cell divisions) within one season or whether it must be a perennial (Bennett 1987). The minimum size of a cell dictated by genome size affects the function of specialized cells, structure of some tissues and, consequently, to some extent also the general appearance of a plant (Knight & Beaulieu 2008, Greilhuber & Leitch 2013). Plants with large genomes tend to have large pollen (Knight et al. 2010), large stomata (Beaulieu et al. 2008, Hodgson et al. 2010, Veselý et al. 2012) and large seeds (Beaulieu et al. 2007). Having large cells, they frequently have rather thick leaves (Stebbins 1950) and are neither trees nor shrubs, probably because of the negative effect of large cells on the mechanical properties of woody tissue (Stebbins 1938). The genome (nucleus) also contains a significant portion of cellular phosphorus (Sterner & Elser 2002), composing the phosphate backbone of DNA. Consequently, large genome size is believed to increase nutrient requirements and, therefore, to limit success in phosphorus limited environments (Leitch & Leitch 2008, Greilhuber & Leitch 2013, Šmarda et al. 2013, Guignard et al. 2016).

Because of its important ecological consequences, information about genome size and/or ploidy level is included in modern biological floras and trait databases such as the Ecological Flora of the British Isles (Fitter & Peat 1994; http://www.ecoflora.co.uk), the Database on Biological and Ecological Traits of the Flora of Germany (BiolFlor; Klotz et al. 2002, www.biolflor.de) and the Global Plant Trait Database (TRY; Kattge et al. 2011, http://www.try-db.org). Information about genome size will also be included in the Database of the Czech Flora and Vegetation (Pladias; www.pladias.cz, Wild et al. 2019). However, there is no genome size data for many widespread species, and the data reported in the literature are sometimes difficult to use for a variety of reasons (see below).

More detailed quantitative information on DNA provides the content of the individual DNA bases: adenine (A), thymine (T), guanine (G) and cytosine (C). The early measurements of the DNA base composition using biochemical methods revealed similar contents of the G and C, or A and T bases (DNA base pairing rules; Chargaff et al. 1949, 1952, Chargaff 1951), which enhanced the discovery of the double helix structure of DNA (Watson & Crick 1953). The GC base pair differs from the AT base pair by a triple hydrogen bonding and stronger stacking interaction between the bases (Yakovchuk et al. 2006), which results in a higher thermal stability of GC-rich DNA when compared with AT-rich DNA (Marmur & Dotty 1962, Mandel & Marmur 1968). This increased thermal stability provides plausible explanation for the high GC contents in the DNA of thermotolerant microorganisms (Nishio et al. 2003, Musto et al. 2006, Basak et al. 2010) or in warm-blooded compared to cold-blooded vertebrates (Bernardi 2000). Nevertheless, several other theories can account for the current diversity in GC content in different groups of organisms (Mann & Phoebe Chen 2010, Šmarda & Bureš 2012). In monocoty-

ledons, for example, increased GC contents are recorded in species from dry and/or cold regions, suggesting a relationship between increased GC content and the tolerance of cells (nuclei) of drought- or freezing-induced desiccation (Šmarda et al. 2014). The biosynthesis of the GC base pair is slightly more energy demanding than that of the AT base pair (Rocha & Danchin 2002), and GC-rich DNA is also less compact than AT-rich DNA (Vinogradov 2003). These relationships provide possible explanations for the low genomic GC contents of plants with very large genomes (Veselý et al. 2012, Šmarda et al. 2014). Due to the greater difficulty of measuring the GC content using flow cytometry (see below), however, GC content data are much scarcer than that on genome size, which accounts for our poorer knowledge of the biological relevance of GC contents in plants.

The measurements of genome size started in the early 1950s (Swift 1950) and the published information is collected in the Plant DNA C-value Database (Bennet & Leitch 2012; http://data.kew.org/cvalues). The latest version of this online database includes information on 8510 species of land plants (Garcia et al. 2014). Since most of these data were collected in Europe, they also include a number of species in the Czech flora. Nevertheless, genome sizes remain still unknown for many widespread species in this country's flora. In some cases, the taxonomic identity of the plants in the Plant C-value Database is uncertain and the reported genome size cannot be unequivocally attributed to the populations found in the Czech Republic. In addition, some earlier measurements were made with error-prone methods (Greilhuber 1998, 2005) and may be used only with care. Even in the case of flow cytometry (FCM), which is currently the most widespread method for measuring genome size (Doležel et al. 1992, 2007, Doležel & Bartoš 2005, Greilhuber et al. 2007), the results obtained by different laboratories can vary by several percent for various technical or methodological reasons (Doležel et al. 1998, Greilhuber et al. 2007, Hoffmann & Wood 2007). Another bias may originate from using different genome size values for the FCM genome size standards.

Although the measuring of GC content and genome size started more or less at the same time, the measurements of GC content data are much scarcer (Meister & Barow 2007). According to our working database, GC contents are known only for about 1200 species of vascular plants, of which more than half was recently measured in our laboratory using FCM (Bureš et al. 2004, Šmarda et al. 2008, 2012, 2014, Veselý et al. 2012, Lipnerová et al. 2013, Veleba et al. 2014, 2017). However, due to the rather broad phylogenetic focus of most studies, these measurements are for only a very limited number of Czech species. As for the genome size data, the methods used to estimate GC content are not always reliable, and values reported in literature cannot be accepted without care; this also includes data from "whole" genome sequencing projects, which are actually hardly ever complete (Šmarda & Bureš 2012, Rieber et al. 2013, Veeckman et al. 2016). Nevertheless, the FCM measurements of GC content also suffer from its own methodological difficulties (Šmarda et al. 2012), which makes it difficult to compare published data obtained using alternative FCM methodology. The major problems of FCM measurements lie (i) in the approximate nature of the FCM estimates based on the assumption of random base patterning in genomes and (ii) the supposed specific behaviour of the base specific DNA dyes used (i.e. binding to a constant number of consecutive base pairs of the same type), which are unlikely to be completely met in real genomes; moreover, (iii) FCM measurements of GC content are usually done by comparing results of measurements using two DNA dyes (see Methods) and thus cumulate errors of both in the final

GC content estimate. For all these reasons, the reported differences in GC content of up to ~1% among related species or up to ~2% among distantly related ones must be interpreted with care (Šmarda et al. 2012).

Flow cytometry was established as an effective method for estimating the genome size of plants during the 1980s (Doležel et al. 1989, 2007). In the Czech Republic this method became widely used in plant systematics particularly during the last fifteen years. Until 2000 there were only a few cytometers available for occasional ecological, taxonomical or phylogenetic research in the Plant Cytogenetics Laboratory in the Institute of Experimental Botany of the Czech Academy of Sciences in Olomouc (J. Doležel Lab; Doležel & Göhde 1995, Doležel et al. 1998, Lysák & Doležel 1998, Suda & Lysák 2001, Šmarda et al. 2005). Nowadays, over 30 flow cytometers (to our knowledge) are used in leading Czech botanical institutions for studies on polyploid evolution, genome size, genomic GC content or breeding systems (e.g. Bureš et al. 2004, Suda et al. 2005, 2007, Šmarda & Bureš 2006, Šmarda et al. 2008, 2013, 2014, 2016, 2018, Chrtek et al. 2009, Kolář et al. 2009, 2014, Duchoslav et al. 2010, Kubešová et al. 2010, Zedek et al. 2010, Šarhanová et al. 2012, 2017, Veselý et al. 2012, Lipnerová et al. 2013, Hanzl et al. 2014, Kabátová et al. 2014, Otisková et al. 2014, Veleba et al. 2014, 2017, Vít et al. 2014, Lepší et al. 2015, Ekrt & Koutecký 2016, Kobrlová et al. 2016, Čertner et al. 2017, Mandák et al. 2017). Since 2003, 173 flow cytometry papers on these topics have been published by authors from the Czech Republic (17.1% of all papers published in 2003–2017), thus the Czech Republic is the world leader in this field of research (Fig. 1, Electronic Appendix 1).

We have taken advantage of the current favourable position of our laboratory in terms of equipment and personnel to reduce the lack of data on vascular plants in the Czech flora by measuring the genomic parameters in a large number of species. The goal was to prepare comprehensive and cross-comparable data on genome sizes, genomic GC contents and associated metrics (ploidy level and monoploid genome size) of Czech vascular plants. This work will (i) serve as the base for the genomic and karyological traits section of the Database of the Czech Flora and Vegetation (Pladias; www.pladias.cz, Wild et al. 2019) and will (ii) enable the testing of hypotheses about the evolution and biological relevance of these parameters.

Materials and methods

The Czech Republic and its flora

The Czech Republic is a landlocked country in central Europe with an area of 78,870 km^2. The altitudes span from 115 m to 1603 m a.s.l. Major parts of this country are hilly and situated at middle altitudes, but there are lowlands in the basins of large rivers, and mountains mainly along its borders. Most of the bedrock in this country is acidic, especially at high altitudes, but there are also areas with a base-rich bedrock, including limestone, loess and some eruptive rocks, and at some places outcrops of serpentinite. The patches of base-rich bedrock cause a substantial increase in the diversity of the local flora. There are only a few small natural lakes in the mountains, but many shallow artificial fishponds that were established since the Middle Ages at low and middle altitudes, which have become a dominant feature of the landscape in some parts of this country. This country's climate is temperate oceanic in the west, gradually changing to temperate continental in
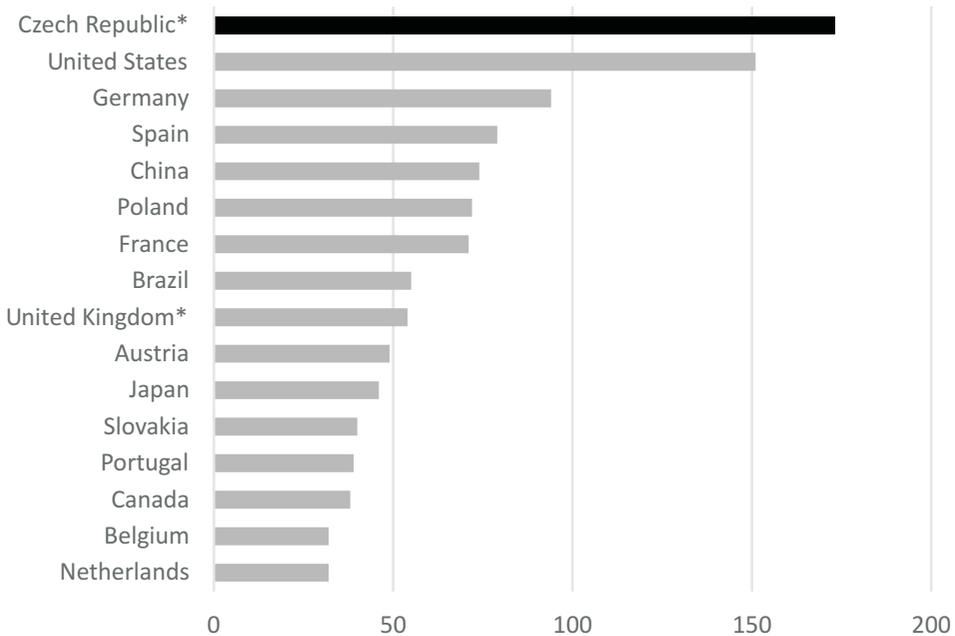
Fig. 1. – Countries with the highest numbers of flow cytometric studies on genome size, polyploidy or breeding systems in plants published in 2003–2017. This comparison is based on a search of the publications in Plant Sciences research area indexed by Thomson Reuters in the Web of Science database. Searching criteria considered also alternative country names that frequently appeared in author addresses (examples of searching formula for countries marked with asterisk are found in Electronic Appendix 1).

the east. Most of the lowlands are warm and dry, with a mean annual temperature of about 8.0–9.5 °C (January mean –2 to 0 °C, July mean 18–20 °C), while the highest areas in the mountains have a mean annual temperature of 1–2 °C (January mean –7 to –6 °C, July mean 8–10 °C). Precipitation is positively correlated with altitude: the lowlands receive 400–600 (–800) mm of precipitation annually, whereas in the highest areas in the mountains it is 1200–1400 mm (Chytrý 2017b).

The Czech Republic is situated in the zone of broad-leaved deciduous forests. Since the Neolithic Age, and especially since the Middle Ages, the majority of the country's surface was deforested and converted to pastures, arable land and meadows. Forests now cover about 34% of the country and frequently consist of plantations of conifers, especially at middle and high altitudes. In the driest and warmest lowlands in the north and south-east of the country, forest-steppe biome is assumed to be the extrazonal natural vegetation, at least in some places continuously persisting from the last glaciation. Extrazonal patches of coniferous taiga are preserved in the mountains with *Picea abies* and in the lowlands with *Pinus sylvestris* as dominant tree species. On the highest summits of the mountains in the north of the country, small remnants of alpine tundra are preserved. In the driest and warmest lowlands, mainly in the south-east, small patches of inland halophytic vegetation once occurred, but are now reduced to a few remnants (Chytrý 2017a).

The Czech flora is one of the best explored floras in the world, based on a long tradition of botanical research and its relatively low diversity (Kaplan 2017). According to the latest statistics (Danihelka et al. 2012, Danihelka 2013a, b), it comprises 2565 permanently present species of vascular plants (i.e. native plus naturalized and invasive archaeophytes and neophytes, and excluding 124 species considered as extinct or missing). Since the late 18th century, a further 872 species have been recorded as casual, escaped from cultivation or accidentally introduced. The five most species-rich families are *Asteraceae* (528 permanent species), *Rosaceae* (244), *Poaceae* (169), *Cyperaceae* (115) and *Fabaceae* (107). The 11 most species-rich genera include *Taraxacum* (215 permanent species), *Rubus* (121), *Carex* (83), *Hieracium* (51), *Pilosella* (40), *Veronica* (33), *Alchemilla* (23), *Potentilla* (23), *Festuca* (21), *Galium* (21) and *Viola* (21). The remarkable species richness of dominantly apomictic genera (the five genera listed here, *Taraxacum*, *Rubus*, *Hieracium*, *Pilosella* and *Alchemilla*, comprise altogether 450 permanent species, i.e. about 17% of the permanent flora) partly reflects the glacial and postglacial history of the Czech flora, but it may also partly be an artefact as these genera are particularly well explored in this country. The number of endemics in the Czech flora is rather small: Kaplan (2017) recently listed 82 endemic or subendemic species and subspecies (of which 45 are apomicts).

Investigated plants

Plants for the measurements were collected in the majority of cases in the wild in the Czech Republic. Rare and protected plants were sampled in botanical gardens or other specialized collections, namely the Collection of Aquatic and Wetland Plants in Třeboň (http://www.butbn.cas.cz/sbirkavk) and the living collection of endangered plants of the Krkonoše National Park in Vrchlabí. If known, the original locality of these cultivated plants is given in Electronic Appendix 2. Vouchers of measured plants are stored in the Herbarium of Masaryk University (BRNU). Altogether, 1908 samples of 1700 species were collected and measured (Electronic Appendix 2). The taxonomy and nomenclature follow Danihelka et al. (2012).

Along with new measurements, our survey also includes data for 225 species measured in our former studies (Bureš et al. 2004, Šmarda et al. 2008, 2014, Kočková 2012, Veselý et al. 2012, 2013, Lipnerová et al. 2013, Veleba et al. 2014, 2017), starting from Šmarda et al. (2008), obtained using the same method and flow cytometers as in the present study. Results from studies by Bureš et al. (2004), Šmarda et al. (2008) and Kočková (2012), which used different genome sizes and CG content values for the FCM standards, were recalculated to be fully compatible with the present data (Electronic Appendix 2).

Flow cytometry measurements

Genome size and genomic GC content were measured using flow cytometry and standard protocols consistently applied in our laboratory over the last decade (e.g. Šmarda et al. 2008). It is based on the widespread two step procedure of Otto (1990; further referred to as Otto) and combines measurements of a sample using this method and two different dyes: DNA intercalating propidium iodide (PI) for measurements of absolute DNA content, and the AT-specific 4',6-diamidino-2-phenylindole (DAPI) for calculation of the AT (or GC) fraction in the genome (Doležel et al. 1992, Barow & Meister 2002). Briefly,

Table 1. – Used FCM genome size and GC content standards. [1]Grown from seeds provided by the International Rice Research Institute, Philippines; [2]Grown from seeds provided by The Centre of Plant Structural and Functional Genomics (AVČR Olomouc, Czech Republic); [3]Plant grown in the Botanical Garden of the Faculty of Science, Masaryk University, Brno, Czech Republic; [4]Calculated from Mbp data based on the conversion formula 1 pg = 978 Mbp (Doležel et al. 2003), neglecting eventual minor difference of this conversion coefficient for plants with different GC content; [5]Whole genome sequencing including gap size estimation (International Rice Genome Sequencing Project 2005); [6]Veselý et al. (2012: Table 1); [7]Šmarda et al. (2014: Table S5); [8]Veleba et al. (2017: Table S2); [9]Calculated using the sample/standard fluorescence ratios given by Doležel et al. (1992: Table 1).

Taxon	Plant source	2C (pg)[4]	2C (Mbp)	GC (%)	Estimation method
Oryza sativa 'Nipponbare'	IRRI[1]	0.795	777.64	43.60	whole genome sequencing[5]
Carex acutiformis	wild clone	0.818	799.93	36.46	FCM with *L. esculentum*[6]
Raphanus sativus 'Saxa'	Doležel lab.[2]	0.997	975.55	40.30	FCM with *O. sativa*[7]
Lycopersicum esculentum 'Stupické polní tyčkové rané'	Doležel lab.[2]	1.735	1 696.81	38.72	FCM with *O. sativa*[6]
Glycine max 'Polanka'	Doležel lab.[2]	2.077	2 030.89	37.89	FCM with *L. esculentum*[8]
Bellis perennis	wild clone	3.159	3 089.89	39.54	FCM with *L. esculentum*[7]
Epipremnum aureum	commercial clone	7.991	7 815.39	42.70	FCM with *B. perennis*[6]
Pisum sativum 'Ctirad'	Doležel lab.[2]	8.018	7 841.27	41.77	FCM with *L. esculentum* + human + *Zea mays*[6, 9]
Vicia faba 'Inovec'	Doležel lab.[2]	23.796	23 272.88	41.15	FCM with *P. sativum*[6, 9]
Haemantus albiflos	BG MU Brno[3]	59.143	57 841.66	38.76	FCM with *V. faba*[6]

about 0.5 cm^2 pieces of young leaves of the sample and the standard (Table 1) were chopped together using a sharp razor blade in a Petri dish containing 1 ml of Otto I buffer (0.1 M citric acid and 0.5% Tween 20). Subsequently, an additional 1 ml of Otto I buffer was added. The crude nuclear suspension was filtered through a 50-µm nylon mesh. The filtered suspension was divided between two sample tubes to which either 1 ml of Otto II buffer (0.4 M Na$_2$HPO$_4$.12 H$_2$O) supplemented with DAPI (final concentration 2.0 µg/ml) or 1 ml of Otto II buffer containing PI (final concentration 50 µg/ml) was added. Both samples were measured simultaneously on two different flow cytometers, one (CyFlow SL or CyFlow ML, Sysmex Partec GmbH) equipped with a green, 532 nm, 100–300 mW, solid state laser (for measurements with PI; enabling the measurement of absolute genome size), and the other equipped with a UV light emitting diode (365 nm, Sysmex Partec GmbH) (for measurements with DAPI; enabling the measurement of the AT fraction of the genome, necessary for calculating the DNA base composition). This practice shortens the time needed for sample preparation and enables a partial correction of the possible bias between PI and DAPI measurements that may appear in some species due to differences in the content of metabolic compounds that interfere with the DNA dyes used. In general only a single plant was measured in each species sample. Measurements of each plant were repeated at least three times on different days and afterwards averaged. At least 5000 nuclei were analysed in each measurement, with peaks accepted generally only if their coefficient of variance was below 3.5 (measurements with DAPI) or 5 (measurements with PI). The maximum difference in average sample/standard ratios obtained for the same samples (four different pairs of FCM standards measured from 2011 to 2014) in each of the four flow cytometers used for measurements with PI was 1.024-fold, and 1.020-fold by comparing the three flow cytometers used for the measurements with DAPI (tested in the same way in 2012).

GC content was calculated from the flow cytometry data (sample/standard fluorescence ratio with DAPI divided by the sample/standard fluorescence ratio with PI) using the regula falsi method in the Excel spreadsheet by Šmarda et al. (2008; available at http://sci. muni.cz/botany/systemgr/download/Festuca/ATGCFlow.xls). Calculations of GC contents were done separately for each pair of DAPI + PI measurements (i.e. individually for each sample chopped in the same Petri dish) and only afterwards averaged. Given the base pairing rules (see above), the reported GC contents can be easily converted into AT contents: AT(%) = 100–GC(%), which are used instead of GC contents to express DNA base composition in some papers.

The first step of sample preparation and the composition of Otto I buffer were frequently modified for measurements of species containing metabolic compounds that may interfere with the DNA dyes used or in any other way prevent the proper staining of DNA and of obtaining accurate measurements using flow cytometry (Table 2, Electronic Appendix 2). In general, increasing the acidity by adding HCl, HNO_3, or acetic acid to the Otto I buffer helped to dissolve mucilaginous or other substances otherwise preventing proper nuclei release (altogether used for more than 700 species, namely from the *Rosaceae*, *Liliaceae*, *Crassulaceae* and *Boraginaceae*). The chemical reaction of some intracellular metabolites released during tissue cutting (namely when the nuclei suspension gradually darkens, as in *Myricaria* and *Geranium lucidum*) may be slowed down by cooling the sample during its preparation (using ice-cooled sample, Petri dish, buffers, sample cuvette and filter) and measurement (i.e. placing the cuvette with nuclei suspension in ice-cold water during the FCM measurement). In some species we also modified the Otto I buffer by increasing the concentration of the detergent (Tween 20) or occasionally also by using a different detergent (Triton X-100). If a strong background signal was present in FCM histograms, it proved effective to remove the debris using centrifugation (usually 4500 rpm for 15 min) and working further only with nuclei-containing pellet redissolved in Otto I buffer (used for example in *Azolla filiculoides*, *Daphne cneorum*, *Epilobium lamyi*, *Hylotelephium maximum*, *Jovibarba globifera*, *Lythrum salicaria*, *Malva sylvestris*, *Morus* spp., *Scilla kladnii*, *Sedum hispanicum* and *Tephroseris aurantiaca*). In species where the concentration of secondary metabolites was too high in the leaves, the measurements were done on tissues from other parts of a plant, which empirically enable the most accurate measurements; e.g. flower stalks (in *Fragaria*, *Viola* and *Hylotelephium maximum*), young stems (*Pulmonaria*, *Trapa* and *Illecebrum*), seeds (*Geranium molle* and *Omphalodes*), petioles (*Salix*) and bulbs (*Scilla*). In species with a high level of endopolyploidy we always searched for peaks of 2C, i.e. non-replicated nuclei in the FCM histograms, mostly by measuring the youngest, just developing leaves, developing inflorescence, young seeds (*Brassicaceae* and *Crassulaceae*) or young, immature ovaries (*Orchidaceae*; cf. Trávníček et al. 2015).

For the measurements we used the traditional FCM reference standards, as introduced and provided by the laboratory of J. Doležel in Olomouc (Doležel et al. 1992, 2007, Doležel & Greilhuber 2010) as well as a few others that have been used since then (Table 1): *Oryza sativa* (Šmarda et al. 2012), *Carex acutiformis*, *Haemanthus albiflos*, *Epipremnum aureum* (Veselý et al. 2012) and *Bellis perennis* (Morgan-Richards et al. 2004, Schönswetter et al. 2007). The genome sizes and GC contents of these standards were derived based on a comparison with that of the fully sequenced cultivar of *Oryza sativa* 'Nipponbare' (2C = 777.64 Mbp, GC = 43.6%). The actual size of the *O. sativa*

Table 2. – Methods used for preparing samples for FCM (some developed by L. Horová); for clarity the original amounts of chemicals used (providing one ±2 ml sample stained with PI and ±2 ml sample stained with DAPI) are recalculated here to enable preparation of a single ±2 ml sample. Usage and eventual modifications of these methods in particular species are indicated in Electronic Appendix 2.

lh-1	Chopping in 0.5 ml OTTO I buffer → 0.5 ml OTTO I buffer added → filtering → 1 ml OTTO II buffer added (standard method according to Doležel et al. 2007; see the text)
lh-2	Chopping in 0.5 ml acidified OTTO I buffer (OTTO I: 0.1M HCl 3:1 with three drops of Tween 20 added to each 2 ml of the buffer) → 0.5 ml of the modified OTTO I buffer added → filtering → 1ml OTTO II buffer added
lh-3	Chopping in 0.5 ml acidified OTTO I buffer (OTTO I: 0.1M HCl 1:1 with three drops of Tween 20 added to each 2 ml of the buffer) → 0.5 ml of the modified OTTO I buffer added → filtering → 1ml OTTO II buffer added
lh-4	Chopping in 0.5 ml acidified OTTO I buffer (OTTO I: 0.1M acetic acid 1:1 with three drops of Tween 20 added to each 2 ml of the buffer) → 0.5 ml of the modified OTTO I buffer added → filtering → 1ml OTTO II buffer added
lh-5	Chopping in 0.5 ml acidified OTTO I buffer (OTTO I: 0.1M HNO_3 1:1 with three drops of Tween 20 added to each 2 ml of the buffer) → 0.5 ml of the modified OTTO I buffer added → filtering → 1ml OTTO II buffer added
lh-6	Chopping in 1 ml 0.1M HCl with two drops of Tween 20 → filtering → 1 ml OTTO II buffer added
lh-7	Chopping in 1 ml 0.1M HNO_3 with three drops of Tween 20 → filtering → 1 ml OTTO II buffer added
lh-8	Chopping in 1 ml woody plant buffer (Loureiro et al. 2007) → 1 ml woody plant buffer added → filtering → 100 μl PI or 80 μl DAPI added
lh-9	Chopping in 1 ml Tris $MgCl_2$ buffer (Pfosser et al. 1995) with extra 100 μl PI or 80 μl DAPI (i.e. chopped directly with the dye) → filtering → further 1 ml of the Tris $MgCl_2$ buffer (without extra PI or DAPI) added
lh-10	Nuclei isolated and stained using CyStain PI kit (Sysmex) according to the manufacturer's instructions
lh-11	Chopping in 0.5 ml modified OTTO I buffer (0.1M citric acid with 1% Triton X-100: 0.1M HNO_3, 1:1) → 0.5 ml of the modified OTTO I buffer added → filtering → 1 ml OTTO II buffer added
lh-12	Chopping in 1 ml seed buffer (Matzk et al. 2001) (containing either 100 μl PI or 80 μl DAPI) → filtering → further 1 ml of the seed buffer added (without PI or DAPI but containing NaCl)
lh-13	Chopping in 0.5 ml modified OTTO I buffer (0.1M citric acid with 1% Triton X-100: 0.1M HCl, 1:1) → 0.5 ml of the modified OTTO I buffer added → filtering → 1 ml OTTO II buffer added
lh-14	Chopping in 1 ml (0.1M HCl with 1% Triton X-100) → filtering → 1 ml OTTO II buffer added
lh-15	Chopping in 1 ml (0.1M HNO_3 with 1% Triton X-100) → filtering → 1 ml OTTO II buffer added
lh-16	Chopping in Galbraith et al. (1983) buffer with extra 1% Triton X-100 and extra 1% PVP (pH = 9.5)
lh-17	Chopping in 1 ml (0.1M HCl with 2% Triton X-100) → filtering → 1 ml OTTO II buffer added

'Nipponbare' genome was carefully determined and the sizes of existing sequence gaps exactly estimated using various independent methods and included into its genome size estimate (International Rice Genome Sequencing Project 2005, Vij et al. 2006). The reliability of this estimate is supported also by a more recent *de novo* sequencing project, reporting a very similar genome size and GC content for another cultivar of *O. sativa* (2C = 780.64 Mbp, GC = 43.57%, Du et al. 2017).

The genomic parameters of internal standards were calculated directly from measurements with *Oryza sativa* (for *Solanum lycopersicum* and *Raphanus sativus*) or derived from measurements with *S. lycopersicum* (for *Carex acutiformis* and *Bellis perennis*) (Table 1; cf. Veselý et al. 2012, Šmarda et al. 2014). Genomic parameters of standards with large genome sizes were derived from the standard sample ratios using the corresponding dyes reported by Doležel et al. (1992: Table 1). It is always necessary to keep in mind that FCM results are still to some extent hypotheses about the actual genome size,

namely in species with rather large genomes. Here we depend on a comparison with reference FCM standards, the genome sizes of which were obtained by comparison with a different standard (Table 1) and the cumulative error of these cascade comparisons makes the exact genome sizes of these standards relatively uncertain. Any FCM study should therefore report the raw primary result, which is the ratio of sample/standard nuclei fluorescence, to allow for any future corrections of reported genome size estimates. Apart from purely methodological reasons, sometimes genome size cannot be exactly estimated for some taxa because of the existence of intra-specific variation in genome size (Šmarda & Bureš 2006, 2010, Trávníček et al. 2013).

Ploidy level, chromosome number and monoploid genome size (1Cx)

The major differences in genome size are due to differences in the content of non-coding, repetitive DNA represented mainly by retrotransposons (Bennetzen et al. 2005, Wendel et al. 2016). Nevertheless, genome size may be easily increased through polyploidy (i.e. whole genome duplication), which simultaneously multiplies also the content of all genes. Therefore, polyploidy is very important in plant evolution (Müntzing 1936, Stebbins 1950, 1971, Levin 2002, Soltis & Soltis 2012, Soltis et al. 2014). However, the saltation nature of polyploidy makes it difficult to be properly addressed in phylogenetic analyses of genome size evolution. Effects of recent polyploidization, therefore, used to be filtered out in genome size analyses. This filtering could be done most easily by using only the size of a single chromosome set (i.e. the monoploid genome size, 1Cx), which is calculated by dividing 2C genome size by the respective ploidy level (Greilhuber et al. 2005).

The size of one chromosome set (x) or the "base chromosome number", used for calculating the ploidy level, is generally derived from the lowest chromosome number recorded for a genus or a group of closely related genera (e.g. Raven 1975). This practice is, however, limited in taxa with holocentric chromosomes (of the Czech species the *Cyperaceae*, *Juncaceae*, *Drosera*, *Cuscuta* sect. *Cuscuta* and *C.* sect. *Grammica*), where chromosome number is not always positively correlated with ploidy level due to possible chromosomal fissions (agmatoploidy) or chromosomal fusions (symploidy; Bureš et al. 2013). It may also be difficult even in taxa with monocentric chromosomes if there is no diploid species within the group concerned due to extinction, taxonomical bias (i.e. delimitation of diploid progenitors and their polyploid descendants as separate genera), or if such a species has not been karyologically investigated. Here we attempt to overcome some of these difficulties by a joint analysis of chromosome number and genome size data. We consider a taxon as polyploid whenever both its chromosome number and its genome size are ± double (or other multiple) if compared with diploid taxa in the genus concerned or phylogenetically closely related genera. This practice enabled us to derive the ploidy level for monotypic or species-poor genera (e.g. *Asperugo*, *Glaux* and *Luronium*), for many agmatoploid species and for groups with limited karyological information, and derive the polyploidy status of several genera despite the absence of diploids (e.g. *Ajuga*, *Origanum*, *Soldanella*; Electronic Appendix 2). To handle the chromosomal fusions in *Luzula*, we additionally also used chromosome size categories as defined by Nordenskiöld (1951) to estimate the actual ploidy level. Ploidy level estimates in highly polyploid genomes of *Viola* follow Marcussen et al. (2015).

Information on chromosome numbers were obtained from volumes 1–8 of the Flora of the Czech Republic (Hejný et al. 1988–1992, Slavík et al. 1995–2004, Štěpánková et al. 2010) and the Chromosome Counts Database (Rice et al. 2015; http://ccdb.tau.ac.il). If several chromosome counts were available for a species, the ploidy of the measured sample was estimated by comparing its genome size with the genome sizes of its most closely related species for which the ploidy level was known, assuming similarity of monoploid genome sizes in closely related species. Whenever the genome size of a sample contradicted the chromosome numbers or ploidy levels recorded within the species, we repeated measurements on another sample of that species and had the identification of the plant confirmed by experts. If discrepancies between karyological and genome size data remained, we counted the chromosomes of the measured plants. These chromosome counts were done in root tips, using the rapid squash aceto-orcein method as described by Rotreklová et al. (2011), sometimes with pre-treatmentin ice cold water (Mandáková & Lysák 2016).

If there was no information on chromosome number for a given species and, in the same time, its ploidy could be derived from genome size, the chromosome number typical of the given ploidy in closely related taxa is indicated in Electronic Appendix 2; this was the case for seven taxa.

Results and discussion

General aspects and coverage of the study

Altogether we measured genome size and genomic GC contents for 1908 samples belonging to 1700 species of which 46 species are represented by two or more subspecies and/or cytotypes. For more than 1000 species, genome size is published here for the first time (according to the Plant DNA C-value Database; Bennet & Leitch 2012). Along with rare species, these new estimates include those for many widespread species, for which data were lacking due to the presence of specific secondary metabolites (particularly in *Rosaceae* and *Boraginaceae*), which prevent measurements using standard FCM procedures and buffers. GC contents are provided here for the first time for the vast majority of the species measured (> 1500), which more than doubles the number of existing GC content data for vascular plants. Considering also our former data, the genome sizes and GC contents are now available for 1910 species of the Czech vascular flora (Electronic Appendix 2).

Despite our efforts to provide a complete survey, there are many species not covered by our study. First of all, these include numerous apomictic microspecies of *Alchemilla*, *Hieracium*, *Pilosella*, *Rubus* and *Taraxacum* (only 77 of 495 species measured). However, monoploid genome sizes and GC contents of these taxa are very similar at least within particular sections or other infrageneric taxa and missing information may be substituted by values of the closest relatives. Thus the efforts and costs necessary for a complete coverage would not be counterbalanced by the benefit. Excluding the apomicts, the data presented in Electronic Appendix 2 cover ~83% of the permanent flora, including almost all widespread species. The missing species will be investigated in the near future and the results will be added to the existing dataset.
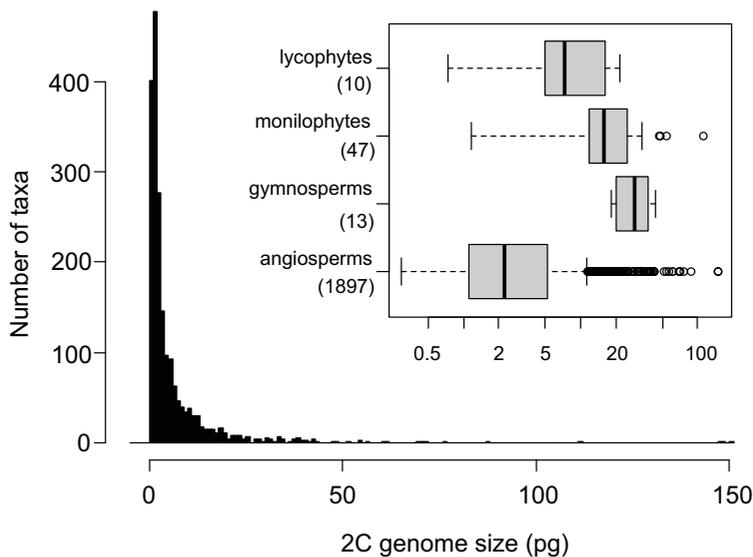
Fig. 2. – Distribution of 2C genome sizes in the genome dataset for Czech vascular plants (histogram) and separately for the four major phylogenetic groups (boxplots). Small genomes clearly prevail in the entire dataset making the distribution highly positively skewed. This is particularly because of the predominance of angiosperms (flowering plants) in the dataset, the genomes of which are smaller than those of gymnosperms, monilophytes and lycophytes. Boxplots show: median (thick horizontal line), interquartile range (box), range of non-outlying values (whiskers) and outliers (circles); numbers of taxa (including all subspecies and/or ploidy levels of a species) are in brackets below the group names.

Variation in genome size

The reported genome sizes of Czech vascular plants vary 505-fold (*Viscum album* compared to *Arabidopsis thaliana*). Small genomes clearly prevail in the dataset making the distribution highly positively (right) skewed (Fig. 2). This is particularly because angiosperms (flowering plants) are predominant in the dataset and their genome sizes are generally smaller than those of the non-flowering vascular plants (Fig. 2). The 10 species with the smallest genome are *Arabidopsis thaliana* (2C = 0.30 pg), *Groenlandia densa* (2C = 0.33 pg), *Spirodela polyrhiza* (2C = 0.35 pg), *Valerianella locusta* (2C = 0.36 pg), *Aruncus dioicus* (2C = 0.36 pg), *Cardamine impatiens* (2C = 0.36 pg), *Valerianella carinata* (2C = 0.37 pg), *Neslia paniculata* (2C = 0.37 pg), *Lepidium campestre* (2C = 0.37 pg) and *Valerianella dentata* (2C = 0.39 pg). Alltogether, 384 species in the dataset (about 20%), all of which are angiosperms, have genomes smaller than 1 pg. These belong to 60 families (Electronic Appendix 2), particularly the *Amaranthaceae*, *Betulaceae*, *Brassicaceae*, *Boraginaceae*, *Caryophyllaceae*, *Cyperaceae*, *Hypericaceae* (*Hypericum*), *Juncaceae*, *Lentibulariaceae* (*Utricularia*), *Onagraceae* (*Epilobium*), *Papaveraceae* (*Fumaria*), *Plantaginaceae* (*Veronica*), *Potamogetonaceae*, *Rosaceae*, *Salicaceae*, *Scrophulariaceae* (*Verbascum*), *Typhaceae* and *Valerianaceae* (*Valerianella*). The gymnosperms, monilophytes and lycophytes with the smallest genomes are *Chamaecyparis lawsoniana* (2C = 18.72 pg), *Azolla filiculoides* (2C = 1.20 pg) and *Selaginella selaginoides* (2C = 0.75 pg), respectively.
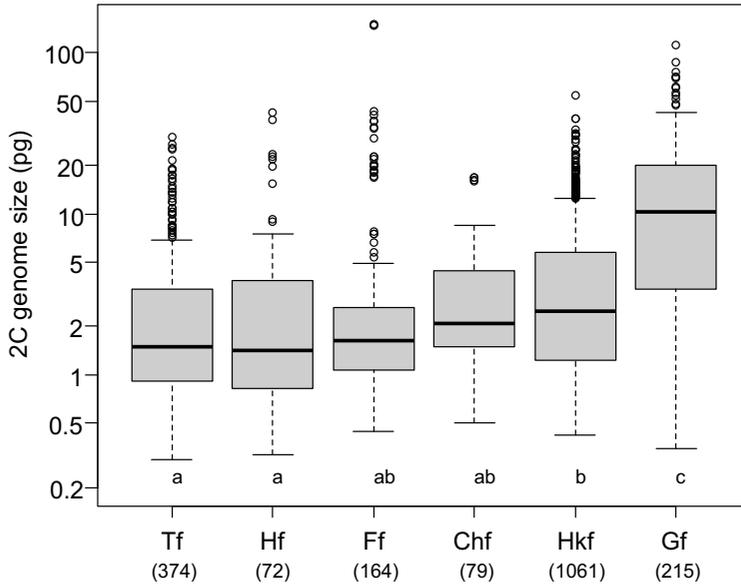
Fig. 3. – Distribution of 2C genome sizes of plants according to their life forms (Raunkiær 1934; information from Kaplan et al. 2019): Tf – therophytes, Hf – hydrophytes, Ff – phanerophytes, Chf – chamaephytes, Hkf – hemikryptophytes (incl. facultative therophytes), Gf – geophytes. While large genomes are typical of geophytes, small genomes prevail in therophytes, hydrophytes and woody plants. Boxplots show: median (thick horizontal line), interquartile range (box), range of non-outlying values (whiskers) and outliers (circles). Numbers in brackets show number of taxa (including all subspecies and/or ploidy levels of a species) in each category. Boxplots marked with the same letter at the bottom of the graph do not differ significantly (P > 0.05, Tukey HSD test with ln-transformed values).

Consistent with theory, annual species have small genomes (Fig. 3), as large genomes and the associated prolonged duration of mitosis, meiosis and cell division would prevent them from completing their development in one season (Bennett 1987). Nevertheless, in several phylogenetic groups with mainly large genomes, such as *Ranunculales* s. str., *Melampyrum* and the *Bromeae+Triticeae* clade of *Poaceae*, even annual species can have genomes that are relatively large, with the two most extreme sizes those for the hexaploid bread wheat, *Triticum aestivum* (2C = 30.7 pg) and tetraploid *Adonis aestivalis* (2C = 27.8 pg).

Small genomes are also typical of angiosperm trees and shrubs (macrophanerophytes and nanophanerophytes; Fig. 3). In woody plants, large genomes may have a negative effect on the size of cambial cells and mechanical properties of woody tissues (Stebbins 1938, 1950) or on the size of stomata and their maximum achievable conductance needed to maintain the transpiration stream in tall trees (Beaulieu et al. 2008). Remarkable exceptions to this trend are the obligate hemiparasites *Viscum album* (2C = 153 pg) and *Loranthus europaeus* (2C = 17.2 pg). In non-parasitic angiosperm phanerophytes, genome size rarely exceeds 7 pg, even that of the 22-ploid *Morus nigra* (2C = 6.7 pg). The only exceptions to this limit are the *Ranunculaceae* vines *Clematis vitalba* (2C = 17.5 pg) and *C. viticella* (2C = 18.3 pg) and woody *Adoxaceae*, with the genome size of *Sambucus nigra* (2C = 23.2 pg) and *S. racemosa* (2C = 23.1 pg) being the largest in the world for any tree-like or shrubby woody angiosperm.

The Czech plant with the largest genome is *Viscum album* subsp. *austriacum* (2C = 153 Gbp; Electronic Appendix 2). The 10 species with the largest genomes in our dataset are *Viscum album* (2C = 153 pg), *Ophioglossum vulgatum* (2C = 114 pg), *Paris quadrifolia* (on average 2C = 88.3 pg), *Lilium martagon* (2C = 78.1 pg), *Cypripedium calceolus* (2C = 72.7 pg), *Lilium bulbiferum* (2C = 72.1 pg), *Leucojum vernum* (2C = 70.7 pg), *L. aestivum* (2C = 63.0 pg), *Galanthus nivalis* (2C = 62.1 pg) and *Allium oleraceum* (2C = 58.0 pg). The largest genome in gymnosperms and lycopods is found, respectively, in *Pinus nigra* (2C = 44.6 pg) and *Isoëtes lacustris* (2C = 22.3 pg). Altogether 294 species (15%) of the taxa studied have genomes larger than 10 pg (Electronic Appendix 2). These are typical of non-flowering vascular plants (Fig. 2; except for aquatic ferns), geophytic monocotyledons (*Amarylidaceae*, *Alismataceae*, *Alliaceae*, *Asparagaceae*, *Butomaceae*, *Liliaceae*, *Melanthiaceae* and *Orchidaceae*) and *Ranunculaceae* subfamily *Ranunculoideae* (Electronic Appendix 2). Species with genomes larger than 10 pg are also common in *Adoxaceae*, *Asteraceae*, *Fabaceae* and *Poaceae*, some hemiparasitic groups (*Viscum*, *Loranthus* and *Melampyrum*), the herbaceous *Ericaceae* (*Moneses* and *Orthilia*) and in an incidental manner in a further 12 angiosperm families (Electronic Appendix 2).

In general, the association between large genome size and geophytic life cycle in angiosperms (Fig. 3), may be due to (i) the advantage of large cells (associated with large genomes) for fast development in spring (Grime & Mowforth 1982, Veselý et al. 2012) and/or (ii) the DNA synthesis required to replicate large genomes may continue when there is a continuous supply of resources from storage organs (Veselý et al. 2013).

Incidence of polyploidy

About 43% of the species in Electronic Appendix 2 are polyploid (42.8% when subspecies and/or different ploidy levels are included as species; 42.7% when only the major ploidy level is considered). This estimate seems to be slightly lower than the estimates of polyploid frequency in surrounding countries or in Central Europe as a whole (cf. Baquar 1976: Table 1, Grant 1981: Table 24.1). This could be due to the limited amount of information on polyploidy status that was available in the past as well as the lack of many apomictic species in our dataset, which are mostly polyploid. Another reason is that this estimate is based on data that also include the holocentric families *Juncaceae* and *Cyperacae* (both relatively poor in polyploid species), which used to be excluded from ploidy frequency calculations due to uncertainties about their actual ploidy level; such inference is possible only when the genome size is known (see Methods, and Lipnerová et al. 2013).

Despite being somewhat low, the reported frequency of polyploids in the Czech flora may be considered typical for a latitude of 50°N, which is higher than in the tropics or the Mediterranean regions but lower than in boreal and arctic regions (Rice et al. 2019). One of the complex reasons for this may be the relatively high proportion of herbaceous plants in the Czech flora, which are much richer in polyploids than woody and annual plants, which are more abundant in tropical and Mediterranean regions, respectively. In contrast, high latitudes (repeatedly glaciated in the past) may provide more free niches for polyploids to establish (Stebbins 1985), and the harsher climates of these regions may support a disproportionately larger immigration and survival success of polyploids than diploids, or even positively affect the frequency of polyploids origin (Brochmann et al. 2004).
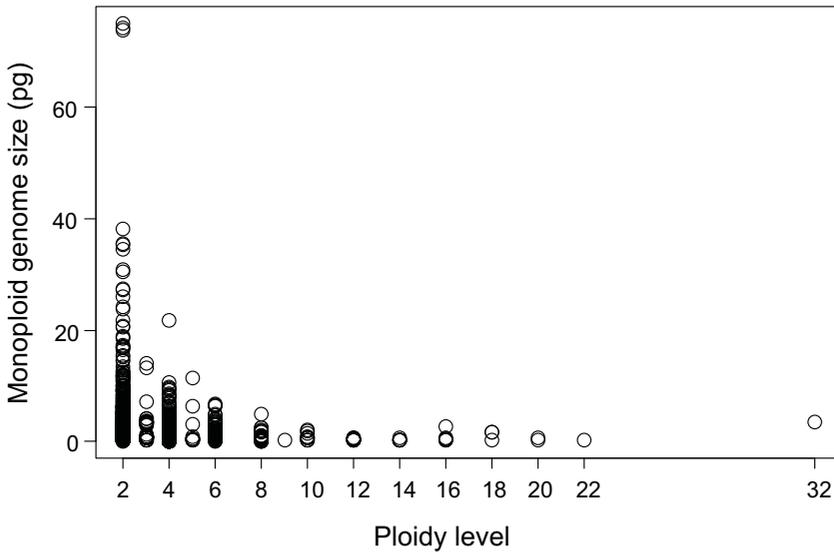
Fig. 4. Relationship between ploidy level and monoploid genome size, which indicates a decreasing tendency to form polyploids with increase in monoploid genome size.

The highest ploidy in the Czech vascular flora is found in *Ophioglossum vulgatum* (32x). Other strikingly high polyploids are *Morus nigra* (22x), *Rumex hydrolapathum* and *Viola lutea* subsp. *sudetica* (both 20x), *Alchemilla fissa*, *Helictochloa pratensis* and *H. planiculmis* (all 18x), *Cerastium fontanum*, *C. holosteoides*, *C. lucorum*, *Ranunculus lingua*, *Trientalis europaea*, *Trifolium pannonicum* and *Viola arvensis* (all 16x).

In plants with very large genomes polyploidy is generally rare. Ploidy level is therefore inversely correlated with monoploid genome size (Fig. 4, P < 0.001, Kendall non-parametric correlation test), which is in accordance with the observations of Grif (2000). This pattern may have resulted from the existence of certain, lineage specific or enviromentally driven, upper limits to genome size. Consequently, it is more likely that such a limit would be reached by polyploidization in plants with large genomes, whereas plants with small genomes do not reach it even after repeated polyploidizations. Though polyploidy is known also in plants with very large genomes (Hidalgo et al. 2017), the upper limits are more likely to be reached by a gradual increase in DNA via retrotransposon amplification, which provides species more time to adapt to the ongoing changes, rather than by a sudden increase in genome size via polyploidy. A newly arising polyploid plant must immediately cope with all genome size effects on cell size and function, making it likely to be selectively disadvantaged compared to its diploid progenitor(s) or other competing species, which are presumably well adapted to the respective environment.

If plants with rather small genomes are considered, polyploidy is relatively rare in macrophanerophytes (22.8%), which agrees with the early observations of Stebbins (1938, 1971). In addition, polyploidy in Czech plants also seems to be somewhat less frequent in therophytes (38.7%) and more frequent in chamaephytes (53.9%). Of the larger Czech plant families (with more than 20 species), polyploidy is unusually rare in *Apiaceae*, and also occurs with a low frequency in *Cyperaceae* and *Juncaceae* (though
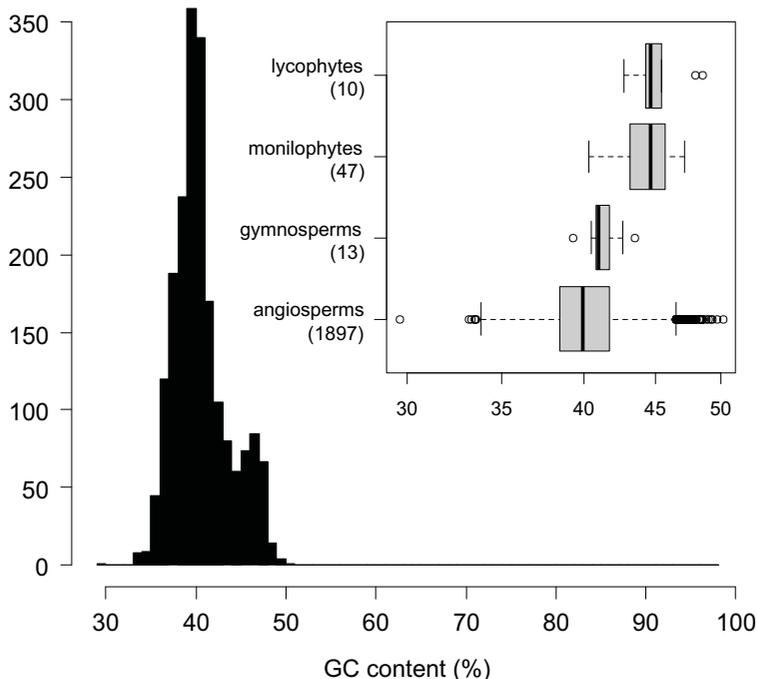
Fig. 5. – Distribution of GC contents in the genome dataset of Czech vascular plants (histogram) and separately for the four major phylogenetic groups (boxplots). The minor peak between 45–50% is mainly due to the high number of GC-rich grasses in the dataset and partly also to GC-rich monilophytes and lycophytes. Boxplots show: median (thick horizontal line), interquartile range (box), range of non-outlying values (whiskers) and outliers (circles); numbers of taxa (including all subspecies and/or ploidy levels of a species) are in brackets below the group names.

this may be partly an artefact of ploidy detection problems in these families; see the Methods), *Orobanchaceae* and *Orchidaceae* (probably as consequence of their relatively large genomes). In contrast, polyploidy is particularly frequent in *Violaceae*, *Onagraceae*, *Poaceae*, *Polygonaceae* and *Rosaceae*.

Genomic GC content

The GC contents in Czech vascular plants vary from 29.0% in *Amaranthus blitum* to 50.2% in *Gentianopsis ciliata* and shows an apparently bimodal distribution (Fig. 5). The highest GC contents are mainly found in *Poaceae* (Fig. 6), in which it ranges from 44.6% in *Eragrostis pilosa* to 49.7% in *Helictotrichon desertorum*. Apart from *Poaceae*, high GC contents are typical also of ferns and lycopods (Fig. 5) and species with a very high GC content (> 45%) also frequently occur in three other monocotyledonous families (*Araceae*, *Asparagaceae*, *Orchidaceae*). High GC contents are also occasionally recorded in *Aristolochia*, *Buglossoides*, *Butomus*, *Colchicum*, *Diphasiastrum*, *Gagea* (3 species), *Gentianopsis*, *Geranium*, *Loranthus*, *Oenothera* (5 species), *Pulsatilla* (2 species), *Ranunculus*, *Reseda*, *Rumex*, *Selaginella*, *Streptopus*, *Tofieldia*, *Triglochin*, and *Veratrum*. Apart from *Poaceae*, the highest GC contents are found in *Gentianopsis ciliata* (50.2%),
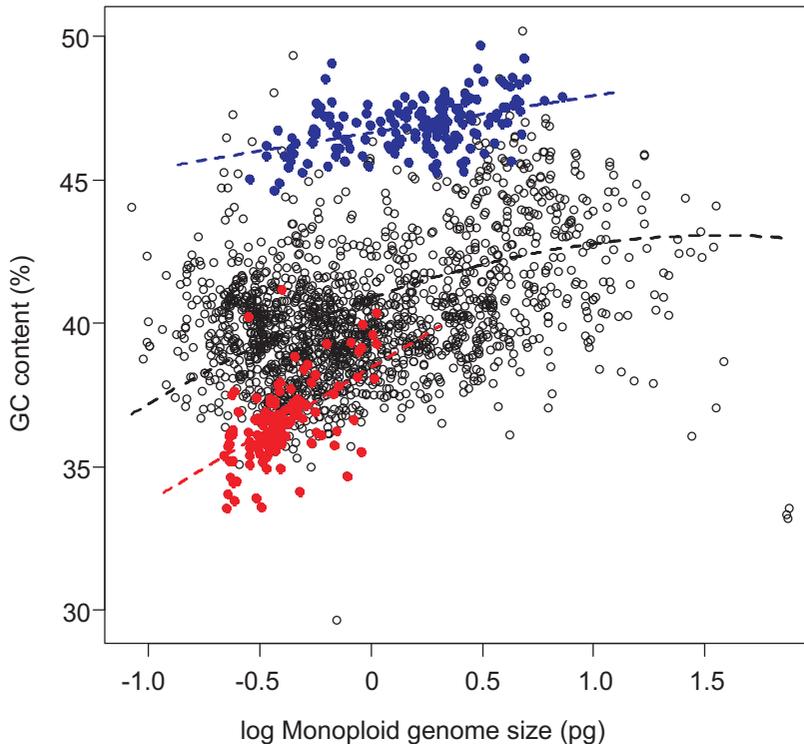
Fig. 6. – Relationship between the genomic GC content and monoploid genome size (1Cx). The general relationship is fitted using a quadratic polynomial function (P < 0.001, black interrupted line). The positive correlation between both genomic parameters in species with small monoploid genome sizes may be documented in both of the GC extreme groups, *Poaceae* (blue circles and linear regression line; P < 0.001) and *Cyperaceae+Juncaceae* (red circles and linear regression line; P < 0.001).

Aristolochia clematitis (49.3%), *Selaginella selaginelloides* (48.0%), *Streptopus amplexifolius* (48.0%), *Muscari botryoides* (47.2%), *Blechnum spicant* (47.1%), *Woodsia alpina* (47.1%), *Asplenium ruta-muraria* (47.0%) and *Gagea pusilla* (47.0%).

   Very low GC contents are typical of *Juncaceae* and *Cyperaceae*, both of which include species with holocentric chromosomes (Bureš et al. 2013, Šmarda et al. 2014, Fig. 6). The GC content in *Cyperaceae* varies from 33.5% in *Schoenoplectus pungens* to 41.3% in *Isolepis setacea*; in *Juncaceae* it ranges from 33.6% in *Juncus inflexus* to 38.1% in *J. articulatus* and the outlying value in this family is 40.2% in *J. trifidus*. GC contents lower than in *Cyperaceae* and *Juncaceae* occur only in the giant genome of *Viscum album* (GC = 33.2–33.5%) and *Amaranthus blitum* (GC = 29.0–30.0%). In both of these species the extremely low GC contents were confirmed by samples from several localities and in *Amaranthus blitum* using two alternative FCM standards (Electronic Appendix 2). The GC content in *Amaranthus blitum* is the lowest recorded so far in vascular plants and certainly points to a peculiar composition of the genome of this species, worthy of a future detailed genomic analysis. Apart from the above mentioned high GC contents (*Cyperaceae*, *Juncaceae*, *Viscum album*, *Amaranthus blitum*), very low GC contents (35–37%) are also quite frequent in *Amaranthus*, *Impatiens*, *Elaeagnus*,

Trifolium, *Malva*, *Tilia*, *Morus*, *Oleaceae*, *Primulaceae* and *Salix*, with the lowest GC contents documented for *Elaeagnus angustifolia* (35.0%), *Mercurialis annua* (35.3%), *Amaranthus powellii* (35.5%), *A. retroflexus* (35.6%) and *Trifolium striatum* (35.6%).

We have previously reported that genomic GC content is quadratically correlated with genome size, with low GC contents being found in plants with either small or very large genomes (Veselý et al. 2012, Šmarda et al. 2014). This relationship is also present in the data for Czech vascular plants (P < 0.001; Fig. 6).

New reports of cytotypes, chromosome counts and comments on related peculiarities of genome size

Genome size measurements revealed DNA ploidy in several species in which, to our knowledge, ploidy has never been previously reported, *Thesium dollineri* (2n~4x), *Juncus tenageia* (2n~2x), *Taraxacum ochrochlorum* (2n~3x), *Gentianella praecox* subsp. *bohemica* (2n~4x), *Stipa dasyphylla* (2n~4x) and *S. glabrata* (2n~4x), although for the last four taxa these counts are not surprising given that the same ploidy levels are also reported for their closest relatives.

Our survey revealed several cytotypes that have not yet been reported in the Czech Republic. While *Genista tinctoria* is known to be diploid (2n = 2x = 48) in Bohemia (Skalická in Slavík et al. 1995: 350–354), a tetraploid cytotype was found in southern Moravia at three localities (Brno-Nový Lískovec, Hnanice and Drásov). These findings accord with records of this cytotype in southern Slovakia (Murín & Neischlová 1973 as *G. hungarica*) and Lower Austria (Kiehn et al. 1991). The inferred ploidy level was confirmed also by direct chromosome counts (2n = 4x = ca 96) for plants from another site in southern Moravia (Omice). The existence of diploid and tetraploid cytotypes is also further confirmed for *Tephroseris integrifolia* and *Veratrum album* subsp. *album*. The detailed distribution of cytotypes and their taxonomic value are being evaluated in ongoing studies (Roleček et al. in prep., Grulich et al. in prep.). Hexaploids of *Salix bicolor* occur at its only known locality in the Krkonoše Mts, which contradicts the assumption by Chmelař & Koblížek (1990). The diploid ploidy level is also newly inferred for *Fumaria schleicherii* for which there are currently only two tetraploid chromosome counts; one for *F. schleicherii* subsp. *microcarpa* (Lidén 1986) and the other from the Caucasus (Gvinianidze & Avazneli 1982). This inference is based on the on the similarity of the genome size of *F. schleicheri* and that of the diploid *F. rostellata* and the very stable monoploid genome size in the whole genus.

The fact that a species has genome size that is double the size of that of its relatives, however, does not necessarily mean it is polyploid (cf. Piegu et al. 2006). *Lamium maculatum*, for example, has a genome size exactly double (2C = 2.46–2.49 pg) that of its close relative *L. purpureum* (2C = 1.25–1.33 pg). For both species, however, only diploid chromosome numbers (2n = 2x = 18) are reported a dozen times in the Chromosome Count Database (Rice et al. 2015). Our own chromosome count of one measured plant of *L. maculatum* confirmed that it is also diploid (2n = 18), indicating that genome size doubling may have occurred in this species in ways other than polyploidy. Similarly, chromosome counts for a FCM measured plant of *Salvia glutinosa* (2n = 2x = 16) verified its nearly doubled genome size (2C = 1.93–1.94 pg compared to 2C = 0.85–1.22 pg in *S. pratensis*, *S. nemorosa* and *S. verticillata*), but does not relate to the tetraploidy, which

was (therefore wrongly) assumed for a sample of this species with a similar genome size from the Balkan Peninsula (Siljak-Yakovlev et al. 2010: Appendix 1). A striking two-fold difference in genome size is also revealed by comparing the genome sizes of *Symphytum officiale* (2C = 3.89 pg) and *S. tuberosum* subsp. *angustifolium* (2C = 1.90 pg). However, both species are considered to be tetraploid based on their chromosome numbers, and this polyploidy-like genome size difference may be a result of a different karyotype evolution in these species, with each representing a different phylogenetic lineage in *Symphytum* (M. Hroneš, pers. comm.). There is another remarkable situation in *Plantago*, in which the genome sizes of *P. major* and *P. uliginosa* (2C = 1.30–1.33 pg) are about half that of the other diploid *Plantago* species, *P. lanceolata* and *P. maritima* (2C = 2.13–2.44 pg) or a quarter of that of their tetraploid relative *P. media* (2C = 4.95 pg). A chromosome count done on one measured *P. major* plant (2n = 2x = 12), however, confirmed that this is not a haploid but a regular diploid.

In addition to the chromosome counts for the above mentioned species, *Genista tinctoria* (2n = 4x = ca 96), *Lamium maculatum* (2n = 2x =18), *Salvia glutinosa* (2n = 2x =16), *Plantago major* (2n = 2x = 12), chromosomes were counted in order to check the ploidy level of FCM measured plants in 11 other species, including *Bromus erectus* (2n = 8x = ca 56), *Cardamine dentata* (2n = 9x-4 = 68), *Geranium macrorrhizum* (2n = 4x = 56), *Glyceria maxima* (2n = 6x = 60), *Potentilla arenaria* (2n = 4x = 28), *Sanguisorba officinalis* (2n = 8x = ca 56), *Poa angustifolia* (2n = 6x = 38–42), *Sedum reflexum* (2n = 14 = ca 106), *Symphyotrichum lanceolatum* (2n = 8x = 64), *Trifolium rubens* (2n = 2x = 16) and *Vicia pannonica* (2n = 2x = 12). Yet there are some other FCM measurements in our dataset that might be for rare or yet unknown cytotypes and for which chromosome counts are needed to confirm our inference about their ploidy level (e.g. *Aconitum anthora*, *Alchemilla* spp., *Anemonastrum narcissiflorum*, *Arenaria serpyllifolia*, *Campanula rapunculoides*, *Comarum palustre*, *Draba muralis*, *Hierochloë repens*, *Iris graminea*, *Lysimachia nemorum*, *L. thyrsiflora*, *Malva alcea*, *Muscari* spp., *Saxifraga paniculata*, *Pseudognaphalium luteoalbum*, *Scabiosa canescens*, *Sonchus arvensis*, *S. palustris*, *Stratiotes aloides*, *Thalictrum flavum* and *T. simplex* subsp. *galioides*; Electronic Appendix 2).

Perspectives

The present dataset provides the most comprehensive information on genomic parameters ever provided for a flora of a particular country. These data are now incorporated into the Database of the Czech Flora and Vegetation (Pladias; www.pladias.cz, Wild et al. 2019) and will be updated when new measurements become available. This dataset will enable robust tests of the effect of individual genomic parameters on species anatomy, physiology, morphology and ecology (e.g. cell size, growth rate, plant size, life strategy, drought tolerance and nutrient demands). Some preliminary results have shown that genome size may have significant effect on growth rate in herbaceous plants (Huang et al., in prep.), and many other relationships may be identified in the future. Based on this new dataset, these relationships can be studied not only at the species level but also at the level of plant communities, i.e. by considering the effect of species interactions on realized ecological niches.

See www.preslia.cz for Electronic Appendices 1–2.

Acknowledgements

Souhrn

Obsah DNA v somatických, nereplikovaných jádrech buněk (velikost genomu) a bázové složení DNA (GC obsah) jsou základní genomické parametry, které lze měřit pomocí průtokové cytometrie. Velikost genomu, a tedy i ploidní úroveň mohou ovlivňovat četné vlastnosti rostlin, a proto se uvádějí v moderních biologických flórách a databázích biologických vlastností druhů. Velikost genomu se dosud měří převážně pro biosystematické účely a navzdory velkému rozvoji průtokové cytometrie v České republice tato informace pro mnohé druhy české květeny chybí. Biologický význam GC obsahu je oproti významu velikosti genomu zatím méně jasný, navíc se měří obtížněji, a proto jej neznáme u velké většiny druhů. V této práci přinášíme měření obou genomických parametrů pro většinu cévnatých rostlin české květeny. Celkem jsme naměřili nově velikosti genomu a genomický GC obsah pro 1908 vzorků 1700 druhů. Velikost genomu je zde poprvé uvedena pro více než 1000 druhů a obsah GC bází poprvé pro více než 1500 druhů, což více než zdvojnásobuje množství dosud známých údajů o obsahu GC bází u cévnatých rostlin. Spolu s dříve uveřejněnými údaji získanými v naší laboratoři na stejných přístrojích a stejnou metodikou jsou tak tyto údaje dostupné pro 1910 druhů české květeny (~83 % současné flóry bez apomiktů, přechodně zavlečených a vyhynulých druhů). Tyto údaje jsou uspořádány do tabulky a doplněny údaji o počtu chromozomů, ploidní úrovni a vypočtené monoploidní velikosti genomu. V článku přinášíme popisnou analýzu tohoto datového souboru a diskutujeme o některých cytogeograficky pozoruhodných poznatcích. Tento přehled představuje největší a nejúplnější soubor genomických dat pro květenu jednoho státu. Zde prezentované výsledky tvoří základ karyologické části české databáze druhových vlastností Pladias (www.pladias.cz) a budou využity pro analýzy evoluce a biologického významu obou měřených genomických parametrů.

References

Baquar S. R. (1976) Polyploidy in the flora of Pakistan in relation to latitude, life form, and taxonomic groups. – Taxon 25: 621–627.

Barow M. & Meister A. (2002) Lack of correlation between AT frequency and genome size in higher plants and the effect of nonrandomness of base sequences on dye binding. – Cytometry 47: 1–7.

Basak S., Mukhopadhyay P., Gupta S. K. & Ghosh T. C. (2010) Genomic adaptation of prokaryotic organisms at high temperature. – Bioinformation 4: 352–356.

Beaulieu J. M., Leitch I. J., Patel S., Pendharkar A. & Knight C. A. (2008) Genome size is a strong predictor of cell size and stomatal density in angiosperms. – New Phytologist 179: 975–986.

Beaulieu J. M., Moles A. T., Leitch I. J., Bennett M. D., Dickie J. B. & Knight C. A. (2007) Correlated evolution of genome size and seed mass. – New Phytologist 173: 422–437.

Bennett M. D. (1987) Variation in genomic form in plants and its ecological implications. – New Phytologist 106: 177–200.

Bennett M. D. & Leitch I. J. (2012) Plant DNA C-values database, release 6.0. –
    URL: http://data.kew.org/cvalues/.

Bennetzen J. L., Ma J. & Devos K. M. (2005) Mechanisms of recent genome size variation in flowering plants.
    – Annals of Botany 95: 127–132.

Bernardi G. (2000) Isochores and the evolutionary genomics of vertebrates. – Gene 241: 3–17.

Brochmann C., Brysting A. K., Alsos I. G., Borgen L., Grundt H. H., Scheen A.-C. & Elven R. (2004)
    Polyploidy in arctic plants. – Biological Journal of the Linnean Society 82: 521–536.

Bureš P., Wang Y.-F., Horová L. & Suda J. (2004) Genome size variation in central European species of
    Cirsium (*Compositae*) and their natural hybrids. – Annals of Botany 94: 353–363.

Bureš P., Zedek F. & Marková M. (2013) Holocentric chromosomes. – In: Wendel F., Greilhuber J., Doležel J.
    & Leitch I. J. (eds), Plant genome diversity 2: 187–208, Springer, Vienna etc.

Čertner M., Fenclová E., Kúr P., Kolář F., Koutecký P., Krahulcová A. & Suda J. (2017) Evolutionary dynamics
    of mixed-ploidy populations in an annual herb: dispersal, local persistence and recurrent origins of
    polyploids. – Annals of Botany 120: 303–315.

Chargaff E. (1951) Structure and function of nucleic acids as cell constituents. – Federation Proceedings 10:
    654–659.

Chargaff E., Lipshitz R. & Green C. (1952) Composition of the desoxypentose nucleic acids of four genera of
    sea-urchin. – Journal of Biological Chemistry 195: 155–160.

Chargaff E., Vischer E., Doniger R., Green C. & Misani F. (1949) The composition of the desoxypentose
    nucleic acids of thymus and spleen. – Journal of Biological Chemistry 177: 405–416.

Chmelař J. & Koblížek J. (1990) *Salix* L. – vrba. – In: Hejný S., Slavík B., Hrouda L. & Skalický V. (eds),
    Květena České republiky [Flora of the Czech Republic] 2: 458–489, Academia, Praha.

Chrtek J., Zahradníček J., Krak K. & Fehrer J. (2009) Genome size in *Hieracium* subgenus *Hieracium*
    (*Asteraceae*) is strongly correlated with major phylogenetic groups. – Annals of Botany 104: 161–178.

Chytrý M. (2017a) Current vegetation of the Czech Republic. – In: Chytrý M., Danihelka J., Kaplan Z. & Pyšek
    P. (eds), Flora and vegetation of the Czech Republic, p. 229–337, Springer, Cham.

Chytrý M. (2017b) Physical geography of the Czech Republic. – In: Chytrý M., Danihelka J., Kaplan Z. &
    Pyšek P. (eds), Flora and vegetation of the Czech Republic, p. 1–23, Springer, Cham.

Danihelka J. (2013a) Botanické součty, rozdíly a podíly [Botanical counts]. – Živa 61: 69–72.

Danihelka J. (2013b) Jak se dělá seznam [Making a checklist]. – Živa 61: XXI–XXV.

Danihelka J., Chrtek J. Jr. & Kaplan Z. (2012) Checklist of vascular plants of the Czech Republic. – Preslia 84:
    647–811.

Doležel J. & Bartoš J. (2005) Plant DNA flow cytometry and estimation of nuclear genome size. – Annals of
    Botany 95: 99–110.

Doležel J., Bartoš J., Voglmayr H. & Greilhuber J. (2003) Nuclear DNA content and genome size of trout and
    human. – Cytometry 51: 127–128.

Doležel J., Binarová P. & Lucretti S. (1989) Analysis of nuclear DNA content in plant cells by flow cytometry.
    – Biologia Plantarum 31: 113–120.

Doležel J. & Göhde W. (1995) Sex determination in dioecious plants *Melandrium album* and *M. rubrum* using
    high resolution flow cytometry. – Cytometry 19: 103–106.

Doležel J. & Greilhuber J. (2010) Nuclear genome size: are we getting closer? – Cytometry Part A 77: 635–642.

Doležel J., Greilhuber J., Lucretti S., Meister A., Lysák M. A., Nardi L. & Obermayer R. (1998) Plant genome
    size estimation by flow cytometry: inter-laboratory comparison. – Annals of Botany 82: 17–26.

Doležel J., Greilhuber J. & Suda J. (2007) Estimation of nuclear DNA content in plants using flow cytometry. –
    Nature Protocols 2: 2233–2244.

Doležel J., Sgorbati S. & Lucretti S. (1992) Comparison of three DNA fluorochromes for flow cytometric esti-
    mation of nuclear DNA content in plants. – Physiologia Plantarum 85: 625–631.

Du H., Yu Y., Ma Y., Gao Q., Cao Y., Chen Z., Ma B., Qi M., Li Y., Zhao X., Wang J., Liu K., Qin P., Yang X.,
    Zhu L., Li S & Liang C. (2017) Sequencing and de novo assembly of a near complete indica rice genome. –
    Nature Communications 8: 15324.

Duchoslav M., Šafářová L. & Krahulec F. (2010) Complex distribution patterns, ecology and coexistence of
    ploidy levels of *Allium oleraceum* (*Alliaceae*) in the Czech Republic. – Annals of Botany 105: 719–735.

Ekrt L. & Koutecký P. (2016) Between sexual and apomictic: unexpectedly variable sporogenesis and produc-
    tion of viable polyhaploids in the pentaploid fern of the *Dryopteris affinis* agg. (*Dryopteridaceae*). –
    Annals of Botany 117: 97–106.

Fitter A. H. & Peat H. J. (1994) The Ecological Flora Database. – Journal of Ecology 82: 415–425,
    URL: http://www.ecoflora.co.uk.

Galbraith D. W., Harkins K. R., Maddox J. M., Ayres N. M., Sharma D. P. & Firoozabady E. (1983) Rapid flow cytometric analysis of the cell cycle in intact plant tissues. – Science 220: 1049–1051.

Garcia S., Leitch I. J., Anadon-Rosell A., Canela M. Á., Gálvez F., Garnatje T., Gras A., Hidalgo O., Johnston E., Mas de Xaxars G., Pellicer J., Siljak-Yakovlev S., Vallès J., Vitales D. & Bennett M. D. (2014) Recent updates and developments to plant genome size databases. – Nucleic Acids Research 42: D1159–D1166.

Grant V. (1981) Plant speciation. – Columbia University Press, New York.

Greilhuber J. (1998) Intraspecific variation in genome size: a critical reassessment. – Annals of Botany 82, Suppl. A: 27–35.

Greilhuber J. (2005) Intraspecific variation in genome size in angiosperms: identifying its existence. – Annals of Botany 95: 91–98.

Greilhuber J., Doležel J., Lysák M. A. & Bennett M. D. (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.

Greilhuber J. & Leitch I. J. (2013) Genome size and the phenotype. – In: Greilhuber J., Doležel J. & Wendel J. (eds), Plant genome diversity 2: 323–344, Springer, Vienna etc.

Greilhuber J., Temsch E. M. & Loureiro J. C. M. (2007) Nuclear DNA content measurement. – In: Doležel J., Greilhuber J. & Suda J. (eds), Flow cytometry with plant cells, p. 67–101, Willey-VCH, Weinheim.

Grif V. G. (2000) Some aspects of plant karyology and karyosystematics. – International Review of Cytology 196: 131–175.

Grime J. P. & Mowforth M. A. (1982) Variation in genome size: an ecological interpretation. – Nature 299: 151–153.

Guignard M. S., Nichols R. A., Knell R. J., Macdonald A., Romila C.-A., Trimmer M., Leitch I. J. & Leitch A. R. (2016) Genome size and ploidy influence angiosperm species' biomass under nitrogen and phosphorus limitation. – New Phytologist 210: 1195–1206.

Gvinianidze Z. I. & Avazneli A. A. (1982) Khromosomnye chisla nekotorykh predstavitelej vysokogornykh floristicheskikh kompleksov Kavkaza [Chromosome numbers in some representatives of high-mountain flora of the Caucasus]. – Soobshcheniya Akademii Nauk Gruzinskoi SSR, Institut Botaniki, Trudy, Seriya Geobotanika 106: 577–580.

Hanzl M., Kolář F., Nováková D. & Suda J. (2014) Nonadaptive processes governing early stages of polyploid evolution: insights from a primary contact zone of relict serpentine *Knautia arvensis* (*Caprifoliaceae*). – American Journal of Botany 101: 935–945.

Hejný S., Slavík B., Chrtek J., Tomšovic P. & Kovanda M. (eds) (1988) Květena České socialistické republiky [Flora of the Czech Socialist Republic]. Vol. 1. – Academia, Praha.

Hejný S., Slavík B., Hrouda L. & Skalický V. (eds) (1990) Květena České republiky [Flora of the Czech Republic]. Vol. 2. – Academia, Praha.

Hejný S., Slavík B., Kirschner J. & Křísa B. (eds) (1992) Květena České republiky [Flora of the Czech Republic]. Vol. 3. – Academia, Praha.

Hidalgo O., Pellicer J., Christenhusz M., Schneider H., Leitch A. R. & Leitch I. J. (2017) Is there an upper limit for genome size? – Trends in Plant Science 22: 567–573.

Hodgson J. G., Sharafi M., Jalili A., Díaz S., Montserrat-Martí G., Palmer C., Cerabolini B., Pierce S., Hamzehee B., Asri Y., Jamzad Z., Wilson P., Raven J. A., Band S. R., Basconcelo S., Bogard A., Carter G., Charles M., Castro-Díez P., Cornelissen J. H. C., Funes G., Jones G., Khoshnevis M., Pérez-Harguindeguy N., Pérez-Rontomé M. C., Shirvany F. A., Vendramini F., Yazdani S., Abbas-Azimi R., Boustani S., Dehghan M., Guerrero-Campo J., Hynd A., Kowsary E., Kazemi-Saeed F., Siavash B., Villar-Salvador P., Craigie R., Naqinezhad A., Romo-Díez A., de Torres Espuny L. & Simmons E. (2010) Stomatal vs. genome size in angiosperms: the somatic tail wagging the genomic dog? – Annals of Botany 105: 573–584.

Hoffman R. A. & Wood J. C. S. (2007) Characterization of flow cytometer instrument sensitivity. – Current Protocols in Cytometry 40: 1.20.1–1.20.18.

International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. – Nature 431: 931–945.

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. – Nature 436: 793–800.

Kabátová K., Vít P. & Suda J. (2014) Species boundaries and hybridization in central-European *Nymphaea* species inferred from genome size and morphometric data. – Preslia 86: 131–154.

Kaplan Z. (2017) Flora and phytogeography of the Czech Republic. – In: Chytrý M., Danihelka J., Kaplan Z. & Pyšek P. (eds), Flora and vegetation of the Czech Republic, p. 89–163, Springer, Cham.

Kaplan Z., Danihelka J., Chrtek Jr. J., Kirschner J., Kubát K., Štěpánek J. & Štech M. (eds) (2019) Klíč ke květeně České republiky [Key to the flora of the Czech Republic]. – Academia, Praha (in press).

Kattge J., Díaz S., Lavorel S., Prentice I. C., Leadley P., Bönisch G., Garnier E., Westoby M., Reich P. B., Wright I. J., Cornelissen J. H. C., Violle C., Harrison S. P., Van Bodegom P. M., Reichstein M., Enquist B. J., Soudzilovskaia N. A., Ackerly D. D., Anand M., Atkin O., Bahn M., Baker T. R., Baldocchi D., Bekker R., Blanco C. C., Blonder B., Bond W. J., Bradstock R., Bunker D. E., Casanoves F., Cavender-Bares J., Chambers J. Q., Chapin III. F. S., Chave J., Coomes D., Cornwell W. K., Craine J. M., Dobrin B. H., Duarte L., Durka W., Elser J., Esser G., Estiarte M., Fagan W. F., Fang J., Fernández-Méndez F., Fidelis A., Finegan B., Flores O., Ford H., Frank D., Freschet G. T., Fyllas N. M., Gallagher R. V., Green W. A., Gutierrez A. G., Hickler T., Higgins S. I., Hodgson J. G., Jalili A., Jansen S., Joly C. A., Kerkhoff A. J., Kirkup D., Kitajima K., Kleyer M., Klotz S., Knops J. M. H., Kramer K., Kühn I., Kurokawa H., Laughlin D., Lee T. D., Leishman M., Lens F., Lenz T., Lewis S. L., Lloyd J., Llusià J., Louault F., Ma S., Mahecha M. D., Manning P., Massad T., Medlyn B. E., Messier J., Moles A. T., Müller S. C., Nadrowski K., Naeem S., Niinemets Ü., Nöllert S., Nüske A., Ogaya R., Oleksyn J., Onipchenko V. G., Onoda Y., Ordoñez J., Overbeck G., Ozinga W. A., Patiño S., Paula S., Pausas J. G., Peñuelas J., Phillips O. L., Pillar V., Poorter H., Poorter L., Poschlod P., Prinzing A., Proulx R., Rammig A., Reinsch S., Reu B., Sack L., Salgado-Negret B., Sardans J., Shiodera S., Shipley B., Siefert A., Sosinski E., Soussana J.-F., Swaine E., Swenson N., Thompson K., Thornton P., Waldram M., Weiher E., White M., White S., Wright S. J., Yguel B., Zaehle S., Zanne A. E. & Wirth C. (2011) TRY – a global database of plant traits. – Global Change Biology 17: 2905–2935.

Kiehn M., Vitek E., Hellmayr E., Walter J., Tschenett J., Justin C. & Mann M. (1991) Beiträge zur Flora von Österreich: Chromosomenzählungen. – Verhandlungen der Zoologisch-botanischen Gesellschaft in Österreich 128: 19–39.

Klotz S., Kühn I. & Durka W. (ed.) (2002) BIOLFLOR: Eine Datenbank zu biologisch-ökologischen Merkmalen der Gefäßpflanzen in Deutschland. – Schriftenreihe für Vegetationskunde 38: 1–334, URL: http://www.biolflor.de.

Knight C. A. & Beaulieu J. M. (2008) Genome size scaling through phenotype space. – Annals of Botany 101: 759–766.

Knight C. A., Clancy R. B., Götzenberger L., Dann L. & Beaulieu J. M. (2010) On the relationship between pollen size and genome size. – Journal of Botany 2010: 612017.

Kobrlová L., Hroneš M., Koutecký P., Štech M. & Trávníček B. (2016) *Symphytum tuberosum* complex in central Europe: cytogeography, morphology, ecology and taxonomy. – Preslia 88: 77–112.

Kočková L. (2012) Velikost genomu a poměr bazí v genomu v čeledi *Ranunculaceae* [Genome size and base ratio in *Ranunculaceae*]. – Master thesis, Dept. of Botany and Zoology, Masaryk University, Brno, URL: https://is.muni.cz/th/209653/prif_m/Diplomova_praceFINAL.pdf.

Kolář F., Lučanová M., Koutecký P., Dortová M., Knotek A. & Suda J. (2014) Spatio-ecological segregation of diploid and tetraploid cytotypes of *Galium valdepilosum* in central Europe. – Preslia 86: 155–178.

Kolář F., Štech M., Trávníček P., Rauchová J., Urfus T., Vít P., Kubešová M. & Suda J. (2009) Towards resolving the *Knautia arvensis* agg. (*Dipsacaceae*) puzzle: primary and secondary contact zones and ploidy segregation at landscape and microgeographic scales. – Annals of Botany 103: 963–974.

Kubešová M., Moravcová L., Suda J., Jarošík V. & Pyšek P. (2010) Naturalized plants have smaller genomes than their non-invading relatives: a flow cytometric analysis of the Czech alien flora. – Preslia 82: 81–96.

Leitch I. J. & Leitch A. R. (2013) Genome size diversity and evolution in land plants. – In: Greilhuber J., Doležel J. & Wendel J. (eds), Plant genome diversity 2: 307–322, Springer, Vienna etc.

Leitch I. J., Soltis D. E., Soltis P. S. & Bennett M. D. (2005) Evolution of DNA amounts across land plants (*Embryophyta*). – Annals of Botany 95: 207–217.

Leitch A. R. & Leitch I. J. (2008) Genomic plasticity and the diversity of polyploid plants. – Science 320: 481–483.

Lepší M., Lepší P., Koutecký P., Bílá J. & Vít P. (2015) Taxonomic revision of *Sorbus* subgenus *Aria* occurring in the Czech Republic. – Preslia 87: 109–162.

Levin D. A. (2002) The role of chromosomal change in plant evolution. – Oxford University Press, New York.

Lidén M. (1986) Synopsis of *Fumarioideae* (*Papaveraceae*) with a monograph of the tribe *Fumarieae*. – Opera Botanica 88: 1–133.

Lipnerová I., Bureš P., Horová L. & Šmarda P. (2013) Evolution of genome size in *Carex* (*Cyperaceae*) in relation to chromosome number and genomic base composition. – Annals of Botany 111: 79–94.

Loureiro J., Eleazar Rodriguez E., Doležel J. & Santos C. (2007) Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. – Annals of Botany 100: 875–888.

Lysák M. A. & Doležel J. (1998) Estimation of nuclear DNA content in *Sesleria* (*Poaceae*). – Caryologia 51: 123–132.

Mandák B., Krak K., Vít P., Pavlíková Z., Lomonosova M. N., Habibi F., Wang L., Jellen E. N. & Douda J. (2017) How genome size variation is linked with evolution within *Chenopodium* sensu lato. – Perspectives in Plant Ecology, Evolution and Systematics 23: 18–32.

Mandáková T. & Lysák M. A. (2016) Chromosome preparation for cytogenetic analyses in *Arabidopsis*. – Current Protocols in Plant Biology 1: 43–51.

Mandel M. & Marmur J. (1968) Use of ultraviolet absorbance-temperature profile for determining the guanine plus cytosine content of DNA. – Methods in Enzymology, part B, 12: 195–206.

Mann S. & Phoebe-Chen Y. P. (2010) Bacterial genomic G+C composition-eliciting environmental adaptation. – Genomics 95: 7–15.

Marcussen T., Heier L., Brysting A. K., Oxelman B. & Jakobsen K. S. (2015): From gene trees to a dated allo-polyploid network: insights from the angiosperm genus *Viola* (Violaceae). – Systematic Biology 64: 84–101.

Marmur J. & Doty P. (1962) Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. – Journal of Molecular Biology 5: 109–118.

Matzk F., Meister A., Brutovská R. & Schubert I. (2001) Reconstruction of reproductive diversity in *Hypericum perforatum* L. opens novel strategies to manage apomixis. – The Plant Journal 26: 275–282.

Meister A. & Barow M. (2007) DNA base composition of plant genomes. – In: Doležel J., Greilhuber J. & Suda J. (eds), Flow cytometry with plant cells, p. 177–215, Willey-VCH, Weinheim.

Morgan-Richards M., Trewick S. A., Chapman H. M. & Krahulcová A. (2004) Interspecific hybridization among *Hieracium* species in New Zealand: evidence from flow cytometry. – Heredity 93: 34–42.

Müntzing A. (1936) The evolutionary significance of autopolyploidy. – Hereditas 21: 363–378.

Murín A. & Neischlová E. (1973) Chromosomenzahlen einiger Taxa aus dem Verwandtschaftskreis von *Genista tinctoria* L. – Biológia (Bratislava), ser. A, 28: 493–497.

Musto H., Naya H., Zavala A., Romero H., Alvarez-Valin F. & Bernardi G. (2006) Genomic GC level, optimal growth temperature, and genome size in prokaryotes. – Biochemical and Biophysical Research Communications 347: 1–3.

Nishio Y., Nakamura Y., Kawarabayasi Y., Usuda Y., Kimura E., Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K. & Gojobori T. (2003) Comparative complete genome sequence analysis of the amino acid replacements responsible for the thermostability of *Corynebacterium efficiens*. – Genome Research 13: 1572–1579.

Nordenskiöld H. (1951) Cytotaxonomical studies in the genus *Luzula*. – Hereditas 37: 325–355.

Otisková V., Koutecký T., Kolář F. & Koutecký P. (2014) Occurrence and habitat preferences of diploid and tetraploid cytotypes of *Centaurea stoebe* in the Czech Republic. – Preslia 86: 67–80.

Otto F. (1990) DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. – Methods in Cell Biology 33: 105–110.

Pfosser M., Amon A., Lelley T. & Heberle-Bors E. (1995) Evaluation of sensitivity of flow cytometry in detecting aneuploidy in wheat using disomic and ditelosomic wheat-rye addition lines. – Cytometry 21: 387–393.

Piegu B., Guyot R., Picault N., Roulin A., Sanyal A., Kim H., Collura K., Brar D. S., Jackson S., Wing R. A. & Panaud O. (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.

Raunkiær C. (1934) The life forms of plants and statistical plant geography. – Clarendon Press, Oxford.

Raven P. H. (1975) The bases of angiosperm phylogeny: cytology. – Annals of the Missouri Botanical Garden 62: 724–764.

Rice A., Glick L., Abadi S., Einhorn M., Kopelman N. M., Salman-Minkov A., Mayzel J., Chay O. & Mayrose I. (2015) The chromosome counts database (CCDB): a community resource of plant chromosome numbers. – New Phytologist 206: 19–26.

Rice A., Šmarda P., Novosolov M., Drori M., Glick L., Sabath N., Meiri S., Belmaker J. & Mayrose I. (2019) The global biogeography of polyploid plants. – Nature Ecology and Evolution 3: 265–273.

Rieber N., Zapatka M., Lasitschka B., Jones D., Northcott P., Hutter B., Jäger N., Kool M., Taylor M., Lichter P., Pfister S., Wolf S., Brors B. & Eils R. (2013) Coverage bias and sensitivity of variant calling for four whole-genome sequencing technologies. – PLoS ONE 8: e66621.

Rocha E. P. C. & Danchin A. (2002) Competition for scarce resources might bias bacterial genome composition. – Trends in Genetics 18: 291–294.

Rotreklová O., Bureš P., Řepka R., Grulich V., Šmarda P., Hralová I., Zedek F. & Koutecký T. (2011) Chromosome numbers of *Carex*. – Preslia 83: 25–58.

Šarhanová P., Sharbel T. F., Sochor M., Vašut R. J., Dančák M. & Trávníček B. (2017) Hybridization drives evolution of apomicts in *Rubus* subgenus *Rubus*: evidence from microsatellite markers. – Annals of Botany 120: 317–328.

Šarhanová P., Vašut R. J., Dančák M., Bureš P. & Trávníček B. (2012) New insights into the variability of reproduction modes in European populations of *Rubus* subgen. *Rubus*: how sexual are polyploid brambles? – Sexual Plant Reproduction 25: 319–335.

Schönswetter P., Suda J., Popp M., Weiss-Schneeweiss H. & Brochmann C. (2007) Circumpolar phylogeography of *Juncus biglumis* (*Juncaceae*) inferred from AFLP fingerprints, cpDNA sequences, nuclear DNA content and chromosome numbers. – Molecular Phylogenetics and Evolution 42: 92–103.

Siljak-Yakovlev S., Pustahija F., Šolić E. M., Bogunić F., Muratović E., Bašić N., Catrice O. & Brown S. C. (2010) Towards a genome size and chromosome number database of the Balkan flora: C-values in 343 taxa with novel values for 242. – Advanced Science Letters 3: 190–213.

Slavík B., Chrtek J. jun. & Štěpánková J. (eds) (2000) Květena České republiky [Flora of the Czech Republic]. Vol. 6. – Academia, Praha.

Slavík B., Chrtek J. jun. & Tomšovic P. (eds) (1997) Květena České republiky [Flora of the Czech Republic].Vol. 5. – Academia, Praha.

Slavík B., Smejkal M., Dvořáková M. & Grulich V. (eds) (1995) Květena České republiky [Flora of the Czech Republic]. Vol. 4. – Academia, Praha.

Slavík B., Štěpánková J. & Štěpánek J. (eds) (2004) Květena České republiky [Flora of the Czech Republic].Vol. 7. – Academia, Praha.

Šmarda P. & Bureš P. (2006) Intraspecific DNA content variability in *Festuca pallens* on different geographical scales and ploidy levels. – Annals of Botany 98: 665–678.

Šmarda P. & Bureš P. (2010) Understanding intraspecific variation in genome size in plants. – Preslia 82: 41–61.

Šmarda P. & Bureš P. (2012) The variation of base composition in plant genomes. – In: Wendel F., Greilhuber J., Doležel J. & Leitch I. J. (eds), Plant genome diversity 1: 209–235, Springer, Vienna.

Šmarda P., Bureš P., Horová L., Foggi B. & Rossi G. (2008) Genome size and GC content evolution of *Festuca*: ancestral expansion and subsequent reduction. – Annals of Botany 101: 421–433.

Šmarda P., Bureš P., Horová L., Leitch I. J., Mucina L., Pacini E., Tichý L., Grulich V. & Rotreklová O. (2014) Ecological and evolutionary significance of genomic GC content diversity in monocots. – Proceedings of the National Academy of Sciences of the United States of America 111: E4096–E4102.

Šmarda P., Bureš P., Šmerda J. & Horová L. (2012) Measurements of genomic GC content in plant genomes with flow cytometry: a test for reliability. – New Phytologist 193: 513–521.

Šmarda P., Hejcman M., Březinová A., Horová L., Steigerová H., Zedek F., Bureš P., Hejcmanová P. & Schellberg J. (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., Horová L., Knápek O., Dieck H., Dieck M., Ražná K., Hrubík P., Orlóci L., Papp L., Veselá K., Veselý P. & Bureš P. (2018) Multiple haploids, triploids, and tetraploids found in modern-day "living fossil" *Ginkgo biloba*. – Horticulture Research 5: 55.

Šmarda P., Müller J., Vrána J. & Kočí K. (2005) Ploidy level variability of some Central European fescues (*Festuca* subg. *Festuca*, *Poaceae*). – Biologia (Bratislava) 60: 25–36.

Šmarda P., Veselý P., Šmerda J., Bureš P., Knápek O. & Chytrá M. (2016) Polyploidy in a 'living fossil' *Ginkgo biloba*. – New Phytologist 212: 11–14.

Soltis D. E., Visger C. J. & Soltis P. S. (2014) The polyploidy revolution then…and now: Stebbins revisited. – American Journal of Botany 101: 1057–1078.

Soltis P. & Soltis D. E. (eds) (2012) Polyploidy and genome evolution. – Springer, Vienna.

Stebbins G. L. (1938) Cytological characteristics associated with the different growth habits in the dicotyledons. – American Journal of Botany 25: 189–198.

Stebbins G. L. (1950) Variation and evolution in plants. – Columbia University Press, New York.

Stebbins G. L. (1971) Chromosomal evolution in higher plants. – Edward Arnold, London.

Stebbins G. L. (1985) Polyploidy, hybridization, and the invasion of new habitats. – Annals of the Missouri Botanical Garden 72: 824–832.

Štěpánková J., Chrtek J. jun. & Kaplan Z. (eds) (2010) Květena České republiky [Flora of the Czech Republic]. Vol. 8. – Academia, Praha.

Sterner R. W. & Elser J. J. (2002) Ecological stoichiometry. The biology of elements from molecules to biosphere. – Princeton University Press, Princeton and Oxford, UK.

Suda J., Krahulcová A., Trávníček P., Rosenbaumová R., Peckert T. & Krahulec F. (2007) Genome size variation and species relationships in *Hieracium* sub-genus *Pilosella* (*Asteraceae*) as inferred by flow cytometry. – Annals of Botany 100: 1323–1335.

Suda J., Kyncl T. & Jarolímová V. (2005) Genome size variation in Macaronesian angiosperms: forty percent of the Canarian endemic flora completed. – Plant Systematics and Evolution 252: 215–238.

Suda J. & Lysák M. A. (2001) A taxonomic study of the *Vaccinium* sect. *Oxycoccus* (Hill) W. D. J. Koch (*Ericaceae*) in the Czech Republic and adjacent territories. – Folia Geobotanica 36: 303–320.

Swift H. (1950) The constancy of desoxyribose nucleic acid in plant nuclei. – Proceedings of the National Academy of Sciences of the United States of America 36: 643–654.

Trávníček P., Kirschner J., Chudáčková H., Rooks F. & Štěpánek J. (2013) Substantial genome size variation in *Taraxacum stenocephalum* (*Asteraceae*, *Lactuceae*). – Folia Geobotanica 48: 271–284.

Trávníček P., Ponert J., Urfus T., Jersáková J., Vrána J., Hřibová E., Doležel J. & Suda J. (2015) Challenges of flow-cytometric estimation of nuclear genome size in orchids, a plant group with both whole-genome and progressively partial endoreplication. – Cytometry 87: 958–966.

Veeckman E., Ruttink T. & Vandepoele K. (2016) Are we there yet? Reliably estimating the completeness of plant genome sequences. – Plant Cell 28: 1759–1768.

Veleba A., Bureš P., Adamec L., Šmarda P., Lipnerová I. & Horová L. (2014) Genome size and genomic GC content evolution in the miniature genome-sized family *Lentibulariaceae*. – New Phytologist 203: 22–28.

Veleba A., Šmarda P., Zedek F., Horová L., Šmerda J. & Bureš P. (2017) Evolution of genome size and genomic GC content in carnivorous holokinetics (*Droseraceae*). – Annals of Botany 119: 409–416.

Veselý P., Bureš P. & Šmarda P. (2013) Nutrient reserves may allow for genome size increase: evidence from comparison of geophytes and their sister non-geophytic relatives. – Annals of Botany 112: 1193–1200.

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.

Vij S., Gupta V., Kumar D., Vydianathan R., Raghuvanshi S., Khurana P., Khurana J. P. & Tyagi A. K. (2006) Decoding the rice genome. – Bioessays 28: 421–432.

Vinogradov A. E. (2003) DNA helix: the importance of being GC-rich. – Nucleic Acids Research 31: 1838–1844.

Vít P., Wolfová K., Urfus T., Tájek P. & Suda J. (2014) Interspecific hybridization between rare and common plant congeners inferred from genome size data: assessing the threat to the Czech serpentine endemic *Cerastium alsinifolium*. – Preslia 86: 95–117.

Watson J. D. & Crick F. H. C. (1953) Genetical implications of the structure of deoxyribonucleic acid. – Nature 171: 964–967.

Wild J., Kaplan Z., Danihelka J., Petřík P., Chytrý M., Novotný P., Rohn M., Šulc V., Brůna J., Chobot K., Ekrt L., Holubová D., Knollová I., Kocián P., Štech M., Štěpánek J. & Zouhar V. (2019) Plant distribution data for the Czech Republic integrated in the Pladias database. – Preslia 91: 1–24.

Wendel J. F., Jackson S. A., Meyers B. C. & Wing R. A. (2016) Evolution of plant genome architecture. – Genome Biology 17: 37.

Yakovchuk P., Protozanova E. & Frank-Kamenetskii M. D. (2006) Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. – Nucleic Acids Research 34: 564–574.

Zedek F., Šmerda J., Šmarda P. & Bureš P. (2010) Correlated evolution of LTR retrotransposons and genome size in the genus *Eleocharis*. – BMC Plant Biology 10: 265.