Univerzita Karlova
Přírodovědecká fakulta

Studium velikosti genomu rostlin – od vnitrodruhové variability po ekologické důsledky

Genome size studies in plants – from intraspecific variation to ecological consequences

Magdalena Lučanová

Disertační práce

Školitelé: Jan Suda a Pavel Trávníček

Praha 2018

DECLARATION

I hereby declare that I have compiled this thesis by myself, using the cited references, and that I have never submitted or presented any part of this thesis for the award of any other degree or diploma.

Prohlašuji, že jsem tuto práci zpracovala nezávisle, za využití zmíněných zdrojů a nepodala jsem ani neprezentovala žádnou část této práce pro získání jiného titulu či diplomu.

České Budějovice, 9. 5. 2018

Magdalena Lučanová

ACKNOWLEDGEMENTS

My greatest thanks go to my late supervisor Jan Suda, who was an incredibly inspirational and enthusiastic person. I immensely appreciate that in his busy life of a great scientist he always had time for discussions and to help me in different ways. He was the guru of plant flow cytometry and introduced me to the topic of genome size. I will forever be grateful that he has given me opportunities to be involved in interesting and successful projects. I wholeheartedly thank Pavel Trávníček that he was willing to become my supervisor during the final stage of the writing of this thesis. His comments on its last draft were very valuable. I also gratefully acknowledge the helpful advice of Filip Kolář and his regular invitations to participate on his projects. To my laboratory co-workers Tomáš Urfus and Petr Vít I am grateful that they kindly accepted me as a full-fledged member of their inner circle and that they let me into some of the secrets of flow cytometry. I am also grateful to all co-authors of the articles included in this dissertation that I could have been part of their teams applying different methods and approaches, and to all other prominent team leaders – especially Karol Marhold and Petr Pyšek – that they asked me to join their teams. Jana Krejčíková deserves many thanks for helping me with flow cytometry, and my thanks also go to all my other friends that supported me morally and technically. For ample advice, mainly concerning statistics, I am indebted to Petr Koutecký. Finally, I am grateful to my parents, who supported me in all aspects of my life and never stopped believing that one day I would finish my PhD studies. My thanks also belong to Vašek Kouba for his patience and to my son Vašík, who is the best reason not to work all the time.

CONTENT

AUTHOR CONTRIBUTION STATEMENT

I hereby confirm that I has contributed to the manuscript and published papers concerned in the below-mentioned scope:

I. Lučanová M., Loureiro J., Suda J. (manuscript): Genome size variation correlates with fitness indicators in *Taraxacum stenocephalum* (Asteraceae)

Study design, experimental work, crossings, analysing and processing of flow cytometric data, growing plants, manuscript preparation. Total contribution: 80 %.

II. 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

Field work, collecting and processing of flow cytometric data, participation on manuscript preparation. Total contribution: 40 %.

III. Kolář F., **Lučanová M.**, Záveská E., Fuxová G., Mandáková T., Španiel S., Senko D., Svitok M., Kolník M., Gudžinskas Z., Marhold K. (2015): Ecological segregation does not drive the intricate parapatric distribution of diploid and tetraploid cytotypes of the *Arabidopsis arenosa* group (Brassicaceae). – Biological Journal of the Linnean Society, doi 10.1111/bij.12479

Field work, collecting and processing of flow cytometric data, participation on manuscript preparation. Total contribution: 40 %.

IV. 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

Collecting and processing of flow cytometric data, processing results, participation on manuscript preparation. Total contribution: 70 %.

V. Dušková E., Kolář F., Sklenář P., Rauchová J., **Kubešová M.**, Fér T., Suda J. & Marhold K. (2010): Genome size correlates with growth form, habitat and phylogeny in the Andean genus *Lasiocephalus* (Asteraceae). – Preslia 82: 127-148

Collecting and processing of flow cytometric data, participation on manuscript preparation. Total contribution: 20 %.

VI. Kolář F., **Lučanová M.**, Těšitel J., Loureiro J. & Suda J. (2012): Glycerol-treated nuclear suspensions – an efficient preservation method for flow cytometric analysis of plant samples. – Chromosome Research 20(2): 303-315 https://doi.org/10.1007/s10577-012-9277-0

Study design, field work, colecting and processsing of flow cytometric data, participation on manuscript preparation. Total contribution: 40 %.

ABSTRACT

Nuclear DNA content (genome size) is one of the basic characteristics of living organisms. In the Angiosperms, the range of genome size is 2,300-fold, which raises questions about the causes and consequences of this tremendous variation. This thesis deals with genome size in plants from the level of intraspecific homoploid variation, through intraspecific ploidy variation, to interspecies comparisons. On various study systems we investigated the dynamics and ranges of genome size variation, tried to reveal possible associations between genome size and selected biological traits, and assessed the extent to which differences in genome size are manifested at the ecological and evolutionary level.

As a means of estimating genome size we applied flow cytometry (FCM). In *Taraxacum stenocephalum* we conducted a detailed study of its enormous genome size variation. We carried out crossings of parents with various genome sizes and compared these parental genome sizes with those of F1 offspring. We also attempted to reveal the association of genome size with various growth traits. In *Galium valdepilosum* and *Arabidopsis arenosa* we carried out an extensive flow-cytometric ploidy level screening and compared the distribution and ecological preferences of detected cytotypes. We studied the Andean genus *Lasiocephalus* growing in different habitats spanning a wide range of elevations. Using FCM we determined both relative and absolute genome size across the genus, and, with the use of ITS sequencing, attempted to identify the sources of genome size variation within it. On the set of allien species we studied the relationship of genome size and invasiveness. Last but not least, we attempted to overcome the need for fresh samples in flow-cytometric genome size measurements. We modified the protocol of nuclei preservation in glycerol and verified the new protocol in a time-scale laboratory experiment and a field experiment.

We have confirmed the existence of substantial genome size variation within the species *Taraxacum stenocephalum* and found a strong correlation between parental and F1 genome size. We have also found correlations between genome size and certain biological traits that might influence the establishment of populations. The two ploidy-variable species under study differ in the ecological preferences of their cytotypes. In *Lasiocephalus* the main factor determining genome size variation is phylogeny. We found the association of genome size and invasiveness. We designed a new protocol for the preservation of nuclei which enables to estimate genome size in samples from remote areas.

This thesis unambiguously shows that genome size is associated with various aspects of the lives of plants and that it therefore represents a useful marker in plant studies. Genome size can not only indicate important species-forming processes, such as polyploidization, selection and hybridization, but it can also help us better understand the evolution of taxa. The effect of genome size differs from species to species and at different levels (e.g. anatomical, morphological, phenological or ecological), and to answer particular questions it is necessary to select a suitable model system and to meaningfully set the scope of the study.

ABSTRAKT

Velikost genomu je jednou ze základních charakteristik živých organizmů. U krytosemenných rostlin dosahuje rozsahu pěti řádů, což vždy fascinovalo vědce a vyvolávalo mnoho otázek. Tato práce se zabývá různými aspekty velikosti genomu u rostlin. Počínaje detailním studiem na homoploidní vnitrodruhové úrovni, přes studium ploidně variabilních druhů až po srovnávání velikosti genomu mezi druhy, jsme se snažili zjistit, jaký vztah má velikost genomu k biologickým vlastnostem rostlin, stanovit dynamiku a rozsah variability velikosti genomu a detekovat dosah velikosti genomu až do úrovně ekologické a evoluční.

Stěžejní metodou stanovení velikosti genomu v celé práci je průtoková cytometrie (FCM). U druhu *Taraxacum stenocephalum* vykazujícího extrémní variabilitu velikosti genomu jsme provedli opylovací experiment, kde jsme křížili rodiče s různou velikostí genomu. Porovnali jsme velikost genomu rodičů a F1 generace. U potomků jsme také sledovali souvislost velikosti genomu a růstových charakteristik. U ploidně variabilních druhů *Galium valdepilosum* a *Arabidopsis arenosa* jsme se zabývali rozšířením zjištěných cytotypů a jejich ekologickými nároky. Na souboru druhů andského rodu *Lasiocephalus* jsme studovali relativní i absolutní velikost genomu, provedli jsme ITS sekvenování a otestovali možné hlavní příčiny této variability. Na souboru nepůvodních druhů jsme porovnáním s jejich neinvazními příbuznými testovali vztah velikosti genomu a invazivnosti rostlin. Protože cytometrické stanovení velikosti genomu má i svá omezení, snažili jsme se najít způsob, jak se vypořádat s potřebou čerstvých vzorků a otestovali protokol v laboratoři i v terénu.

Potvrdili jsme výjimečnou vnitrodruhovou variabilitu v rostlinách *Taraxacum stenocephalum* a ukázali, že velikost genomu potomstva je u tohoto druhu zřejmě určená průměrnou hodnotou velikostí genomu rodičů. Objevili jsme vztah mezi velikostí genomu a některými růstovými vlastnostmi. Studované cytotypy ploidně variabilních druhů se liší ve svých ekologických preferencích podle druhu, ke kterému patří. U rodu *Lasiocephalus* jsme zjistili, že velikost genomu koreluje s fylogenezí skupiny. Potvrdili jsme souvislost mezi velikostí genomu a invazivností. A podařilo se nám sestavit laboratorní protokol fixace jader, který umožňuje zpracovávat vzorky i ze vzdálených oblastí.

Tato práce přináší důkazy o tom, že velikost genomu může ovlivnit různé stránky života rostlin, a je proto důležitým markerem ve studiu rostlin. Velikost genomu nám například může poukazovat na různé evoluční procesy jako polyploidie či hybridizace. Její vliv a význam se ale liší u jednotlivých skupin rostlin a také podle úrovně studia (buňky, fenotyp, ekologie). Je proto vždy nutné přesně definovat otázky a vybrat si k jejich zodpovězení vhodný předmět studia.

INTRODUCTION

Genome size variation and its origin in plants

Genome size (hereafter GS) expressed as the total amount of nuclear DNA in picograms (pg) or number of base pairs (Mbp) – 1 pg = 978 Mbp – is among the basic characteristics of living organisms. The whole DNA content of a whole chromosome complement is called holoploid genome size and it is usually expressed as a C-value. The term 1C-value refers to the amount of DNA in the unreplicated reduced nucleus whereas 2C-value corresponds with the DNA content of the unreplicated non-reduced nucleus. Monoploid genome size is the averaged DNA content of one chromosome set with the chromosome base number x. The term Cx-value is used for monoploid genome size (2C-value divided by the ploidy level); in diploids 1Cx is identical with 1C-value (Greilhuber et al. 2005). C-values vary from 106 bp in Protists to 1011 in plants. Based on the tremendous variation of C-value in many different groups of organisms, Gregory (2001) formulated the term C-value enigma, which relates to the amount of non-coding DNA in nuclear genomes. In eukaryotes we have data only for a small fraction of species (about 10,000 species) and the variation in genome size ranges up to 66,000-fold (Pellicer et al. 2010). This enormous variation has fascinated scientist and gives rise to two questions: (i) How is it generated? and (ii) Why does it exist and how is it maintained? Despite the tremendous amount of studies that have been published and theories which have been formulated, some aspects of genome size have not been studied sufficiently or the studies have brought contradictory results. Many questions about genome size therefore still remain unanswered. This chapter summarizes the existing knowledge about the GS in flowering plants.

In angiosperms the range of genome size variation reaches up to more than 2,300-fold (Dodsworth et al. 2015). The smallest DNA content in angiosperms has been found up to now in tiny carnivorous plant *Genlisea tuberosa* – 61 Mbp (1C = 0.0624 pg) of DNA from the eudicot Lentibulariaceae family (Fleischmann et al. 2014). The largest genome size was revealed so far in *Paris japonica* – c. 150,000 Mbp (1C = 152.23 pg) from the monocot family Melanthiaceae (Pellicer et al. 2010) and it represents the largest eukaryotic genome at the same time. There is no direct proportional correlation of genome size, chromosome number or ploidy level with the complexity of organisms or even their taxonomic or phylogenetic position in the system (Gregory 2001). Although this variation in genome size is huge, it is not continuous and distinct patterns can be found between major evolutionary lineages. At the lower taxonomical level, genome size can be suitable means for distinguishing of closely related species, genera or families (e.g. Bureš et al. 2004, Chrtek et al. 2009, Prančl et al. 2014) and for revealing of evolutionary processes (e.g. polyploidy, introgression, hybridization, genome downsizing), which are not recognizable at first glance (e.g. Vítová et al. 2015). With increasing application of molecular sequencing, genome size has become an important value (Kelly et al. 2012).

The staggering GS variation is caused by changes at the chromosomal as well as molecular level. Genome size is determined by deletions of parts or the whole chromosomes, but polyploidy has a bigger impact at the chromosomal level. Polyploids are defined as

organisms with more than two complete chromosome sets (Otto & Whitton 2000, Comai 2005). From genomic point of view, genome doubling has occurred at least once during the evolutionary history of all angiosperms (Wendel et al. 2016), so polyploidy has affected all plants and represents the main driver of sympatric speciation in plants (Otto & Whitton 2000) and the major evolutionary force in plant diversification (Soltis et al. 2009, Weiss-Schneeweiss et al. 2013). The other process which contributes to genome enlargement is the activity of repetitive DNA, mainly LTR retrotransposons (SanMiguel et al. 1996, Vicient et al. 1999, reviewed in Bennetzen 2000, 2005, Neumann et al. 2006). Conversely, unequal homologous recombination, illegitimate recombination and deletion-biased double strand break (DSB) repair pathways cause the shrinking of genomes by small deletions (Shirasu 2000, Devos et al. 2002, Wicker et al. 2003, Bennetzen et al. 2005, Schubert & Vu 2016, Wendel et al. 2016). Some authors take the view that for organisms the loss of genome has greater significance than genome amplification, and this has been demonstrated in several species – *Laupala* cricket (Petrov et al. 2000), *Arabidopsis thaliana* (Devos et al. 2002, Ma et al. 2004, Bennetzen et al. 2005), five angiosperms (Vitte & Bennetzen 2006).

Nucleotypic theory and its manifestation in biological traits

From the fifties and sixties of the twentieth century, GS has become the subject of many studies considering its possible influence on plant life. This has lead to the formulation of the nucleotypic theory (Bennett 1972), which says that the nucleotype (i.e. nuclear mass itself) can influence the phenotype regardless of the information in the DNA. This theory has triggered further research testing the validity of the theory. The impact of nuclear DNA content, which is directly proportional to the nucleus volume, is most evident at the cellular level during the developmental stages of organisms and in conditions when the development is the slowest (Bennett 1972). Bennett has shown that genome size is associated with cell cycle time, meiotic duration and minimum generation time as a consequence of nuclei size, cell size and the cost of division. He studied the duration of the cell cycle and genome size in 31 angiosperms. Annuals, which had smaller genome size, had significantly lower mean cycle time than perennials with larger genome size. Moreover, ephemerals had significantly lower genome size than other annuals. Large-genome plants were obligate perennials whereas small-genome plants were represented in all life history strategies. Interestingly, facultative perennials did not differ from annuals. The study by Šímová & Herben (2012) tested how it works at the cellular level. The authors show that the relationship GS-cell size is determined by geometrical constraints, with the exponent 1/3 for the relationship between GS and cell diameter. For cell cycle duration, they distinguished two phases – GS-dependent S phase duration with exponent 1/3 and GS-independent G1 and G2 phase duration, which does not have the expected exponent.

Many of further studies testing the nucleotypic theory followed, revealing GS reflection also into minimum generation time (or time to flowering) (Leitch and Bennett 2007), seed characteristics (Grotkopp et al. 2004, Beaulieu, Moles, et al. 2007, Krahulcová et al. 2017) growth rate of seedlings (Grotkopp et al. 2004), specific leaf area (Morgan & Westoby 2005, Beaulieu, Leitch, et al. 2007), stomatal size and density (Beaulieu et al. 2008,

Hodgson et al. 2010), pollen size (Knight et al. 2010), flower size (Meagher & Vassiliadis 2005). Knight & Beaulieu (2008) tried to generalize the relationship between genome size and phenotype. They studied eight traits of different phenotypic scale – cell size (guard and epidermal), stomatal density, seed mass, leaf mass per unit area (LMA), wood density, photosynthetic rate and, finally, maximum plant height. They included 100–1,222 species of angiosperms and gymnosperms to test the significance of particular trait. The authors found that genome size affects the phenotype at the cellular level; there was a significant positive correlation of guard cell length and epidermal cell area and a significant negative correlation of GS and stomatal density. The rate of photosynthesis was weakly negatively correlated with genome size in gymnosperms, but there was no significant correlation across the angiosperms. At the higher phenotypic scale, the weaker association of genome size with the phenotypic traits. For LMA and seed mass, only a weak correlation was found and no correlation for wood density was recorded. Testing of maximum plant height showed that there is a significant negative relationship within the angiosperms. Nevertheless, small-genome angiosperm can reach large height whereas large-genome angiosperms have never been associated with large height (although the association of GS and plant height was significant in regression analyses and not in the case of independent contrast). The study shows the need for independent contrast statistics, which incorporate phylogenetic relationships of study species when different species are compared. Without these robust comparative methods, the authors would get an incorrect common result for all species pool, although there is opposite relationship for gymnosperms and angiosperms in some traits. Generally, small-genome species usually exhibit wide variation in phenotypic traits whereas large-genome species tend to have limited ranges of trait values. This has been verified, for example, in the cases of seed size, photosynthetic rate and SLA (Knight et al. 2005).

Strong correlations of genome size with traits can be found at the cellular level, but on the higher phenotype level they disappear (Knight & Beaulieu 2008), which can produce contradictory results. Indeed, some studies did not reveal a negative correlation of genome size with flowering period (e.g. Ohri & Pistrick 2001 – in 75 species in *Allium* – but all the *Allium* species under study have relatively large genomes, so the reason could be that no small genomes were included). The nucleotypic effect is exhibited even at the homoploid level (Loureiro et al. 2010), but it is the most apparent when different ploidy levels are compared. Snodgrass et al. (2017) studied the nucleotypic effect in diploid and allotetraploid *Gossypium* and they found positive correlation of GS with guard cell length, EPC surface area, and pollen grain diameter. However, not all research reached the same results. Tsukaya (2013) studied the size of cells of the leaf mesophyll, petal epidermis and pollen grains in diploids and tetraploids of *Arabidopsis thaliana*, and he did not reveal any correlation. However, this study was done on mutants and transgenic material in unnatural conditions, so it is a question whether selection did not play a role and sort out only some individuals in nature. Francis et al. (2008) carried out a more extensive survey of cell cycle times in root apical meristems and DNA content in 110 species of angiosperm (monocots as well as eudicots). They found positive correlation of GS and cell cycle time regardless of the ploidy level for all species as well as for the monocot group and the eudicot group.

Methodological problems of genome size estimation

Reliable assessment of genome size is a basic prerequisite of GS research. For many years we did not have a precise methodology and uniform stringent rules for GS estimation. In the past, Feulgen microdensitometry was the main means of plant DNA content estimation. This slide-based quantitative procedure quantifies the amount of light absorbed by isolated nuclei quantitatively stained with a DNA-specific dye (usually the Schiff's reagent) (Greilhuber & Temsch 2001). Over the last decades, the situation has changed and flow cytometry has started to prevail. Flow cytometry (FCM) is an effective and high-throughput method of analysing optical parameters of particles, which have been stained with fluorescent dye (Shapiro 2004). The particles individually pass in suspension through a narrow light beam which excites the stain, and the emitted light is captured with specific filters and converted to a digital signal. The amount of emitted light is directly proportional to the amount of stain or DNA, respectively (Shapiro 2004). After the introduction of flow cytometry into botany, this method became dominant for ploidy determination as well as absolute GS estimation, and the use of flow cytometry has increased dramatically in last decades (Doležel & Bartoš 2005; Greilhuber et al., 2007; Leitch & Bennett, 2007). Methodological comparative studies (Doležel et al. 1998, Vilhar et al. 2001) have revealed that FCM is more precise and that there are not many critical points during the procedure which could influence the results. Many old studies using Feulgen microdensitometry brought contradictory results and many of them have been refuted (see Greilhuber 2005). Greilhuber & Temsch (2001) studied the basic aspects of Feulgen densitometry and they critically revised some older studies and summarized the critical points of the Feulgen procedure (Greilhuber 2005, 2008). On the other hand, the big advantage of Feulgen microdensitometry is the possibility to visually check nuclei and its applicability for single nuclei.

Flow cytometry is nowadays more popular for GS estimation because it is a fast method with a simple protocol allowing for extensive population screenings. However, although flow cytometry is an efficient, high throughput and cheap method (Doležel 1991), it has also several limitations. Only differences greater than approximately 1.04-fold can be unambiguously detected as double peaks (Doležel & Göhde 1995), which means that we are still not able to reveal the full extent of GS variation. To detect reliably even small variation it is necessary to use the same internal standard during measurements, to use different stains (intercalating propidium iodide as well as AT-specific DAPI), to repeat measurements on different days and during the different seasons, and to use different machines if possible, and the best way how to prove the discovered variation is to produce double peaks by simultaneously running samples (Greilhuber 2005, Doležel et al. 2007, Greilhuber et al. 2007). Moreover, the choice of internal standards can influence the accuracy of measurements (Doležel et al. 2003, Doležel & Bartoš 2005). GS estimation also varies because of the presence of metabolic compounds, that affect the staining of DNA (Noirot et al. 2000, 2005, Price et al. 2000,(Price et al. 2000)(Price et al. 2000) Loureiro et al. 2006, Walker et al. 2006, Greilhuber et al. 2007, Temsch et al. 2008). Generally, it is necessary to use the best practice protocol (Doležel & Bartoš 2005), and for tiny differences in genome size make all measurements in single laboratory using the same standard (Doležel et al. 1998, 2007).

The big advantage of flow cytometry is the small amount of material needed. This make this method non-destructive, applicable even on endangered species or small seedlings

at the early ontogenetic stage (Doležel 1991). One disadvantage of FCM is the requirement of fresh material for best-practice FCM (Doležel & Bartoš 2005, Doležel et al. 2007). This condition cannot be met during long field trips to remote areas, where it is not possible to keep the material fresh until it is transported to the laboratory with a cytometer. One solution is to work with seeds – flow cytometry is applicable to germinating seedlings as well as to resting seeds. In many cases, however, genome size (including ploidy) of offspring is not identical with parental ones because of hybridization, introgression or production of unreduced gametes. Therefore, seeds can be used only for certain purposes, e.g. the study of reproductive pathways (Matzk et al. 2000). Another possibility is to use fixed plant material. Nuclei, cell or tissues can be fixed in ethanol (Jarret et al. 1995), ethanol:acetic acid or in formaldehyde. Hülgenhof et al. (1988) analysed cereal nuclei fixed in a 1:1 mixture of a 0.9% NaCl solution with ethanol and samples processed with ethidium bromide and olivomycine staining. For longer preservation the nuclei preserved in 30% glycerol and stored at 20°C were used (Chiatante et al. 1990). The modified protocol was used by Hopping (1993). Reasonable FCM histograms can also be achieved by analysing rapidly frozen plant tissues (Dart et al. 2004; Nsabimana and Van Staden 2006; Halverson et al. 2008; Cires et al. 2009), which can also be problematic in the field without electricity. However, none of these protocols has become familiar and they are not used for FCM of plants.

Rapid desiccation of plant tissues in silica gel appeared to be an effective method for relative GS estimation and thus for ploidy studies (Suda & Trávníček 2006a,b), and it has been used in many studies – e.g. for ploidy screening in *Senecio carniolicus* (Suda et al. 2007). But in most species it is not applicable to absolute GS estimation, because a shift of fluorescence can be present (Suda & Trávníček 2006a, Bainard et al. 2011). The use of herbarium vouchers works only for relatively fresh vouchers (only several years old) and it gives high-quality histograms only in some species; it was also successfully applied to mosses (Voglmayr 2000).

C-values are accumulated in the Plant DNA C-value database (Bennett & Leitch 2012), which also provides information about data source as well as an estimation method (Feulgen densitometry vs FCM or others). Up to date this database contains information about genome size for 8,510 plant species (7,135 angiosperms), which represents only about 2 % of currently known plant species (374,000 – Christenhusz & Byng 2016). Because of the mentioned methodological problems, data from the database should be used critically, as it also contains old data obtained using different protocols. Therefore, the data are not suitable for revealing intraspecific GS variation or for comparisons of small differences. The database is a useful tool mainly for extensive comparative studies revealing general trends at higher taxonomic levels.

Intraspecific genome size variation

Reliable detection of very small differences in GS and knowledge of intrapopulation GS dynamics is essential for understanding the early phases of the diversification of genome size among related species (Loureiro et al. 2010). Intraspecific variation is quite rare in plants (Loureiro et al. 2010) and most previous studies have been refuted because they used an incorrect methodological procedure. More than 200 papers have been published about intraspecific GS variation (Šmarda & Bureš 2010), but the vast majority of studies did not report real variation and demonstrated only artefacts. Only few studies adhering to stringent

methodological rules plausibly demonstrate GS variation within species and reasonably test its correlations. Recently, with the application of a best-practice FCM protocol and extensive screenings of genome size, the number of reliable records has increased.

Here are several examples of real GS variation at the homoploid level: *Anthoxanthum* (Chumová et al. 2015), *Arabidopsis thaliana* (Schmuths et al. 2004), *Curcuma* (Leong-Škorničková et al. 2007), *Dasypyrum villosum* (Obermayer & Greilhuber 2005), *Fagus sylvatica* (Paule et al. 2018), *Festuca pallens* (Šmarda & Bureš 2006), *Lollium perenne* (Sugiyama et al. 2002), *Senecio carniolicus* (Suda et al. 2007), *Picris hieracioides* (Slovák et al. 2009), *Taraxacum stenocephalum* (Trávníček et al. 2013) and *Zea mays* (Díez et al. 2013).

Most studies have only revealed existing variation at the homoploid level and did not study the GS variation itself. However, detailed studies of *Festuca pallens* populations represent ones of the most detailed recent studies of intraspecific GS variation and its aspects in wild population (Šmarda et al. 2007, 2010, Šmarda, Bureš, Horová & Rotreklová 2008). They studied *Festuca pallens* (Poaceae), a perennial grass with two ploidy levels. Both of them – diploid and tetraploid – exhibit GS variation at the inter- as well as intrapopulation level. The progeny in one tetraploid wild population exhibits large variation in genome size. In 562 seedlings from 17 mother plants, 1.188-fold variation was revealed. The genome size of seedlings was correlated with maternal genome size, but the offspring variation was even greater. The authors also searched for association of genome size with developmental traits in F1 seedlings. They counted leaves as a measure of developmental rate (Nemoto et al., 1995; McMaster 1997, 2005) and found correlation with genome size, which considered to be opposite to the nucleotypic theory (Bennett 1972). Šmarda et al. (2010) conducted, also on *Festuca pallens*, a competition experiment which proved the existence of stabilizing selection on genome size. They grew the seedlings in conditions with or without strong competition and estimated the genome size of individuals which survived. In pots with high competition there was significantly lower genome size variation where both extremes of the genome size distribution were eliminated. The authors assume that genome size may be under direct selection, which in this case is weak yet evolutionarily important process.

Extraordinary intraspecific GS variation was also reported in *Taraxacum stenocephalum* (Asteraceae) (Trávníček et al. 2013). The authors reported variation in DNA content of more than 20% at the inter- and intrapopulation level. They used a best-practice FCM protocol and checked the number of chromosomes, but did not reveal the real source of this variation. The authors assumed that this variation could be caused by the presence of homeologous chromosomes of different sizes as a consequence of retrotransposons activity (Šmarda & Bureš 2010).

Polyploidy and genome size

Polyploidy, or whole-genome duplication (WGD), is a process that directly affects genome size. It leads to the diversification of species and is an important mode of speciation in some groups. Taking into account that the nucleotypic effect is apparent even at the level of homoploid species, in polyploids the effect must be enhanced. Together with the proliferation of transposable elements and activity of small RNAs molecules, which are responsible for regulatory processes, multiple occurrences of polyploidy are responsible for the enormous genome size variation in angiosperms (Wendel et al. 2016).

Genomic research has made great progress over the last two decades and brought new insights into the role of polyploidy in genome architecture. We already know that WGS has occurred at least once during the evolutionary history of all angiosperms (Wendel et al. 2016). Many groups have even undergone multiple instances of polyploidy during their evolutionary history. Polyploidization is accompanied by changes in the genome, ranging from an immediate response at the DNA level (e.g. gene loss or recombination) to expression-level changes influencing the functioning of genes, and these changes can happen over millions of years (Wendel et al. 2016). Polyploidy is always followed by genome rearrangements (Leitch and Bennett, 2004), often accompanied by genome reduction (diploidization), which is sometimes followed by a further polyploidization event that can counteract the loss of indispensable genes during genome downsizing (Vu et al. 2015). Each plant genome is therefore at a different stage of genome development.

Genome size varies in polyploids depending on their age: In young polyploids genome size corresponds to the sum of their progenitors' genome sizes (Rebernig et al. 2012; Soltis et al. 2012) whereas in old polyploids it is significantly lower because of genome downsizing (e.g. Clarkson et al. 2005, Mandáková et al. 2010). Post-polyploidization processes leading to speciation are determined also by the origin of the polyploid. Polyploids that have originated via doubling of a single diploid genome (i.e. within an individual or after the hybridization of closely related individuals), are referred to as autopolyploids. Polyploids resulting from the hybridization of two or more different species are referred to as allopolyploids (Barker et al. 2016). A new gene combination created by hybridization determines many differences between the hybrid and the parental species. Allopolyploids face the fact that they possess at least two unequal genomes, which inevitably leads to genomic rearrangements that can have a wider consequences than the increase in genome size alone. Polyploidy then often goes hand in hand with species diversification, which is necessary for the establishment of a new polyploid entity. Every new polyploid faces obstacles such as various reproductive barriers (Husband 2004). Therefore, within-species ploidy variation might be a more suitable study system for investigating the effects of GS enlargement. In the cases of many species we do not know whether they are auto- or allopolyploid, but it has been estimated that these two types of polyploidy occur with similar frequency in nature (Barker et al. 2016).

Differences between plants of different ploidy levels stem from a whole array of changes, be it genetic, genomic, morphological or ecological, so it is difficult to separate the effect of genome size. However, even without hybridization, polyploidy is linked with genetic and phenotypic novelties (Levin 1983, Flagel & Wendel 2009; Hegarty et al. 2013) and is likely to be associated with changes in gene functioning (Soltis et al. 2014, Van De Peer et al. 2017, Wendel et al. 2016).

Some studies have explored the nucleotypic effect across species and genera differing in ploidy level and revealed positive correlations between genome size and cell size (e.g. Knight et al. 2005, Beaulieu et al. 2008, Knight & Beaulieu 2008, Hodgson et al. 2010, Veselý et al. 2012). They have also revealed negative correlations of GS with stomatal density (Beaulieu et al. 2008, Knight & Beaulieu 2008). Within species or genera, some direct correlations have been found at the cellular level, but the associations differ depending on the type of cells (Katagari et al. 2016). For example, a strong correlation between guard cell

length and pollen grain diameter has been found iare often tightly correlated with n *Gossypium*. However, in the case of epidermal pavement cell surface area the association was not so strong (Snodgrass et al. 2017).

Differences in ploidy level are are often tightly correlated with differences in reproductive traits (Husband et al. 2013) and fitness (Parisod et al. 2010), and polyploids often have different ecological niches and can occupy different regions (e.g. Hijmans et al. 2007, Ramsey 2011, Paule et al. 2017). Contact zones between new polyploids and their progenitors offer good opportunities for studying the effects of polyploidy (Husband 2004), and in mixed-ploidy populations it is possible to directly compare plants having different ploidy levels (and genome size) in the same environmental conditions.

Evolutionary and ecological consequences of genome size variation

With the increasing number of genome size studies, there is no doubt that GS variation has ecological and evolutionary consequences (Biémont 2008). Although such variability in genome size is huge, most plants have small genomes; the distribution of genome size is distinctly right-skewed (Dodsworth et al. 2015). The asymmetric GS distribution in angiosperms raises the question as to why small genome size is preferred in evolution or, converesly, why large genome sizes are disadvantaged in evolution. Knight et al. (2005) formulated the large genome constraint hypothesis, which says that large genomes are constrained because of the cost of possessing a larger genome consisting of unnecessary 'junk' DNA. They looked for support for this theory at three levels – evolution, ecology and phenotype. They found that genera with large genomes are likely to have smaller species diversity. Genome size influences average rates of diversification and extinction risk (Vinogradov 2003), but this relationship is not linear (Knight et al. 2005). At the ecological level, the hypothesis is supported by the fact that large-genome species are under-represented in extreme environments, while small-genome species occur in a wide range of habitats. Large genomes were also found to be rare in environments polluted by heavy metals (Vidic et al. 2009). Reduced maximum photosynthetic rates were found in species with large genomes. The phenotype level they investigated further in detail – see above and Knight & Beaulieu (2008). In general, sophisticated statistical methods are needed to reveal the real influence of genome size at the phenotypic and ecological level, and especially quantile regression methods represent a useful tool for uncovering these levels of complexity (Knight et al. 2005).

The large genome constraint hypothesis cleverly explains the manifestations of phenotypic properties at the level of ecology. The large-genome species has larger seeds, which limits their dispersal ability and their inability to occupy extreme habitats could make it harder to survive in long-term isolation and decrease the likelihood of allopatric speciation (Knight et al. 2005). Relationships between GS and temperature, precipitation, latitude and elevation of habitats have been reported with contradictory results (reviewed in Knight & Ackerly, 2002). The authors show that for elevation and latitude, all possibilities of relationship with genome size (positive, negative, non-significant) have been published, but most likely the northern latitudinal range limit tends to be negatively correlated with genome size (Bennett, 1987; Grotkopp et al., 2004; Knight et al., 2005). For precipitation, only positive or non-significant studies exist, temperature was negatively correlated in the sole

existing study, and seed size was positively correlated with genome size in all studies. Using data from the Plant DNA C-value database and quantile regression, the authors found out that small-genome species occur in all environments whereas large-genome species are more common with decreasing annual precipitation and prefer environments with intermediate mean July maximum temperatures (unimodal distribution of GS and July maximum temperature) (Knight et al. 2005). A recently published study about *Zea mays* reported a negative correlation of genome size with altitude (Díez et al. 2013).

Further impact of genome size has been detected in species occurrence and distribution. Endangered (rare) species have larger genomes than common species (Vinogradov 2003). Herben et al. (2012) reported an association between genome size and regional abundance, which is a consequence of an effect of GS on seed size, which is inversely proportional to the number of seeds, which in turn determines population growth rates in annuals. As DNA is composed mainly of nitrogen and phosphorus, the relationship between genome size and plant communities with different contents of these elements was studied (Šmarda et al. 2013, Guignard et al. 2016). Indeed, large-genome species are less competitive and can dominate only in communities rich in nitrogen and phosphorus, which shows that genome size is projected into level of plant communities.

The association of GS and life strategy (e.g. weed, invasive species) was also studied. A negative correlation of GS with weediness was found in 156 species (Bennett 1987). The association of genome size and invasiveness is based on the assumption that small genomes have faster division and therefore faster growth and minimum generation time (Bennett 1972), which are the advantageous properties for invasive species (see te Beest et al. 2012, Suda et al. 2015 for review). In addition, the ability of plants with small genome size to produce many light seeds and to establish rapidly (Rejmánek 1996) indicates that genome size could be related to plant invasiveness. Bennett (1987) also showed that small genomes are advantaged in time-limited environments. For these reasons, genome size has been considered one of the most important factors in plant invasiveness (Rejmánek 1996, 2000, Rejmánek et al. 2005). This association has been confirmed in the genera *Pinus* (Wakamiya et al. 1993, Grotkopp et al. 2002, 2004), *Senecio* (Lawrence 1985) and *Acacia* (Mukherjee and Sharma 1990). Another study revealed a negative correlation between genome size and invaded spread in three species of *Briza* (Rejmánek 1996). In addition, more extensive studies with interspecific comparisons show that small genomes favour invasiveness (Bennett et al., 1998; Knight & Ackerly, 2002; Kuester et al., 2014; Pandit et al. 2014). Furthermore, genome size-related differences can further influence the the rate of colonization, which was studied at the within-species level in *Phalaris arundinacea* (Lavergne et al. 2010). Finally yet importantly, Pyšek et al. (2018) showed that the amount of DNA in the cell nucleus can also be used to distinguish between native and invasive populations in *Phragmites australis*. They compared the North American wild population with invasive populations established by plants that originated in Europe, and monoploid genome size separated these populations. Plants from invasive populations generally possess more traits that enable them to be successful invaders (long rhizomes, early emerging abundant shoots, resistance to aphid attack, and a low C:N ratio). The role of genome size in plant invasiveness is a hot topic in invasion biology and it deserves further research because it might have even practical implications in the fight against invasive species.

The asymmetric distribution of genome size (Dodsworth et al. 2015) has evoked the suggestion that a large gnome on the one hand is a limitation and on the other that small

genome size constitutes an evolutionary advantage. Large genomes are rare in all eukaryotes (Gregory 2005) and many studies have presented theories about this asymmetry, suggesting that large genomes are selectively removed (Vinogradov 2003, 2004; Knight et al. 2005). Several comprehensive studies have already been published about GS evolution (Soltis et al. 2003, Bennetzen et al. 2005, Leitch et al. 2005, Dodsworth et al. 2015, Pellicer et al. 2018) submitting different scenarios for GS evolution, and there is still no clear opinion regarding the role of natural selection in relation to genome size.

The study by Oliver et al. (2007) demonstrates that the skewed distribution of eukaryotic genome size can be explained with a simple proportional model of evolution. The authors suggest that evolutionary changes in GS are less pronounced in small genomes than in large genomes. Therefore, it is more difficult for small genomes to become and stay large and easier for large genomes to become and stay small. This model shows that there is no need to invoke natural selection to explain the distribution of GS in eukaryotes. Studies in the Brassicaceae demonstrate that neutral evolution through random and passive changes may also be important in shaping genome size (Lysak et al. 2009). Another study claims that GS in angiosperms does not correlate with effective population size (Whitney et al. 2010). Presuming that effective population size is related to selection, this is another indication that selection does not play a big role in genome size evolution. Though the particularly noteworthy study of Oliver et al. (2007) presented a reasonable explanation, further studies are being conducted to reveal the correlation of genome size with phenotypic consequences, which is the logical way how to find out if the suggested model is true.

The large genome constraint hypothesis (Knight et al. 2005) assumes that genome size is adaptive based on interspecific comparison. On the other hand, the study of intraspecific GS variation have shown that genome size is not always adaptive (Šmarda et al. 2007), wherein no difference was found in GS variation in spatial distribution in a highly variable population. The authors conclude that genome size in this case is non-adaptive and not affected by selection.

The association between genome size and phylogeny has been well documented. Another aspect of genome size is its relationships to the phylogenies of taxa under study. In *Orobanche*, differences in genome size correspond with main phylogenetic lineages (Weiss-Schneeweiss et al. 2006). In *Hieracium* subgenus *Hieracium*, two groups of taxa possessing different genome sizes are congruent with two phylogenetic groups (Chrtek et al. 2009). In *Chrysanthemum*, too, a correlation of genome size with phylogenetic groups has been reported (Luo et al. 2017). Anyway, phylogenetic signal can blur the association of genome size with biological parameters and it is necessary to make phylogenetic corrections when comparing different taxa.

Ecological and evolutionary processes are intertwined, the relationships evolve over time. A noteworthy study conducted by (Jordan et al. 2015) demonstrates that environmental adaptation, and not genome size, has determined the stomatal size in the Proteaceae. The authors suggest that ancient changes in GS affected stomatal size during the species' evolution but that adaptation to the environment altered the stoma-GS association. However, GS studies also face another obstacle: When revealing relationships between genome size and certain traits, a significant relation between genome size and some trait can be overshadowed by a

third factor (Beaulieu, Leitch, et al. 2007). Generally, considering the impact of genome size, the different time scale might have to be taken into account. On a developmental/ecological time scale, genome size could have a direct effect whereas over the long-term course of evolution there should be a selective feedback loop from traits to genome size (Herben et al. 2012), which makes the issue even more complicated.

OBJECTIVES AND STUDY SYSTEMS

The main objective of this thesis was to clarify the significance of genome size for plant life and demonstrate the importance of this marker in botanical studies. Summarized below are the main topics and questions touched in the articles composing this dissertation:

1) Dynamics of genome size in natural systems

How is genome size variation transferred between generations? How heritable is parental divergence in genome size?
What is the range of variation at the homoploid and heteroploid levels within species?

2) Genome size and its association with phenotype

Which biological traits or processes are associated with genome size?
Do morphological and phenological correlations with genome size exist at the intra- and interspecific levels?

3) Ecological and evolutionary consequences of genome size

What are the ecological and evolutionary consequences of genome size variation? How is genome size associated with invasiveness? Is genome size determined by phylogeny?

4) Limits of FCM genome size estimation

How to overcome the need for fresh samples during the collection of samples in remote areas or on long field trips?

All these questions mingle in the six case studies, which were focused on the importance of genome size in different groups of plants at different taxonomic levels (intraspecific, intrageneric, interspecific), studying genome size at different levels of complexity (intraspecific genome size vs ploidy levels), various associations of genome size with plant traits, different ecological correlations and consequences (distribution range, habitats, invasiveness), and also addressing possible methodological obstacles (the need for fresh material and combining of different protocols).

Some of the included case studies did not deal exclusively with genome size. Below is an overview of the aspects of genome size each study explored, the model systems studied and the methods used.

I. Lučanová M., Loureiro J., Suda J. (manuscript): Genome size variation correlates with fitness indicators in *Taraxacum stenocephalum* (Asteraceae)
- *Taraxacum stenocephalum* (Asteraceae)
- 775 seedlings from 25 mother plants

- intraspecific genome size variation – its range, heritability, dynamics and correlation with developmental traits
- absolute genome size
- flow cytometry, karyology, crossing experiment

II. 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

- *Galium valdepilosum* (Rubiaceae)
– studying ploidal distribution and different ecological preferences of cytotypes
- relative DNA genome size screening
- interspecific variation in absolute DNA
- flow cytometry

III. Kolář F., Lučanová M., Záveská E., Fuxová G., Mandáková T., Španiel S., Senko D., Svitok M., Kolník M., Gudžinskas Z., Marhold K. (2016): Ecological segregation does not drive the intricate parapatric distribution of diploid and tetraploid cytotypes of the *Arabidopsis arenosa* group (Brassicaceae). – Biological Journal of the Linnean Society 119: 673–688

- *Arabidopsis arenosa group (Brassicaceae)*
- more than 2,900 individuals from 194 populations
- testing of ecological segregation of parapatrically distributed 2x and 4x
- relative DNA genome size screening
- homoploid differentiation in DNA content
- flow cytometry

IV. Dušková E., Kolář F., Sklenář P., Rauchová J., Kubešová M., Fér T., Suda J. & Marhold K. (2010): Genome size correlates with growth form, habitat and phylogeny in the Andean genus *Lasiocephalus* (Asteraceae). – Preslia 82: 127–148

- *Lasiocephalus* (Asteraceae)
- 20 species, 189 individuals from 101 populations
- correlation with growth form, habitat, altitude, and phylogeny
- absolute genome size
- relative genome size
- flow cytometry, karyology, ITS sequencing

V. 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

- 93 alien species from 32 families naturalized in the Czech Republic
- comparison of absolute genome size of naturalized plants and non-invading congeners

- absolute genome size
- flow cytometry

VI. Kolář F., Lučanová M., Těšitel J., Loureiro J. & Suda J. (2012): Glycerol-treated nuclear suspensions – an efficient preservation method for flow cytometric analysis of plant samples. – Chromosome Research 20(2): 303–315 https://doi.org/10.1007/s10577-012-9277-0

- new protocol for nuclei preservation in glycerol
- time-scale laboratory experiment – 6 species from 5 families
- in situ application of glycerol preservation – 21 species from 12 families
- relative and absolute genome size estimation
- flow cytometry

Case study I

Lučanová M., Loureiro J., Suda J.: **Genome size variation correlates with fitness indicators in *Taraxacum stenocephalum* (Asteraceae)**

Genome size variation correlates with fitness indicators in *Taraxacum stenocephalum* (Asteraceae)

Magdalena Lučanová[1,2], João Loureiro[3] & Jan Suda[1,2,*]

[1] The Czech Academy of Sciences, Institute of Botany, Zámek 1, CZ-252 43 Průhonice, Czech Republic
[2] Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, CZ-128 01 Prague, Czech Republic
[3] Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Calçada Martim de Freitas, 3000-456, Coimbra, Portugal
*deceased

E-mail for correspondence: kubesovm@gmail.com

Abstract

Intraspecific variation in plant genome size has only rarely been rigorously documented, mainly due to methodological issues, and its potential adaptive value remains a matter of debate. We tested for the heritability of genome size and its consequences for plant fitness in *Taraxacum stenocephalum*, a sexual tetraploid exhibiting remarkable variation in genome size in its wild populations.

Using flow cytometry complemented with karyology we assessed the range of genome size variation among 150 individuals of the study species. We then carried out controlled crossings of 25 mother plants and 22 pollen donor plants, and estimated the genome sizes of 775 seedlings of the F1 generation. In a common garden experiment we examined several proxies of fitness of progenies varying in nuclear DNA content.

We recorded a strong correlation between parental and F1 offspring genome size. The genome size variation of F1 seedlings was greater than that of parental plants. During chromosome counting we observed the presence of satellite chromosomes of various sizes. The largest satellites occurred only in plants with large genomes, suggesting that they contribute to the genome size variation observed. The analyses revealed correlations of several traits with genome size of F1 plants. Mid-sized genomes were associated with the best fitness.

This study demonstrates that genome size is a heritable trait exhibiting a pattern of segregation that correlates with certain indicators of fitness. Mid-sized genomes correspond with phenotypes exhibiting the greatest fitness, suggesting the occurrence of balancing selection for mid-sized genomes. Genome size variation can be important in early ontogeny, possibly influencing offspring survival and population phenotype display.

Key words: crossing, fitness, flow cytometry, genome size, heritability, intraspecific variation, nuclear DNA content

Introduction

Genome size is one of the basic characteristics of living organisms. In the different kingdoms of life it exhibits massive variation, which in the Angiosperms spans a 2,300-fold range (Dodsworth et al. 2015). This tremendous variation is caused mainly by polyploidy and by the activity of repetitive DNA elements, mainly LTR retrotransposons (SanMiguel et al. 1996, Vicient et al. 1999, SanMiguel & Bennetzen 1998; Bennetzen 2000, 2005; Neumann et al. 2006, Wendel et al. 2016), which lead to increases in genome size. Processes like unequal homologous recombination and illegitimate recombination are also involved; however, through small deletions, these mechanisms cause the shrinking of genomes (Shirasu 2000, Devos et al. 2002, Wicker et al. 2003, Bennetzen et al. 2005).

Genome size can affect the phenotypes of plants. According to the so-called nucleotypic theory (Bennett 1972) it is known to influence the duration of cell division, the size of the nucleus, and thus the whole cell and its minimum generation time (Bennett 1972, Bennett 1973, Edwards & Endrizzi 1975, Cavalier-Smith 1985). Many correlative studies have shown that genome size can be viewed as an adaptive trait influencing minimum generation time or time to flowering (Leitch and Bennett 2007), seed characteristics (Grotkopp et al. 2004, Beaulieu, Moles, et al. 2007, Krahulcová et al. 2017), relative growth rates of seedlings (Grotkopp et al. 2004), specific leaf area (Morgan & Westoby 2005, Beaulieu, Leitch, et al. 2007), stomatal size and density (Beaulieu et al. 2008, Hodgson et al. 2010), pollen size (Knight et al. 2010) and flower size (Meagher & Vassiliadis 2005), as well as their relations with environmental gradients (Knight & Ackerly 2002), including invasiveness (Suda et al. 2015) and their association with phylogeny (Chrtek et al. 2009, Luo et al. 2017). However, most studies deal with the interspecific level and only little is known about the role of genome size in driving phenotypic and possibly fitness changes within particular species with homoploid genome size variation between generations.

This lack of information is caused mainly by the fact that, within a species, variation in homoploid genome size is usually low (Loureiro et al. 2010), and intraspecific variation is considered a rare phenomenon in wild plant species (reviewed in Šmarda & Bureš 2010). Moreover, earlier studies were hindered by methodological obstacles (for reviews see Greilhuber 2005, Šmarda & Bureš 2010), and the only studies that have reliably documented true intraspecific or even intrapopulation variation in genome size using rigorous methodology have been published in the last two decades, for example in *Anthoxanthum* (Chumová et al. 2015), *Arabidopsis thaliana* (Schmuths et al. 2004), *Curcuma* (Leong-Škorničková et al. 2007), *Dasypyrum villosum* (Obermayer & Greilhuber 2005), *Fagus sylvatica* (Paule et al. 2018), *Festuca pallens* (Šmarda & Bureš 2006), *Lollium perenne* (Sugiyama et al. 2002), *Hieracium* sub-genus *Pilosella* (Suda, Krahulcová, et al. 2007), *Senecio carniolicus* (Suda, Weiss-Schneeweiss, et al. 2007), *Picris hieracioides* (Slovák et al. 2009) and *Taraxacum stenocephalum* (Trávníček et al. 2013). Most studies published to date report low intraspecific genome size variation (but see Šmarda, Bureš, Horová, & Rotreklová 2008, Trávníček et al. 2013), raising further questions about the significance of genome size variation. Nevertheless, these studies have shown that the study of intraspecific variation in genome size may provide key clues towards understanding the mechanisms driving genome size variation in wild plant populations. From this point of view, the detailed studies by Šmarda, Bureš, Horová & Rotreklová (2008) and Šmarda et al. (2010) were pioneering. These authors studied genome size dynamics in highly variable populations of *Festuca pallens* and found a correlation between the genome size of offspring and that of mother plants, assuming that

genome size is a heritable trait. Unfortunately, the authors did not determine the genome size of pollen donor plants, as they worked with fruits collected in wild populations.

One rare case of a species with large intraspecific variation in nuclear DNA content that has been rigorously documented is the Eurasian dandelion *Taraxacum stenocephalum* Boiss & Kotschy ex. Boiss. (Asteraceae, Lactuceae; Trávníček et al. 2013). A detailed flow-cytometric survey, following the best practices (Doležel & Bartoš 2005), of two natural populations documented 1,223-fold differences, with karyological analyses ruling out differences in chromosome numbers as an explanation for the observed pattern. Unlike most species of this genus, *T. stenocephalum* is a sexual tetraploid (Kirschner et al. 1994) with rapid ontogenetic development, which makes it an ideal study system for unravelling the dynamics of intraspecific variation in genome size.

The main objectives of this study were to disentangle the dynamics of genome size variation between generations of *T. stenocephalum* and to unveil possible correlations between nuclear DNA content and fitness traits. In particular, using flow cytometry, karyology, crossing experiments and fitness analyses of F1 plants, we aimed to further our understanding of how parental genome size is projected into successive generations and how, if at all, it influences the fitness of seedlings. The meeting of these objectives allowed us to address the following questions: (i) To what extent is genome size heritable? In particular, is there a correlation between the genome sizes of parental individuals and their offspring?; (ii) Which karyological traits, if any, underpin genome size variation; and, (iii) Does genome size affect any phenotypic or fitness traits of F1 individuals?

Materials and methods

Plant material

The present article is based on the study of 150 individuals of *Taraxacum stenocephalum* collected by Jan Kirschner and Frederick Rooks from nine wild subpopulations in two regions of Georgia (Lesser Caucasus – Trialetis Range, Greater Caucasus – Tergi River and Kazbegi). For details see Supplementary Table S1 and Trávníček et al. (2013). The individuals were cultivated from rhizomes collected in the field in the experimental garden at the Institute of Botany in Průhonice, Czech Republic.

Genome size estimations using flow cytometry

Absolute genome size of plant material was ascertained by flow cytometry using a proven two-step protocol (Doležel et al. 2007). Approximately 0.5 cm^2 of fresh tissue of the sample material was chopped together with the same amount of the internal reference standard *Glycine max* cv. Polanka with a sharp razor blade in 0.5 ml of ice-cold Otto I buffer and incubated for approximately 10 min at room temperature. The nuclear solution was filtered through a 42-µm-pore nylon mesh and a 1 ml of Otto II buffer supplemented with propidium iodide (PI), RNase (both at the final concentration of 50 µg/ml) and 2-mercaptoethanol (at the final concentration of 2 µl/ml). After 10 min of incubation, each sample was run through a Partec CyFlow SL flow cytometer (green laser with the central wavelength of 532 nm, Partec GmbH, Münster, Germany), and data on

5,000 particles were acquired. Only intraindividual measurements which differed by less than 2% from other measurements of the individual were accepted.

To verify the genome size of the internal reference standard with the use of our flow cytometer, the genome size of *Glycine max* cv. Polanka was recalculated against that of *Pisum sativum* cv. Ctirad (2C = 9.09 pg of DNA, Doležel et al. 1998) on three different days. According with this evaluation, the genome size of the reference standard was estimated to be 2C = 2.39 pg instead of the value reported by Doležel et al. (1994), i.e. 2C = 2.50 pg. Therefore, values of holoploid genome size (*sensu* Greilhuber et al. 2005) of *Taraxacum* samples were calculated as the ratio of PI fluorescence of sample vs standard G_1 peaks, multiplied by 2.39.

According to the results of genome size analyses, the cultivated (parental) individuals of *Taraxacum stenocephalum* were divided into five categories for the purpose of further crossings and the obtainment of an F1 generation. Specifically, we determined the entire genome size range for all individuals and divided the set of plants by 20%, giving us five groups (I – very small, II – small, III – medium, IV – large, V – very large genome size). Leaves of sterile rosettes of the resulting offspring were analysed following the same procedure as above. One measurement of genome size was made and only the minimum and maximum values were checked by the second and the third replications on different days. In total, the genome size of 559 seedlings (Set 1) raised from seeds from 30 fruit heads (one fruit head is from one parental combination, hereafter referred to as a family), which corresponded to the different types of crossings, was assessed for 10 to 56 individuals per offspring family (19 on average). In addition, 300 achenes (Set 2) were sown for the assessment of fitness and an estimation of genome size of the plants that survived until the adult stage was obtained (216 individuals in total). The nuclear DNA content of these plants was measured on three different days and, as above, when the difference between the three measurements was greater than 2%, the most extreme value was excluded and new analyses were carried out.

The aforementioned differences in nuclear DNA content were further confirmed by the presence of double peaks obtained by simultaneous analyses of individuals where such differences were detected. To this end, and in order to obtain G_1 peaks with the maximum possible resolution, each sample was prepared as above, but the Otto II buffer was enriched with DAPI (at the final concentration of 4 μg/ml) and 2-mercaptoethanol (at the final concentration of 2 μl/ml). These samples were then run until the acquisition of 5,000 data points through a Partec CyFlow ML machine (Partec GmbH, Münster, Germany) equipped with a UV LED diode with the output wavelength of 365 nm. Histograms were evaluated using FloMax Software (version 2.6, Partec).

Karyological analyses

The karyotypes of 27 F1 individuals that resulted from 10 different crossings (I × I, I × II, II × II, II × V, III × II, III × V, IV × IV, IV × I, V × III, V × IV) of 12 mother plants, covering all genome size categories, were investigated. Karyological accessions were obtained from the roots of seedlings. Because this technique is destructive, it was not possible to estimate the genome sizes of individuals which were used for the karyological analyses. The seedlings were pretreated in 0.002M 8-hydroxyquinoline (2.5 hours in the dark at room temperature and 2.5 hours at 4°C). The washed seedlings were fixed in an ethanol:acetic acid solution (3:1) and stored at −20°C until further processing. Before the making of chromosome preparations, the seedlings were macerated in 5N HCl for 30 min. Schiff's reagent was then added and the material was incubated in the dark for one hour at room temperature. The material was then transferred on to a microscopic slide in a drop of 60% acetic acid, non-meristematic tissues were removed, and the meristem was covered

with a coverslip and squashed. The preparations were examined under a ZEISS Axioplan 2 epifluorescence microscope (Zeiss), and digital images were acquired with a cooled CCD camera and further processed with AxioVision analysis software (Zeiss). Only slides where at least five mitoses were found were considered.

Crossing experiments (Set 1)

In the flowering season (June – July), flower buds of parental plants were bagged in non-woven fabric to exclude flower visitors. Open flowers throughout the inflorescences of plants belonging to each of the five genome size categories were pollinated reciprocally. Mother plants belonging to each genome size category were pollinated by pollen from donor plants belonging to all genome size categories, yielding a total of 25 crossing combinations. Each pollination was repeated on two subsequent days if the flowers re-opened. In total, 443 F1 crosses were carried out, comprising on average 17 F1 crosses (between 7 and 54 F1 crosses) for each combination of genome size categories. Besides, 236 bagged plants were left without a crossing treatment and served as controls for apomixis or autogamy. In 11 cases, an autogamy test was done by reciprocally crossing two inflorescences within the same plant. In addition, three plants were emasculated to exclude apomixis. Bags containing mature achenes were collected at the end of the fruiting season, the fruits were sorted, and the success of the crossing experiment was assessed. A subset of the fruits from 30 crossings of 25 mother and 22 donor plants was sowed, and the resulting seedlings (Set 1) were grown in the greenhouse on commercial garden substrate until the rosette stage.

Growth experiments (Set 2)

Another subset of the fruits was used to study the association between genome size and growth traits in F1 seedlings (Set 2). To this end, 300 fruits representing the offspring of crossings involving plants with the smallest and the largest genomes (I × I, I × II, IV × V, V × V) were individually weighed on a pair of Mettler AE 163 scales, soaked in potassium permanganate for 3 min, then sowed on to damp filter paper in Petri dishes and left to germinate. In one month the seedlings were large enough to be transplanted into pots with sterile commercial soil. The pots were placed at random positions in the experimental garden. In winter the pots were transferred to a greenhouse with better controlled conditions. Over the whole duration of the experiment, the positions of the pots were regularly re-randomized.

At regular intervals for 13 months, the following growth traits were recorded: beginning of germination, number of leaves, length of the longest leaf and beginning of flowering. After the last measurement, the total aboveground biomass of each individual was ascertained by drying the plants for 24 hours at 80°C and weighing the dried material on a pair of Mettler AE 163 scales.

Statistical analyses

In analysis of Set 1 we tested the relationship between mean parental and offspring genome sizes using linear regression. The difference in the distribution of successful and

unsuccessful pollinations between crossing types was tested by a t-test for dependent samples. The correlations of genome size with time to germination, time to flowering, weight of achenes and weight of aboveground biomass were tested in analysis of Set 2 also with linear regression.

In analysis of Set 2, the dependency in the number of offspring leaves in relation to genome size and time was fitted with generalized linear mixed-effect models ('lme4' R package, R Core Team 2017) and tested with a likelihood ratio test. First, a model considering the interaction of time and genome size was tested to evaluate whether genome size-related differences change over time. This model was fitted using all 23 measurements taken over the course of the observation period. Afterwards, genome size and time were entered as quantitative predictors, individuals were considered a factor with random effect, and the number of leaves was considered a dependent variable with an assumed Poisson distribution of unexplained variation.

In analysis of Set 2, the progression of offspring leaf lengths was fitted with a linear mixed-effect model ('lme4' package) with log-transformed leaf length. First, the same model as above was tested to analyse the interaction between time and genome size. Then, genome size and time were entered as quantitative predictors, individuals as a factor with random effect and leaf length as a dependent variable. Based on the identified genome size-related changes in leaf length, we fitted the effect of genome size with a second-order polynomial term. We used only the first 12 data points for leaf length, as this variable represented the period of increasing leaf length (see Fig. S1 in the supplementary material).

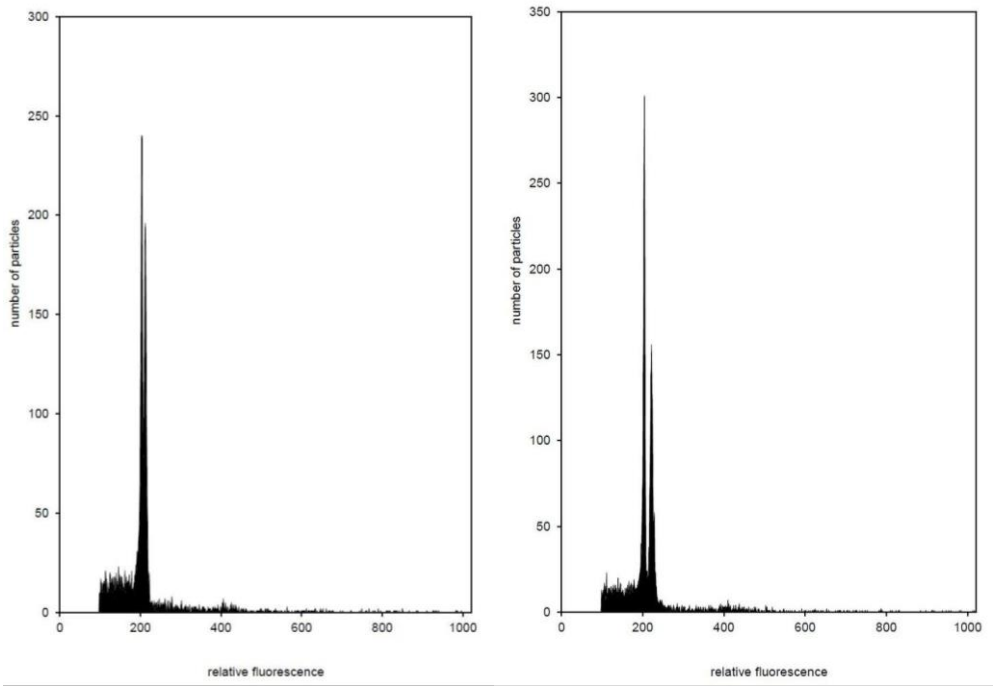
The associations between fruit weight and aboveground biomass and between leaf length and aboveground biomass at the end of the growth experiment were tested in analysis of Set 2 using linear regressions.

Statistical analyses were performed in R (R Core Team 2017) or, alternatively or in addition, in STATISTICA 13 (StatSoft Inc., 2015).

Results

Genome size variation and its dynamics

The nuclear DNA content of the parental plants used in the crossings ranged from 3.785 to 4.469 pg/2C, constituting a 1.181-fold variation in genome size (see Supplementary Table S2 for individual values). The genome size variation among 559 successfully germinated seedlings originating from 30 crossings between 25 mothers and 22 pollen donors was slightly greater (1.228-fold), ranging from 3.707 to 4.553 pg/2C (Set 1, Supplementary Table S2). This genome size variation has to be considered genuine because we acquired similar results repeatedly and on different days, as proved by the presence of double peaks in histograms from simultaneous analyses of individuals with distinct genome sizes (Fig. 1).



A B

Fig. 1 Flow-cytometric histograms of simultaneous analyses of offspring using DAPI. A. Seedlings from crossings between the genome size categories I and IV, coefficient of variation 1.22 and 1.3 for the first and second peak, respectively. B. Seedlings from crossings between the genome size categories I × I and III × V, coefficient of variation 1.13 and 1.78 for the first and second peak, respectively.

The mean genome sizes of offspring families were significantly positively correlated with theoretically expected values (averages for each parental pair – $r^2 = 0.84$, $F(1,28) = 143.47$, $p < 0.001$, Fig. 2). To reveal possible deviations of genome size from expected values derived from differences between parents, we examined the standard deviations of genome sizes of offspring families but only weak and statistically insignificant relationship was found ($r^2 = 0.10$, $F(1,28) = 3.11$, $p = 0.089$, Fig. 3). This indicates that the observed genome sizes vary in accordance with the expected values and that the genome size of the mother or the pollen donor caused no detectable shift. Generally, the nuclear DNA content of offspring differed from the mean parental genome size approximately by a maximum of 3% in both directions (Fig. 4). The overall variation in genome size among all seedlings even exceeded that observed in parental plants (3.785 to 4.469 pg/2C, 1.181-fold). Within seedling families (from 10 to 56 individuals per family) genome size varied 1.019 to 1.113-fold. The greatest difference between two parental plants was 1.17-fold.

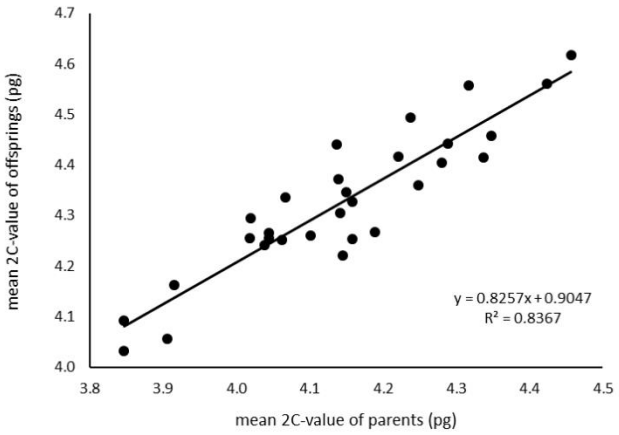


Fig. 2 Mean 2C-values (pg) of parental pairs and mean 2C-values (pg) of offspring families. The relationship is fitted with linear regression.

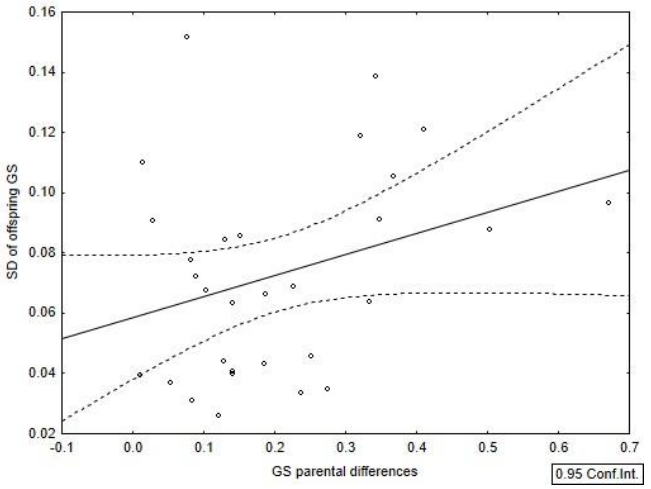


Fig. 3 Standard deviations of offspring genome sizes and differences between parents (in absolute units).

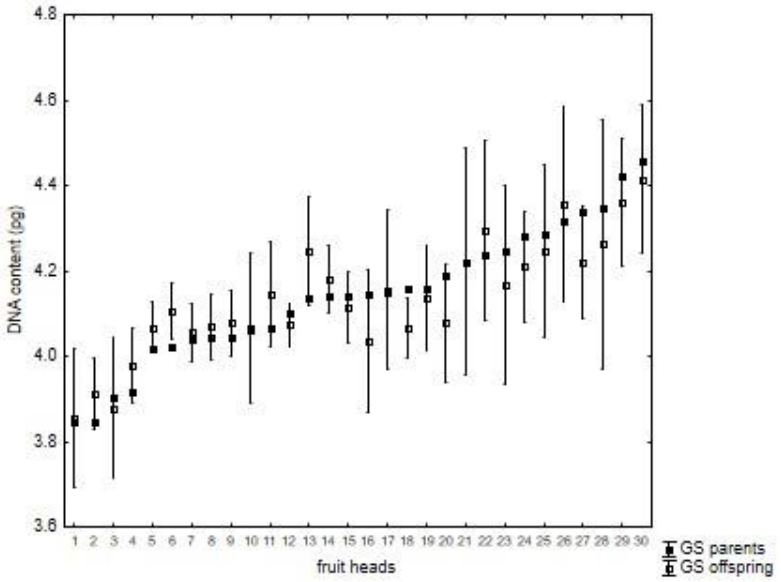


Fig 4. Scatterplot of 2C DNA content of 30 offspring heads (squares) and mean parental genome size (circles). Box = mean, whisker = mean ± 2 SD.

We found no differences in the distribution of crossing combinations between successful and unsuccessful pollinations (t-test for dependent samples, t = 0, d.f. = 24, p = 1), which indicates the absence of a reproductive barrier between plants having different genome sizes. The genome seizes of the 216 F1 seedlings (Set 2) which survived until the end of the experiment ranged from 3.772 to 4.678 pg/2C, amounting to a 1.24-fold variation in genome size. Once again, this range of variation is broader than that observed in parental plants, which ranged from 3.796 to 4.503 pg/2C, i.e. a 1.19-fold difference. The correlation between parental and offspring genome size was tested with linear regression (r^2 = 0.769, F(1,214) = 714.15, p < 0.001), which revealed a significant positive correlation: the greater the mean parental genome size, the greater the genome size of the offspring.

The mean coefficient of variation of G_1 peaks of samples and the internal standard was 2.78 % and 3.52 %, respectively, and the mean difference in genome size between three analyses performed on different days was 1.18%.

Sources of genome size variation

The chromosome number of *Taraxacum stenocephalum* was confirmed to be consistently $2n = 4x = 32$ (Fig. 5) in 108 successfully scored mitotic figures from meristem tissue of 27 individuals. However, we observed satellite chromosomes of different size in 67 mitotic figures (62% of individuals, see Supplementary Table S3).

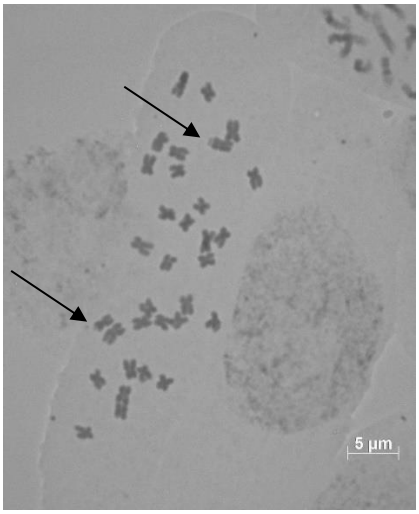


Fig 5. Mitotic chromosomes 2n = 32 from the root tips of seedlings from the crossing of parents belonging to the II x II category (2C mother plant = 4.19 pg, 2C donor plant = 4.19 pg). Arrows point to satellite chromosomes.

We examined the possible correlation between the presence of satellite chromosomes of different size and plant genome size. As it was not possible to flow cytometrically estimate the genome sizes of seedlings used for chromosome counting, because they were too small, we calculated their potential genome size as the arithmetic mean of their parents' nuclear DNA content. This was justified by the fact that we observed a strong correlation between the genome sizes of parents and their offspring (see below). Large satellite chromosomes were only present in plants with potentially large genome size (above 4.22 pg), so they probably contribute to the observed genome size variation. By contrast, small satellites occurred in seven samples that represented the full range of potential genome sizes (3.91–4.35 pg), indicating the lack of an association between the presence of small satellite chromosomes and genome size.

Genome size and offspring fitness

To assess possible correlations between fitness and genome size, we analysed Set 2 to search for a possible correlation between the genome sizes of the 216 surviving individuals and several of their growth traits.

Time to germination was weakly positively correlated with genome size ($r^2 = 0.0196$, $F(1,214) = 4.285$, $p = 0.0396$) (Fig. 6). On the other hand, there was a moderate negative correlation between achene weight and genome size ($r^2 = 0.269$, $F(1,214) = 78.723$, $p < 0.001$). The linear regression of time to germination and achene weight turned out to be non-significant ($r^2 = 0.00264$, $F(1,269) = 0.712$, $p = 0.399$). No correlation was detected between the number of leaves and genome size, as both genome size and the time*genome size interaction were non-significant (GS – $\chi^2_1=0$, n.s.; time:GS – $\chi^2_1=1.4644$, n.s.).

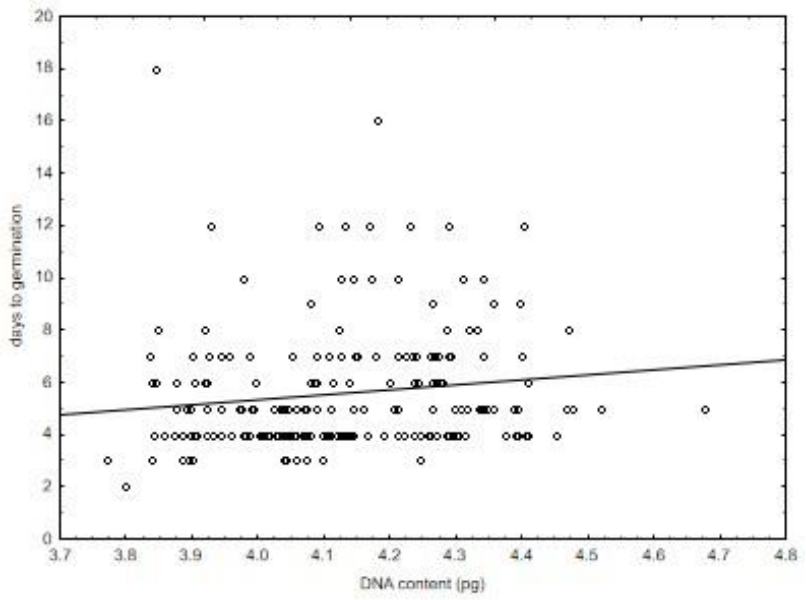


Fig 6. Linear regression of genome size and days to germination.

There was a unimodal dependency of leaf length on genome size over the whole observation period (poly(GS,2): $\chi^2_2 = 8.32$, $p = 0.0156$; Fig. 7). The addition of an interaction between the (polynomial) effect of genome size and time significantly improved the model ($\chi^2_2 = 12.497$, $p=0.0019$), providing evidence for an effect of genome size over time. The obtained results also demonstrate that genome size affected the growth of leaves, with the mid-sized genome size category exhibiting the maximum values both for the number and length of leaves.

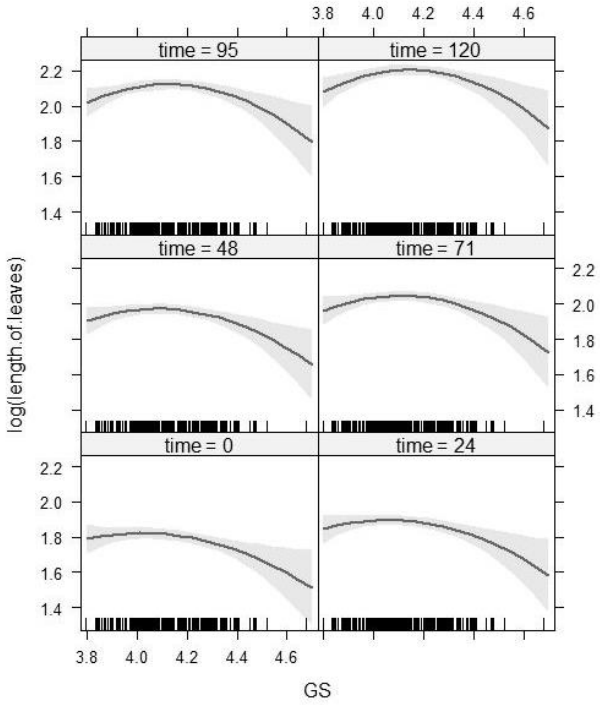


Fig 7. Length of leaves of individuals with varying genome size (2C /pg) at six time points covering the whole study period.

As regards flowering, 69 out of the 216 plants that survived (32 %) started to flower during the experiment. There were significant differences in genome size between flowering and non-flowering plants ($t = -12.4$, d.f. = 214, $p < 0.001$) (Fig. 8), with flowering plants exhibiting a smaller mean genome size (2C = 3.968 ± 0.114 pg) than non-flowering plants (2C = 4.203 ± 0.137 pg). Flowering plants also differed in the number of days they took to start flowering, plants with smaller genome sizes having an earlier onset of flowering than plants with larger genome sizes (Fig. 9).

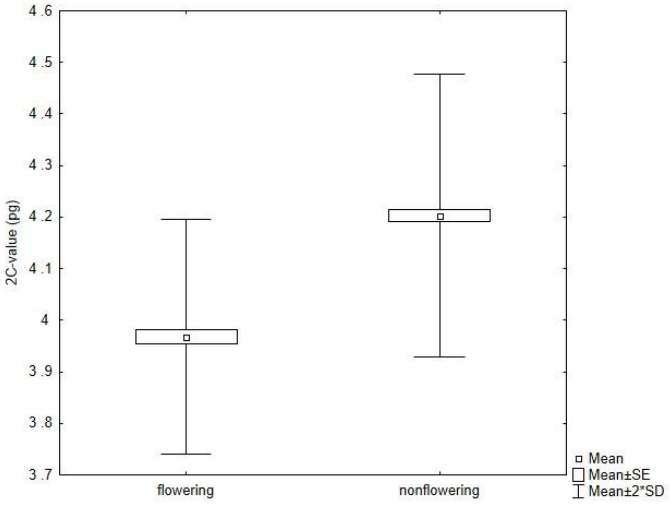


Fig 8. Differences in genome size between flowering and non-flowering plants.

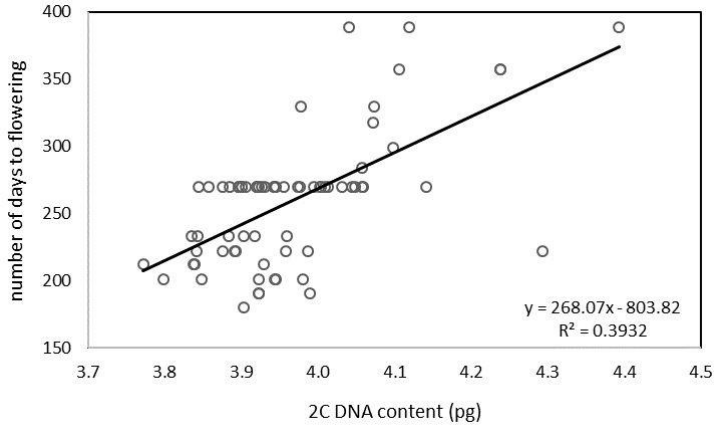


Fig 9. Number of days to the onset of flowering in plants with varying genome size, fitted with a linear regression line. *The linear regression equation is shown in the graph.*

Finally, aboveground biomass was significantly positively correlated with genome size. Polynomial regression provided better results than linear regression ($r^2 = 0.145$, $F(1,213) = 19.3$, $p < 0.001$) (Fig. 10), revealing that plants with intermediate genome size produced the most aboveground biomass. Aboveground biomass and leaf length were strongly positively correlated ($r^2 = 0.468$, $F(1,214) = 188.5$, $p < 0.001$), which indicates that leaf length can be considered a good proxy of biomass when applied to the determination of growth rates.

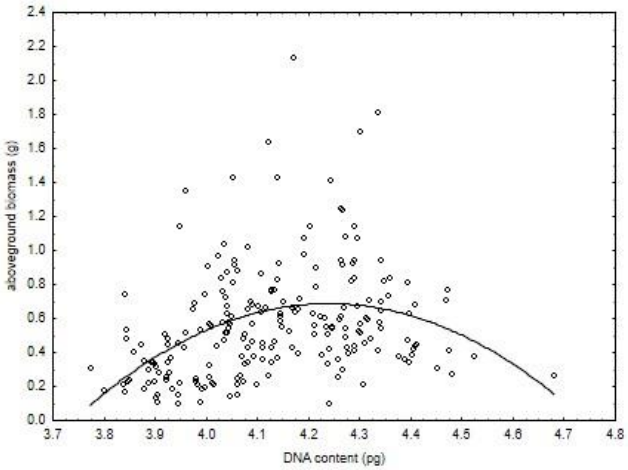


Fig 10. Scatterplot of total aboveground biomass and 2C DNA content fitted with a polynomial curve.

Discussion

Data reliability

Because of inappropriate methodological practices, numerous studies describing intraspecific genome size variation are gradually being refuted (for a review see Šmarda & Bureš 2010). Current studies attempting to prove actual variation in genome size of plant species should stringently observe certain methodological rules (Doležel and Bartoš 2005; Doležel, Greilhuber, and Suda 2007; Šmarda, Bureš, Horová, Foggi, et al. 2008; Šmarda and Bureš 2006, Walker et al. 2006, Greilhuber 2008). Based on these recommendations, and to fully assure that any observed genome size variation was real, we adhered to stringent rules for estimating genome size. In particular, we constantly used the same internal standard, preformed repeated measurements on different days until the variation for an individual was less than 2%, obtained double peaks in simultaneous analyses, and used alternative fluorescent dyes (PI and DAPI) and different flow cytometers.

Range and heritability of genome size variation

The genome size variation observed among all seedlings of *Taraxacum stenocephalum* under study (1.262-fold variation among 775 seedlings), confirmed the exceptional variation initially reported by Trávníček et al. (2013) (1.223-fold). The genome size variation among F1 seedlings was even greater than that among their parents. Similar results were attained by pollination experiments with *Microseris douglasii* (Asteraceae) reported by Price et al. (1983) or in the study of seedlings and mothers of *Festuca pallens* (Šmarda, Bureš, Horová, & Rotreklová 2008, Šmarda et al. 2010). In addition, as observed in *F. pallens*, the genome size variation within offspring families from parents with similar genome size was greater than that of parental individuals.

We observed no strict reproductive barrier between the genomes size categories of *T. stenocephalum*. Šmarda, Bureš, Horová, & Rotreklová (2008) suggested that the genome size variation in *F. pallens*, which is generated continuously, can be maintained by means of segregation of homoeologous chromosomes of different size during gametogenesis, and our results support this suggestion. However, the considerably greater proportion of viable fruits produced by mother plants with mid-sized genomes, together with the smaller dimensions of fruits produced by mother plants with very small, small and very large genomes, might indicate possible problems with the pairing of chromosomes with highly dissimilar lengths.

The larger genome size variation of F1 individuals compared to parental plants from wild populations might indicate that new variation is generated between generations; however, the prevalence of balanced selection might be responsible for keeping the species' genome size more or less constant. Nevertheless, further studies of subsequent generations and backcrosses will be necessary to fully confirm this assumption.

Karyological investigation

Chromosome counts for *Taraxacum stenocephalum* were previously reported by Kirscher et al. (1994) and Trávníček et al. (2013). Kirschner et al. (1994) investigated 13 seedlings from three mother plants collected from one Caucasian population and published karyotypes of *T. stenocephalum* showing variability in the presence of satellite chromosomes. Trávníček et al. (2013) used 16 mature plants with estimated DNA content and detected no satellite chromosomes in plants with small and large genomes. As the detection of satellites or secondary constrictions (NORs) is highly dependent on the condensation of chromosomes or of the squash, the presence of these chromosomes may have gone unnoticed in the work of Trávníček et al. (2013). In the present study, chromosome spreads obtained from 27 seedlings from crossings of plants with 1.18-fold variation in genome size contained up to three pairs of satellite chromosomes varying in size. The largest satellite fragments were only recorded in seedlings with the largest expected genome sizes, so we presume that satellite chromosomes contribute somewhat to the largest genome sizes. Similarly to the authors of previous studies on *T. stenocephalum*, we detected neither aneuploidy nor B chromosomes.

Association of genome size with growth and fitness traits

Correlations of genome size with leaf length, time to germination, time to flowering, aboveground biomass and achene weight suggest that genome size affects the phenotype and that this might have fitness consequences for *T. stenocephalum* individuals. Our results show that, in the study species, time to germination and time to flowering correspond with the nucleotypic theory (Bennett 1972), which postulates that genome size by itself can influence cell size and the duration of cell division, so it can be assumed to influence relative growth rates during early ontogeny. On the other hand, we fitted the relation between leaf length and genome size with a polynomial curve showing that individuals with mid-sized genome size exhibited the best growth rates. The non-linear distribution of genome size is quite often source of contradictory results (Knight & Beaulieu

2008) in different species. However, we did not find the number of leaves, which is usually used as a marker of seedling developmental rate (McMaster 1997, 2005, Šmarda, Bureš, Horová, & Rotreklová 2008), to be correlated with genome size. Šmarda, Bureš, Horová, & Rotreklová (2008) found positive correlations between genome size and developmental rates, which is at odds with the nucleotypic theory. These different results indicate that the relationship is more complex and that factors such as cell type and loccation (Katagiri et al. 2016), rather than merely the rate of division, may be involved.

In addition, we found a strong negative correlation between fruit weight and nuclear DNA content. In *Taraxacum*, fruit mass is directly related to seed mass. Considering that many previous studies trying to search for associations between genome size and seed mass, revealed a positive correlation (Bennett 1987; *Dąbrowska* 1992; Grotkopp et al. 2004; Knight et al. 2005; Knight and Beaulieu 2008; Kim et al. 2015), our results are not in line with the most expected trend. Seeming contradiction of relation of genome size and seed size in *Aesculus* which has relatively large genome and extremely large seeds, was found across the genus (Krahulcová et al. 2017), but using phylogenetic correction the authors found positive correlation of larger genome size and larger within individual clades. However, as has been shown by Beaulieu et al. (2007), this relationship is curvilinear and concave. Their survey of 1,222 species shows that species with small genomes have a wide range of seed mass, while large genome are associated with large seeds.

In conclusion, based on the correlations we observed, it appears that genome size variation, such as that present in *T. stenocephalum*, possibly plays an important role during the establishment of populations in new or unstable environments, where genome size might be a crucial factor affecting the probability of seedling survival.

Acknowledgements

We would like to thank Jan Kirschner and Frederick Rooks for collecting *T. stenocephalum* in the field, Jan Štěpánek for a taxonomical revision of the material, Pavel Trávníček, Jana Krejčíková, Petr Vít, Tomáš Urfus, Dora Čertnerová and Martin Čertner for their help with flow cytometric analyses, Marie Brůnová and Tomáš Černý for helping with plant cultivation, Hanna Weiss-Schneeweiss and Tae-Soo Jang for karyological analyses, Nikol Krejčová for helping with the processing of data, Petr Šmilauer and Petr Koutecký for statistical advice, and Filip Kolář for his useful comments on the manuscript.

This study was supported by long-term research development project no. RVO 67985939 and project SVV 260 439 of *the Czech Ministry of Education, Youth and Sports*.

Literature cited

Beaulieu J. M., Leitch I. J. & Knight C. A. (2007): Genome size evolution in relation to leaf strategy and metabolic rates revisited. – Ann. Bot. 99: 495–505.

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

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

Bennett M. D. (1972): Nuclear DNA content and minimum generation time in herbaceous plants. – Proc. R. Soc. London Ser. B. 181: 109–135.

Bennett M. D. (1973): Nuclear characters in plants. – Brookhaven Symposia in Biology 25: 344–366.

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

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

Cavalier-Smith T. (1985): The evolution of genome size. Wiley, Chichester.

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

Chumová Z., Krejčíková J., Mandáková T., Suda J. & Trávníček P. (2015): Evolutionary and taxonomic implications of variation in nuclear genome size: lesson from the grass genus Anthoxanthum (Poaceae). – PLoS One 10: 1–17.

Dąbrowska J. (1992): Chromosome number and DNA content in taxa of Achillea L. in relation to the distribution of the genus. – Prace Botaniczne 49: 1-83.

Devos K. M., Brown J. K. M. & Bennetzen J. L. (2002): Genome size reduction through illegitimate recombination counteracts genome expansion in Arabidopsis. – Genome Res. 12: 1075–1079.

Dodsworth S., Leitch A. R. & Leitch I. J. (2015): ScienceDirect Genome size diversity in angiosperms and its influence on gene space. – Curr. Opin. Genet. Dev. 35: 73-78.

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

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. – Ann. Bot. 82: 17-26.

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

Edwards G. A. & Endrizzi J. E. (1975): Cell size, nuclear size, and DNA content relationship in Gossypium. – Can. J. Genet. Cytol. 17: 181-186.

Greilhuber J. (2005): Intraspecific variation in genome size in angiosperms: Identifying its existence. – Ann. Bot. 95: 91–98.

Greilhuber J. (2008): Cytochemistry and C-values: The less-well-known world of nuclear DNA amounts. – Ann. Bot. 101: 791–804.

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. – Ann. Bot. 95: 255–260.

Grotkopp E., Rejmánek M., Sanderson M. J. & Rost T. L. (2004): Evolution of genome size in pines (Pinus) and its life history correlates: supertree analysis. – Evolution (N. Y). 58: 1705–1729.

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., Espuny L. D. T. & Simmons E. (2010): Stomatal vs. genome size in angiosperms: the somatic tail wagging the genomic dog? – Ann. Bot. 105: 573–584.

Katagiri Y., Hasegawa J., Fujikura U., Hoshino R., Matsunaga S. & Tsukaya H. (2016): The coordination of ploidy and cell size differs between cell layers in leaves. – Development 143: 1120–1125.

Kirschner J., Štěpánek J., Tichý L., Krahulcová A., Kirschnerová L. & Pellar L. (1994): Variation in Taraxacum bessarabicum and allied taxa of the section Piesis (Compositae): allozyme diversity, karyotypes and breeding behaviour. – Folia Geobot. Phytotaxon. 29: 61–83.

Knight C. A. & Ackerly D. D. (2002): Variation in nuclear DNA content across environmental gradients: a quantile regression analysis. – Ecol. Lett. 5: 66-76.

Knight C. A. & Beaulieu J. M. (2008): Genome size scaling through phenotype space. – Ann. Bot. 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. – J. Bot. 1–7.

Knight C. A., Molinari N. A. & Petrov D. A. (2005): The large genome constraint hypothesis: evolution, ecology and phenotype. – Ann. Bot. 95: 177–190.

Krahulcová A., Trávníček P., Krahulec F. & Rejmánek M. (2017): Small genomes and large seeds: chromosome numbers, genome size and seed mass in diploid Aesculus species (Sapindaceae). – Ann. Bot. 119: 957–964.

Leitch I. J., Bennett M. D. (2007): Genome size and its uses: the impact of flow cytometry. In: Doležel J, Greilhuber J, Suda J, eds. Flow cytometry with plant cells. Analysis of genes, chromosomes and genomes. Weinheim, Germany: Wiley-VCH Verlag, 153–176.

Leong-Škorničková J., Šída O., Jarolímová V., Sabu M., Fér T., Trávníček P. & Suda J. (2007): Chromosome numbers and genome size variation in Indian species of Curcuma (Zingiberaceae). – Ann. Bot. 100: 505–526.

Loureiro J., Trávníček P., Rauchová J., Urfus T., Vít P., Štech M., Castro S. & Suda J. (2010): The use of flow cytometry in the biosystematics, ecology and population biology of homoploid plants. – Preslia 82: 3–21.

Luo C., Chen D., Cheng X., Zhao H. & Huang C. (2017): Genome size estimations in Chrysanthemum and correlations with molecular phylogenies. – Genet. Resour. Crop Evol. 64: 1451–1463.

McMaster G. S. (1997): Phenology, development, and growth of the wheat (Triticum aestivum L.) shoot apex: a review. – Adv. Agron. 59: 63-118.

McMaster G. S. (2005): Phytomers, phyllochrons, phenology and temperate cereal development. – J. Agric. Sci. 143: 137–150.

Meagher T. R. & Vassiliadis C. (2005): Phenotypic impacts of repetitive DNA in flowering plants. – New Phytol. 168: 71–80.

Morgan H. D. & Westoby M. (2005): The relationship between nuclear DNA content and leaf strategy in seed plants. – Ann. Bot. 96: 1321–1330.

Neumann P., Koblížková A., Navrátilová A., Macas J. (2006): Significant expansion of Vicia pannonica genome size mediated by amplification of a single type of giant retroelement. – Genetics 173: 1047-105.

Obermayer R. & Greilhuber J. (2005): Does genome size in Dasypyrum villosum vary with fruit colour? – Heredity (Edinb). 95: 91–95.

Paule J., Paule L. & Gömöry D. (2018): Small genome size variation across the range of European beech (Fagus sylvatica). – Plant Syst. Evol. 304: 577–582.

Price H. J., Chambers K. L., Bachmann K. & Riggs J. (1983): Inheritance of nuclear 2C DNA content variation in intraspecifc and interspecific hybrids of Microseris (Asteraceae). – Am. J. Bot. 70: 1133–1138.

Pyšek P., Skálová H., Čuda J., Guo W. Y., Suda J., Doležal J., Kauzál O., Lambertini C., Lučanová M., Mandáková T., Moravcová L., Pyšková K., Brix H. & Meyerson L. A. (2018): Small genome separates native and invasive populations in an ecologically important cosmopolitan grass. – Ecology 99: 79–90.

SanMiguel P. & Bennetzen J. L. (1998): Evidence that a recent increase in maize genome size was caused by the massive amplication of intergene retrotranposons. – Ann Bot. 82: 37-44.

SanMiguel P., Tikhonov A., Jin Y. K., Motchoulskaia N., Zakharov D., Melake-Berhan A., Springer P. S., Edwards K. J., Lee M., Avramova Z. & Bennetzen J. L. (1996): Nested retrotransposons in the intergenic regions of the maize genome. – Science (80-.). 274: 765–768.

Shirasu K. (2000): A contiguous 66-kb barley DNA sequence provides evidence for reversible genome expansion. – Genome Res. 10: 908–915.

Schmuths H., Meister A., Horres R. & Bachmann K. (2004): Genome size variation among accessions of Arabidopsis thaliana. – Ann. Bot. 93: 317–321.

Slovák M., Vít P., Urfus T. & Suda J. (2009): Complex pattern of genome size variation in a polymorphic member of the Asteraceae. – J. Biogeogr. 36: 372–384.

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. – Ann. Bot. 100: 1323–1335.

Suda J., Meyerson L. A., Leitch I. J. & Pyšek P. (2015): The hidden side of plant invasions: the role of genome size. – New Phytol. 205: 994-1007.

Suda J., Weiss-Schneeweiss H., Tribsch A., Schneeweiss G. M., Trávníček P. & Schönswetter P. (2007): Complex distribution patterns of di-, tetra-, and hexaploid cytotypes in the European high mountain plant Senecio carniolicus (Asteraceae). – Am. J. Bot. 94: 1391–1401.

Sugiyama S., Yamaguchi K. & Yamada T. (2002): Intraspecific phenotypic variation associated with nuclear DNA content in Lolium perenne L. – Euphytica 128: 145–151.

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

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

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

Šmarda P., Bureš P., Horová L. & Rotreklová O. (2008): Intrapopulation genome size dynamics in Festuca pallens. – Ann. Bot. 102: 599–607.

Šmarda P., Horová L., Bureš P., Hralová I. & Marková M. (2010): Stabilizing selection on genome size in a population of Festuca pallens under conditions of intensive intraspecific competition. – New Phytol. 187: 1195–1204.

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

Vicient C., Suoniemi A., Anamthawat-Jónsson K., Tanskanen J., Beharav A., Nevo E. & Schulman A. (1999): Retrotransposon BARE-1 and its role in genome evolution in the genus Hordeum. – Plant Cell 11: 1769–1784.

Walker D. J., Mo I. & Correal E. (2006): Genome size in Bituminaria bituminosa (L .) C . H . Stirton (Fabaceae) populations : separation of " true " differences from environmental effects on DNA determination. – Environ. Exp. Bot. 55: 258-265.

Wendel J. F., Jackson S. A., Meyers B. C. & Wing R. A. (2016): Evolution of plant genome architecture. – Genome Biol. 17: 10.1186/s13059-016-0908-1.

Wicker T., Yahiaoui N., Guyot R., Schlagenhauf E., Liu Z., Dubcovsky J. & Keller B. (2003): Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and Am genomes of wheat. – Plant Cell 15: 1186–1197.

SUPPLEMENTARY MATERIAL

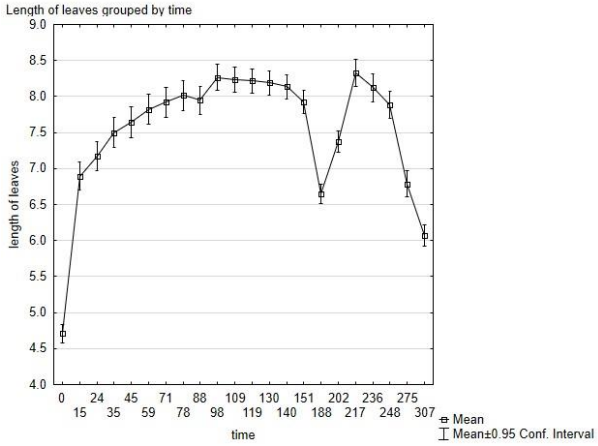


Fig S1. Leaf length during the whole period of the experiment.

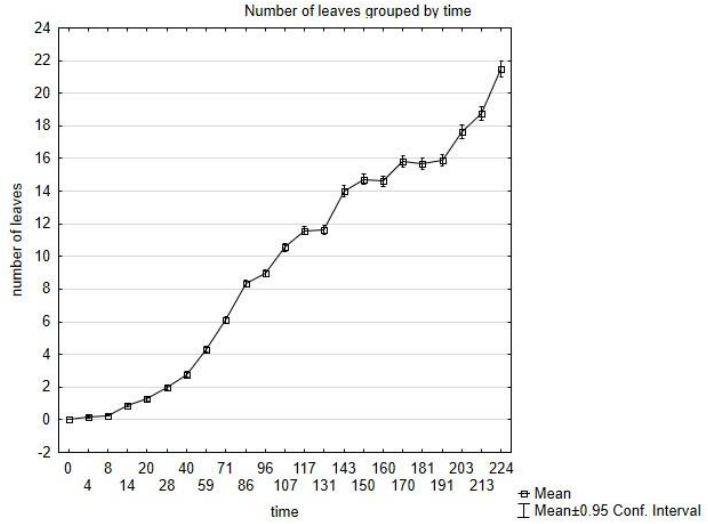


Fig. S2. Number of leaves during the whole period of the experiment.

Case study II

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**

Spatio-ecological segregation of diploid and tetraploid cytotypes of *Galium valdepilosum* in central Europe

Filip Kolář[1,2], Magdalena Lučanová[2,1], Petr Koutecký[3], Markéta Dortová[3], Adam Knotek[1] & Jan Suda[1,2]

[1] Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, CZ-128 01 Prague, Czech Republic, e-mail: filip.kolar@gmail.com;

[2] Institute of Botany, Academy of Sciences of the Czech Republic, Zámek 1, CZ-252 43 Průhonice, Czech Republic; [3]Department of Botany, Faculty of Science,

[3] University of South Bohemia, Branišovská 31, CZ-370 05 České Budějovice, Czech Republic

Abstract

The *Galium pusillum* agg. (*Rubiaceae*), with four species native to the Czech Republic, is a taxonomically challenging complex. Of these, *G. valdepilosum* is particularly interesting because this relict species shows both ploidy (the incidence of diploid and tetraploid cytotypes) and habitat differentiation (occurrence on different soil types, including serpentines). With the aid of DNA flow cytometry, analysis of vegetation samples and a hydroponic cultivation experiment we addressed the cytogeographic pattern, ecological preferences of different cytotypes both across the entire range of distribution and in the contact zone and the plant's response to serpentine edaphic stress. Ploidy distribution in *G. valdepilosum* is parapatric, with a narrow contact zone in southern Moravia. Neither triploids nor mixed 2x-4x populations were found, which together with the restriction of the species to isolated relict habitats, suggest the static character of the contact zone. In general, tetraploids occupied a wider range of habitats and colonized larger geographic areas. Diploids typically occurred in open low-competitive oak-pine forests on acidic soils while their tetraploid counterparts were also able to survive in open basiphilous grasslands with a comparatively higher competitive pressure. Serpentines did not play an important role in ecological sorting of the cytotypes. Cultivation experiments showed that *G. valdepilosum* is likely to be constitutively tolerant to serpentine chemical stress. Relative genome size and ecological data indicate that the serpentine populations from western Bohemia, traditionally referred to as *G. sudeticum*, differ from the type subalpine populations from the Krkonoše Mts and suggest their merger with *G.valdepilosum*.

Keywords: central Europe, contact zone, cytogeography, ecological sorting, flow cytometry, *Galium sudeticum*, *Galium valdepilosum*, ploidy distribution, polyploidy, serpentine

Introduction

Polyploidy, the possession of three or more complete chromosome sets per nucleus, is a prominent and recurring transition in the evolution of eukaryotic organisms, including land plants (Otto & Whitton 2000). Although polyploidization is often associated with species diversification

due to the barriers to gene flow that results from chromosome multiplication, ploidy variation is commonly observed also within taxonomic species (Husband et al. 2013). Many studies of heteroploid species note that different cytotypes have distinct distributions (Suda et al. 2007, Šafářová et al. 2011, Dančák et al. 2012, Krejčíková et al. 2013a). The pattern of ploidy distribution is shaped by the interplay between adaptive and non-adaptive ecological processes (Husband et al. 2013). The adaptive scenario assumes that polyploidy contributes to the acquisition of new genetic, morphological, physiological and/or ecological characteristics (reviewed in Levin 2002) that may modify competitive ability, fitness or ecological tolerance of polyploids compared to their diploid progenitors and ultimately lead to new responses to environmental conditions. As a consequence, different cytotypes can sort along abiotic and/or biotic environmental gradients, both contemporary and past (Husband et al. 2013). Although ecological sorting is widely acknowledged as the key mechanism driving geographic segregation of different cytotypes, several non-adaptive (i.e. environmentally independent) processes can also play a role in shaping ploidy distribution. Among others, spatial segregation of cytotypes can be governed through frequency-dependent mating success, in polyploid systems traditionally referred to as the "minority cytotype disadvantage" (Levin 1975). Present-day ploidy distribution can also reflect the dynamics of genome duplication (e.g. the frequency of unreduced gamete formation) or different dispersal abilities of the cytotypes; for example, widespread cytotypes may have been superior colonizers of habitats that appeared after the retreat of ice shields or due to human activities such as deforestation and agricultural practices (Stebbins 1985, Sonnleitner et al. 2010). However, adaptive and non-adaptive scenarios could not be distinguished on the basis of distributional patterns but the cytotypes should be subjected to a detailed evaluation of their ecological preferences and important biological traits (e.g. vegetation analyses, crossing and transplant experiments, cultivation under manipulated environmental characteristics).

Spatial relationships between cytotypes within species can be categorized as sympatric, parapatric or allopatric, depending on whether they are geographically intermixed, adjacent or disjunct, respectively. When polyploids first arise, they by necessity occur in sympatry with their diploid/lower-polyploid progenitors. Subsequent cytotype expansion or retreat will result in parapatric or allopatric distributions. Contact zones can be quite narrow, eventually comprising only a few populations, as reported in *Chamerion angustifolium* (Husband & Schemske 1998) or *Ranunculus adoneus* (Baack 2004). Cytotype mixtures extending over large areas seem to be less frequent and occur for example in *Galax urceolata* (Burton & Husband 1999), *Solidago altissima* (Halverson et al. 2008) and *Allium oleraceum* (Duchoslav et al. 2010). However, the immediate contact of different cytotypes (i.e. the incidence of mixed-ploidy populations) is often limitedeven in species with geographically extensive and diffuse contact zones, illustrative examples being *Knautia arvensis* (Kolář et al. 2009), *Vicia cracca* (Trávníček et al. 2010), *Aster amellus* (Castro et al. 2012) or *Odontites vernus* (Koutecký et al. 2012). While most contact zones are formed by two ploidy levels, the last years have seen much more complex population structures, with up to five different co-existing cytotypes (Sonnleitner et al. 2010, Trávníček et al. 2011b, 2012). Investigations into the adaptive significance of ploidy shift first require assessment of potential relationship between intraspecific ploidy variation and environmental factors of occupied sites. Detected associations of ploidy levels with both abiotic (Duchoslav et al. 2010, Sonnleitner et al. 2010, Manzaneda et al. 2012) and biotic (Krejčíková et al. 2013b) parameters provide important clues for explaining the observed cytogeographic patterns.

Published studies addressing cytogeographic patterns and underlying mechanisms in heteroploid species in central Europe usually deal with species of semi-ruderal habitats (*Allium oleraceum*: Duchoslav et al. 2010, Šafářová & Duchoslav 2010, Šafářová et al. 2011; *Knautia arvensis*: Kolář et al. 2009; *Pilosella officinarum*: Mráz et al. 2008; *Spergularia echinosperma*: Kúr et al. 2012; *Vicia cracca*: Trávníček et al. 2010) or nonrelict natural sites (*Aster amellus*: Mandáková & Münzbergová 2006, Castro et al. 2012; *Molinia caerulea* agg.: Dančák et al. 2012), whereas species restricted to isolated relict sites, i.e. low-competition habitats with species assemblages usually persisting from the early Holocene, have been largely neglected (but see Suda & Lysák 2001, Suda et al. 2004). Due to their supposed closer association with local environmental conditions, insular-like distribution and long periods of isolation of individual populations, relict species with multiple cytotypes provide novel insights into the structure and dynamics of contact zones between different cytotypes.

A suitable candidate for such an investigation is *Galium valdepilosum* H. Braun (*Rubiaceae*), a diploid-tetraploid member of the *G. pusillum* aggregate (Ehrendorfer 1960, Ehrendorfer et al. 1976). This group, which in central European literature is sometimes treated in a narrower sense as *G. pumilum* aggregate, encompasses four native species in the Czech Republic (Krahulcová & Štěpánková 1998, Štěpánková 2000, Danihelka et al. 2012): (i) widespread octoploid (2n = 8x = 88) *G. pumilum* Murray, (ii) tetraploid (2n = 4x = 44) *G. austriacum* Jacq. restricted to limestone outcrops in Pavlovské vrchy in southern Moravia, (iii) endemic tetraploid *G. sudeticum* Tausch, which has a very unusual distribution pattern, being reported from basiphilous subalpine areas (glacial cirques) in the Krkonoše Mts (historically also from the Hrubý Jeseník Mts) and from comparatively low-lying serpentine outcrops in the Slavkovský les Mts (western Bohemia), and (iv) ploidy-variable *G. valdepilosum*, which includes diploid (2n = 2x = 22) and tetraploid (2n = 4x = 44) populations inhabiting different relict sites (dry grasslands, open forests) on both serpentine and non-serpentine soils. A previous study of the aggregate using conventional chromosome counts (Krahulcová & Štěpánková 1998) provided a rough picture of ploidy distribution in the Czech Republic and its close surroundings and concluded that ploidy variation is not associated with serpentine vs non-serpentine sites. The origin of the tetraploid cytotype (auto- vs allopolyploid) is unclear. Although overall morphological similarities(but with certain quantitative differentiating traits;Štěpánková 2000) and close monoploid genome sizes of both cytotypes (Kolář et al. 2013) would favour autopolyploidy, reticulate patterns of morphological characters, high plasticity and great taxonomic complexity of the whole *G. pusillum* group indicate the need for a multi-species molecular investigation.

The present study builds on our previous research on the *G. pusillum* agg. in deglaciated areas of northern Europe (Kolář et al. 2013) and the karyological investigations in eastern central Europe of Krahulcová & Štěpánková (1998). Using DNA flow cytometry, analysis of habitat preferences and a hydroponic cultivation experiment we addressed the following questions:(i) Whatare the ranges of diploid and tetraploid *G. valdepilosum* and where is the contact zone between these cytotypes located? (ii) Do both cytotypes co-occur in ploidymixed populations? (iii) Do diploid and tetraploid cytotypes differ in their habitat preferences both across the entire range of distribution and in the zone of ploidy contact? (iv) Are there any ploidy-specificdifferences in growth response of *G. valdepilosum* to serpentinechemicalstress? (v) What is the variation in nuclear DNA content within the tetraploid *G. valdepilosum*? Do taxonomically uncertain serpentine populations in western Bohemia, traditionally referred to as *G. sudeticum*, share genome size values with plants of *G. sudeticum* from subalpine type populations or with *G. valdepilosum*?

Materials and methods

Field sampling

Plant material was collected from 2009 to 2013 in Austria (12 sites), the Czech Republic (70 sites), Germany (13 sites) and Poland (nine sites). We covered the entire range of *G. valdepilosum* except for populations in central Denmark that are referred to as an endemic subsp. *slesvicense*(Sterner ex Hylander) Ehrendorfer. In addition to the nominate subspecies of *G. valdepilosum* (94 populations), we also included for comparative purposes four serpentine populations from western Bohemia [traditionally determined as *G. sudeticum*, but showing some morphological differences from typical subalpine populations (Štěpánková 2000), which are ecologically close to *G. valdepilosum*], five highaltitude populations of *G. sudeticum* from the Krkonoše Mts and one taxonomically uncertain population from limestone outcrops in the Králický Sněžník Mts (further referred to as *G. pusillum* agg.; see Appendix 1 for details of individual localities). Whenever possible with respect to population size, shoots from at least 10 plants per population were collected and stored in plastic bags in cold conditions until used in the FCM analysis. To avoid collecting the same genet, the distance between the individuals sampled was at least 0.5 m. Herbarium vouchers are deposited in the Herbarium of Charles University in Prague (PRC).

Floristic composition and selected environmental conditions recorded at 52 localities were characterized using vegetation samples (phytosociological relevés), including those of 46 localities of *G. valdepilosum* (covering the entire range of distribution: 7 and 15 diploid-inhabited sites in Lower Austria and Moravia, respectively, and 10, 1, 2, 6, and 5 tetraploid-inhabited sites in Bavaria, Lower Austria, Bohemia, Moravia and Poland, respectively), two serpentine localities of putative *G. sudeticum*, three subalpine localities of *G. sudeticum* and one locality of a taxonomically uncertain member of the *G. pusillum* agg. One vegetation sample per locality was usually recorded, exceptions being three ecologically diverse sites where two samples from distinct vegetation units were recorded; each sample covered an area of 3 × 3 m in areas with an abundance of *Galium* plants (Electronic Appendix 3). In each plot, relative cover of all vascular plant species was quantified using a modified nine-point Braun-Blanquet scale (Braun-Blanquet 1964) and the following environmental parameters were recorded: total vegetation cover, cover of each vegetation layer, slope inclination and orientation, and proportion of bare rock. At 49 localities (Electronic Appendix 5), mixed rhizosphere soil samples were collected at five microsites within the area of the vegetation sample; pH and concentrations of selected elements (C, N, K, Ca, and Mg) were determined in the Analytical Laboratory of the Institute of Botany, Průhonice, CZ (see Kolář et al. 2013 for methodology details).

Flow cytometry

Relative fluorescence intensities of isolated nuclei were estimated using DNA flow cytometry (FCM) following the simplified two-step protocol with DAPI staining and *Bellis perennis* as internal reference standard as detailed in Kolář et al. (2013). In six selected populations (Appendix 1), one individual per population was subjected to more stringent analysis of relative DNA content (following Kolář et al. 2013). For comparative purposes DNA content values of another 17 individuals (from 17 populations) were taken from Kolář et al. (2013). *Galium* accessions with distinct fluorescence intensities were analysed simultaneously in order to confirm between-plant

differences observed in runs with an internal standard. Chromosome-counted individuals (Kolář et al. 2013) were used as a reference for the interpretation of FCM histograms.

Hydroponic cultivation

Eight populations were subjected to a hydroponic cultivation experiment aimed at assessing the effects of the major chemical factors associated with serpentine conditions (i.e. low Ca/Mg ratio and high Ni concentrations; Brady et al. 2005, Kazakou et al. 2008) on seedling performance. Due to the acidic pH of *G. valdepilosum*-inhabited serpentine stands (mean pH of 5.5) their responses were compared with those of four acidophilous non-serpentine populations. Two diploid and two tetraploid populations were represented in each group (Fig. 1; see Appendix 1 for details). Mature achenes collected along transects at the original sites were germinated on moist filter paper over a period of three weeks. Vital, undamaged seedlings were then carefully fixed to a floating plastic disc (14 cm in diameter) so that there was an equal distance between each of the experimental plants. There were eight plants (one per population) on each disc, which was placed in a 1-L lightimpermeable container filledwith a standard nutrient solution as described in Huss-Danell (1978), with a slight modification: $Co(NO_3)_2$ was used instead of $CoSO_4$ as a cobalt source. The seedlings were grown in this nutrient solution for 11 days prior to the start of the experiment. They were then placed into experimental solutions with manipulated concentrations of Mg^{2+} and Ni^{2+} for the next 22 days ($MgSO_4$ and $NiSO_4$ were used as sources of Mg and Ni, respectively; the pH was approx. 7 during the whole experiment). The solutions were replaced every three days with freshly prepared solution and the plants cultivated in a controlled-environment growth cabinet at the Faculty of Science, University of South Bohemia, Czech Republic (for details see Kolář et al. 2014).

To test the individual and combined effects of Ni and Mg on *G. valdepilosum* populations differing in soil type (factor 'substrate at origin') and ploidy level (factor 'ploidy'), we used a mixed-effect full-factorial experimental design. Four experimental treatments were applied: the control (standard nutrient solution), high Ni^{2+}, high Mg^{2+}, and high Ni^{2+} and Mg^{2+}. Based on a preliminary cultivation experiment, the concentrations of Ni^{2+} were set to 0 μM (control) and 30 μM, while the concentrations of Mg^{2+} were set to 0.55 mM (control) and 5.5 mM (i.e. Ca/Mg ratio of 2 and 0.2, respectively). Each experimental unit (= plastic container filled with one of the four experimental solutions) consisted of eight seedlings, one seedling per population. There were eight replicates of each treatment, resulting in 32 experimental units and 256 seedlings. Total root length was used as a proxy of the plant's response to different experimental treatments; the values were obtained from measurements recorded at the beginning and the end of the experiment (following the method described in Kolář et al. 2014).

Statistical analyses

Differences in relative DNA contents were tested in R version 2.15.2 using one-way ANOVA with post-hoc comparisons (Tukey HSD test).

Habitat preferences were based on the species composition of vegetation samples and recorded biotic and abiotic characteristics of the sites. Ellenberg indicator values (EIV), which

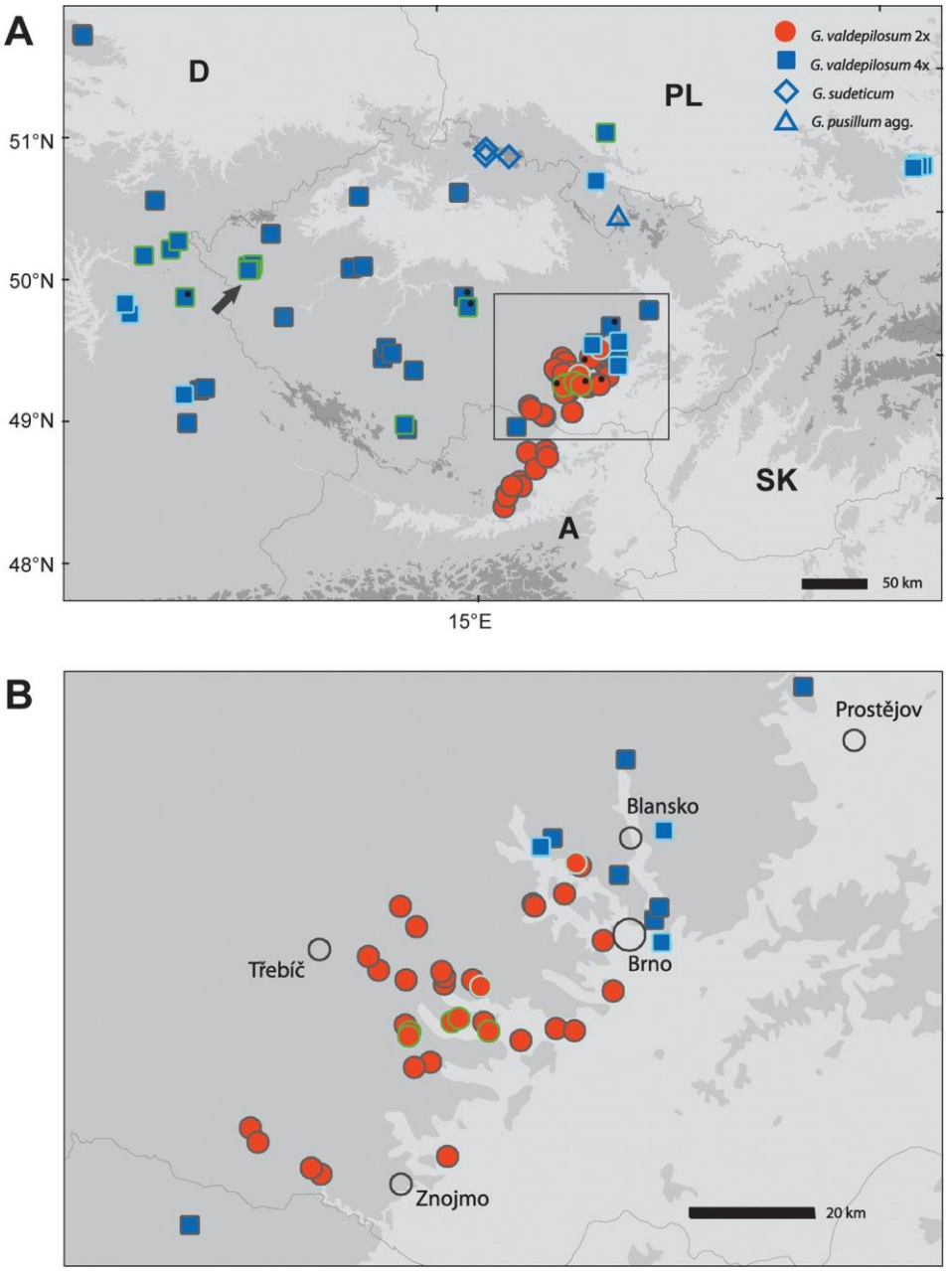


Fig. 1. – Geographic location of populations of *Galium valdepilosum* across the entire study area (A) and in the contact zone in southwestern Moravia (B). Red and blue denote diploids and tetraploids, respectively. Black, light blue and green borders indicate acid, basic and serpentine soils, respectively. The arrow indicates the location of taxonomically reclassified serpentine populations from western Bohemia traditionally referred to as *G. sudeticum*.
Populations marked by a black dot were cultivated hydroponically.

provide estimates of environmental characteristics inferred from species composition data (Ellenberg 1992), were calculated in JUICE 7.0 (Tichý 2002) based on presence/absence data of herbaceous species of plants. Separate analyses were done for (i) all available vegetation samples of *G. valdepilosum* covering the entire distribution range of the species, and (ii) a subset of vegetation samples from the contact zone between di- and tetraploid cytotypes (i.e. within a radius of 50 km around the town of Brno where immediate contact of both ploidies was recorded). Both unconstrained (using the detrended correspondence analysis, DCA) and constrained (using the canonical correspondence analysis, CCA, with forward selection of environmental variables) ordinations in Canoco for Windows, ver. 4.5 (Lepš & Šmilauer 2003) were used to describe the overall vegetation patterns of the *G. valdepilosum* sites studied. Differences in vegetation composition among vegetation samples recorded at sitesof diploid vs tetraploid *G. valdepilosum* were testedin a separate CCA with 'ploidy level' as the only predictor variable. In order to reveal associations of di- vs tetraploid *G. valdepilosum* plants with other plant species, ten co-occurring species with the strongest marginal effects were analysed using the Monte Carlo permutation test (999 permutations, with Bonferroni correction for multiple tests) during the forward-selection linear discriminant analysis in which species abundances (log-transformed) were treated as predictor variables and *Galium* ploidy level as a response (see Lepš & Šmilauer 2003 for details). The biotic characteristics inferred from species composition data (i.e. EIV, species diversity, layer cover) were omitted as predictors in constrained analyses.

Differences in root growth (log-transformed) of *G. valdepilosum* seedlings in response to high concentrations of Mg^{2+} and Ni^{2+} were tested using a hierarchical ANOVA. The effects of substrate at origin, ploidy, Mg and Ni treatments, and all their interactions were tested using a linear model where the experimental container (nested in Mg and Ni treatment interaction) and population of origin (nested in substrate at origin and ploidy interaction) were treated as random and fixed factors, respectively. For comparative purposes, we also performed an analysis aimed at identification of the overall differences in serpentine tolerance among *G. valdepilosum* populations differing in ploidy / soil conditions. A similar ANOVA model was used for this purpose, but with the population of origin (again nested in substrate at origin and ploidy interaction) treated as a factor with random effect. The ANOVA analyses were calculated in Statistica 8 (StatSoft 2007). Note that Statistica uses Satterthwaite's method of denominator synthesis, which finds linear combinations of sources of random variation that serve as appropriate error terms for testing the significance of the respective effect of interest;for this reason the synthesized error mean squares and synthesized error degrees of freedom are also presented.

Results

Cytogeography and variation in relative nuclear DNA content

The FCM analysis of 874 plant samples revealed two different DNA ploidy levels: diploid (338 individuals from 46 localities) and tetraploid (536 individuals from 58 localities). All diploids corresponded to *G. valdepilosum* and were restricted to southern Moravia and Lower Austria. The zone of contact between the plants of the two ploidy levels is located near the town of Brno, where tetraploids in the north-east give way to diploids in the south-west (Fig. 1). Only the tetraploid cytotype of *G. valdepilosum* was recorded in Bohemia, Germany and Poland. One tetraploid population occurred in northern Austria in an area otherwise dominated by diploids. Subalpine

populations of *G. sudeticum* in the Krkonoše Mts were uniformly tetraploid as also were serpentine populations in western Bohemia and a taxonomically-uncertain population on the Polish side of the Králický Sněžník Mts.

While fluorescence intensities of all diploid samples were uniform, there was significant variation in the relative amounts of nuclear DNA ($F_{3,25} = 23.15$, $P < 0.001$) at the tetraploid level. Two groups were identified. The first group encompassed all populations determined as *G. valdepilosum*, four serpentine populations in western Bohemia traditionally referred to as *G. sudeticum* and one calcicolous mountain population in the Králický Sněžník Mts (Fig. 2). The second group with higher fluorescence intensities (mean difference 4.3%) consisted of subalpine populations of *G. sudeticum* in the Krkonoše Mts. Simultaneous FCM analysis (Fig. 3) confirmed the differences in the relative DNA contents of individuals of the putative *G. sudeticum* that originated from the two disjunct geographic areas (western Bohemia and the Krkonoše Mts).

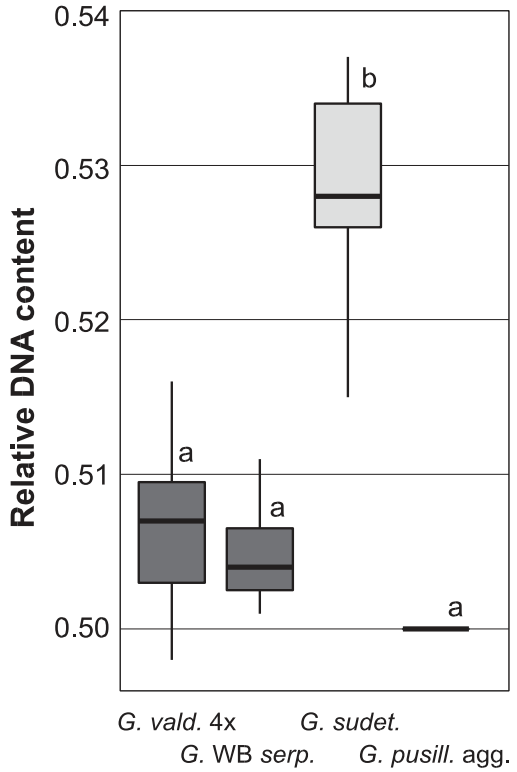


Fig. 2. – Variation in relative nuclear DNA content of *Galium valdepilosum* (23 individuals from 23 populations across the entire range of distribution), *G. sudeticum* from the Krkonoše Mts (four populations), plants inhabiting serpentine sites in western Bohemia traditionally referred to as *G.sudeticum* (four populations) and one taxonomically uncertain *G. pusillum* agg. population from the Králický Sněžník Mts. Fluorescence intensity of *Bellis perennis* was set to a unit value. Each plant was measured three times on different days. Letters indicate significantly different groups at = 0.05. The values represented by lines, boxes and whiskers are median, quartiles and range (min-max), respectively.

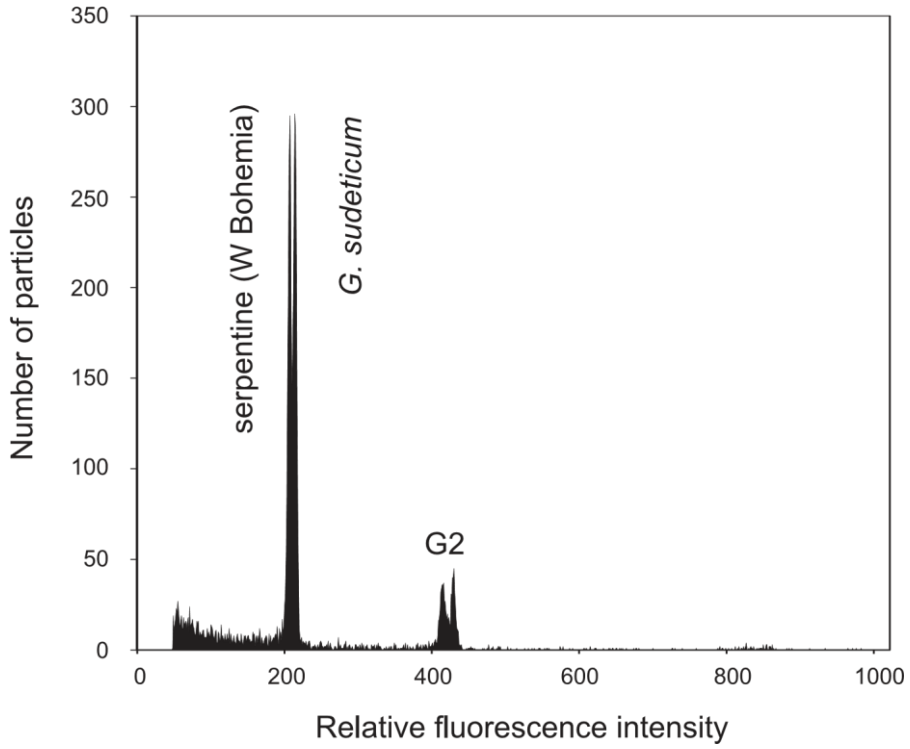


Fig. 3. – Flow cytometric histogram documenting 3.8% divergence in relative nuclear DNA content among simultaneously processed and DAPI-stained accessions of *Galium sudeticum* from the Krkonoše Mts (pop. G172) and plants from serpentine outcrops in western Bohemia traditionally referred to as *G. sudeticum* (pop. G032).

Ecological preferences of different cytotypes

Subalpine populations of *G. sudeticum* in the Krkonoše Mts and the taxonomically uncertain population in the Králický Sněžník Mts are ecologically very distinct from all other populations of *G. valdepilosum* analysed as wellas from populations inhabiting serpentine sites in western Bohemia traditionally referred to as *G. sudeticum* (Electronic Appendix 1) and therefore omitted from the following statistical analyses. In contrast, the western Bohemian populations do not ecologically differ from those of *G. valdepilosum* and both groups were therefore merged and included in subsequent analyses.

Floristic composition of sites inhabited by *G. valdepilosum* is primarily shaped by soil pH, concentration of Ca, organic C content and serpentine-specific Ca/Mg ratio (Monte Carlo test, $P = 0.001$). At these sites five other environmental parameters (concentration of Mg, cover of rocks, tree/shrub and moss layers, and altitude) were marginally significant (i.e. $P < 0.05$ yet not passing the significance level defined by Bonferroni correction).

Sites of di- and tetraploid cytotypes significantly differed in floristic composition both across the entire range of their distribution and in the contact zone (Monte Carlo test, both $P = 0.001$). Despite this differentiation, linear discriminant analysis revealed only a few species that were significantly associated with a particular cytotype of *G. valdepilosum*. *Arrhenatherum elatius*,

Genista pilosa and *Pimpinella saxifraga* were associated with diploids while juvenile *Rubus idaeus* and *Galium album* with tetraploids (in vegetation samples from the entire range and contact zone, respectively).

Diploids of *G. valdepilosum* mostly occurred in open forests on nutrient-poor acidic or serpentine soils and, in general, had a narrower ecological niche than their tetraploid counterparts (Fig. 4). Tetraploids were ecologically more divergent and occupied two major types of habitats across their entire distribution: (i) acidic or serpentine sites and (ii) baserich non-forested sites such as relatively species-rich rocky/continental grassland (see also Table 1). Although both di- and tetraploids grow on serpentine soils the environmental conditions where tetraploids grow differ. Ecological segregation of both cytotypes was more pronounced in the zone where they come into contact (Fig. 4). While diploids usually occurred in acidophilous open forests (including serpentine sites), tetraploids preferred lime-rich stands with a dense herbaceous cover.

Response to serpentine chemical stress

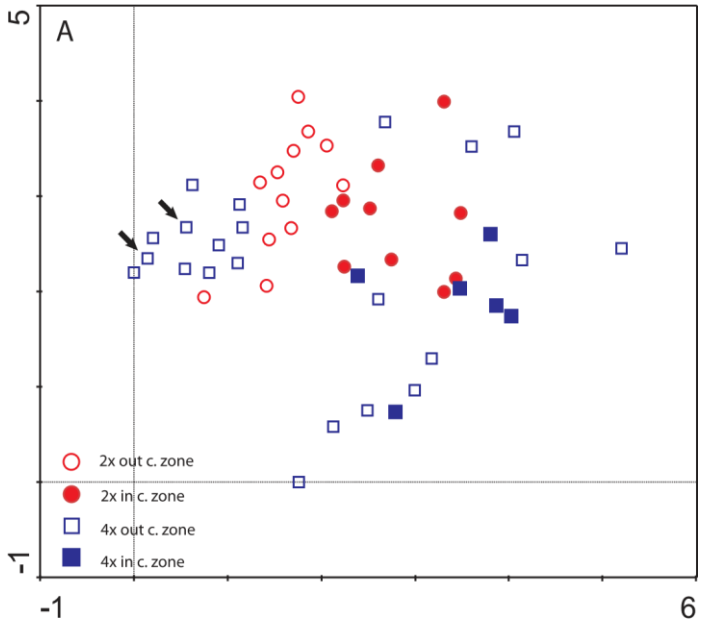
At high concentrations of Mg the roots of seedlings of *G. valdepilosum* grew significantly less, whereas the effect of high Ni was obvious only in its interaction with Mg (slightly better growth at a high Mg + Ni concentration than at a high concentration of Mg; Table 2). In general, *Galium* plants of serpentine vs non-serpentine origin and of different ploidy levels responded to Mg and Ni stress in a similar way (Table 2). The root growth of the two serpentine tetraploid populations was better than that of both their diploid and non-serpentine counterparts, irrespective of the actual concentrations of Mg and/or Ni in the solution (Fig. 5; see also Electronic Appendix 2 for response of individual populations). However, the effects of ploidy level ($F_{1,207}$ = 2.34, P = 0.20) and substrate at origin ($F_{1,207}$ = 6.83, P = 0.06) were not significant in the ANOVA model with population treated as a randomeffect factor, which makes generalizing about this difference tenuous.

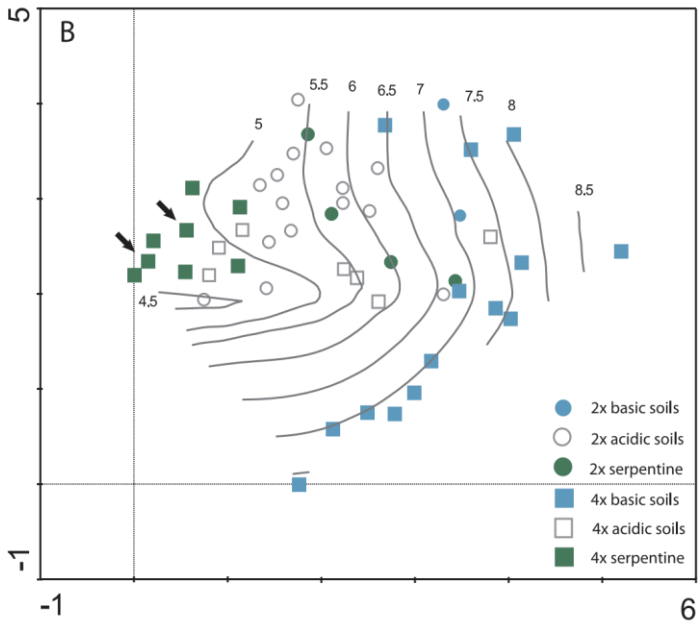
Discussion

This study increased our understanding of the karyological and ecological differentiation of the *G. pusillum* agg. in central Europe, particularly that of *G. valdepilosum*, which is a declining species restricted to various relict habitats, whose centre of distribution is in the Czech Republic. In addition to providing adetailed pictureof thedistributions of individuals with different ploidy levels at various spatial scales, the data also provides the first evidence that the taxonomic relationships of some populations may need to be reassessed.

Fig. 4. – Habitat preferences of di- and tetraploid cytotypes of *Galiumvaldepilosum*. The patterns in floristic composition of 50 vegetation samples are visualized using detrended correspondence analysis (the first and second ordination axes explain 5.4% and 3.8% of the total variation, respectively). (A) Diploid (red) and tetraploid (blue) localities within the contact zone (filled symbols) and beyond (empty symbols). (B) Vegetation samples labelled according to the major soil type (base-rich: blue, acidic: white, and serpentine: green) as determined by geological bedrock, soil pH and Ca/Mg ratio (diploid: circle, tetraploid: square). The contour lines depict pH values modelled by loess smoother from the measured values of individual vegetation samples. (C) Environmental variables significantly (red lines) and marginally significantly (blue lines) influencing floristic composition of *Galium* sites, and variables inferred from species composition data (black

lines) passively projected on the plot. Serpentine populations from western Bohemia traditionally referred to as *G. sudeticum* are marked by an arrow.


A
5
-1
-1
6
2x out c. zone
2x in c. zone
4x out c. zone
4x in c. zone


B
5
-1
-1
6
5
5.5
6
6.5
7
7.5
8
8.5
4.5
2x basic soils
2x acidic soils
2x serpentine
4x basic soils
4x acidic soils
4x serpentine

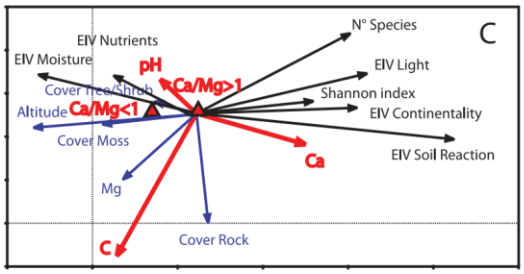

C
EIV Nutrients
EIV Moisture
pH
Ca/Mg>1
Cover Tree/Shrub
Altitude
Ca/Mg<1
Cover Moss
Mg
C
Cover Rock
Ca
N° Species
EIV Light
Shannon index
EIV Continentality
EIV Soil Reaction

Table 1. – Ploidy levels, relative amounts of nuclear DNA and habitat preferences of the taxa/populations of *Galium pussilum* agg. investigated. Relative DNA content is given as mean±SD; fluorescence intensity of *Bellis perennis* (internal reference standard) was set to a unit value. Nomenclature of vegetation units follows Chytrý et al. (2007, 2013).

Taxon/population	No. of popu-lations	Ploidy level	Ralative DNA content*	Geological substrate	Associated vegetation
G. valdepilosum H. Braun	46	2x	0.259±0.006	various silicate rocks, serpentine, rarely basic conglomerate and limestone	*Quercion petraeae, Quercion roboris, Dicrano-Pinion sylvestris, rarely Erico carneae-Pinion* (on serpentines)
	48	4x	0.506±0.005	various silicate rocks, serpentine, limestone, rarely chalk (Poland) and vulcanite	*Quercion roboris, Quercion petraeae, Dicrano-Pinion sylvestris, rarely Diantho lumnitzeri-Seslerion, Cirsio-Brachypodion pinnati* (in Poland), *Erico carneae-Pinion* (in Bavaria)
G. pusillum agg. from serpentines in western Bohemia traditionally referred to as *G. sudeticum*	4	4x	0.505±0.004	serpentine	*Dicrano-Pinion*
G. sudeticum Tausch	5	4x	0.528±0.008	base-rich substrates in glacial cirques (erlan, carbonate)	*Agrostion alpinae*
G. pusillum agg. from the Králický Sněžník Mts	1	4x	0.500	limestone	cf. *Tilio platyphyli-Acerion*

Table 2. – The effects of different concentrations of Mg and Ni, ploidy level and soil from which the plants originated (serpentine vs non-serpentine) on the total root length of *Galiumvaldepilosum* plants in hydroponic cultivation. Statistically significant results are in bold: *P < 0.05, ***P < 0.001. Dependent variables were log transformed prior to the analysis.

Factor/Interaction	Effect	Effect df	Synthesized error df	MS	Synthesized error MS	F
Experimental container	random	28	207	0.191	0.108	**1.77***
Population	fixed	4	207	0.635	0.108	**5.89*****
Mg	fixed	1	28	1.303	0.191	**6.82***
Ni	fixed	1	28	0.065	0.191	0.34
Ploidy	fixed	1	207	1.489	0.108	**13.80*****
Substrate at origin	fixed	1	207	4.339	0.108	**40.22*****
Mg × Ni	fixed	1	28	1.199	0.191	**6.27***
Ploidy × Mg	fixed	1	207	0.003	0.108	0.03
Ploidy × Ni	fixed	1	207	0.005	0.108	0.05
Substrate at origin × Mg	fixed	1	207	0.044	0.108	0.4
Substrate at origin × Ni	fixed	1	207	0.039	0.108	0.36
Ploidy × Substrate at origin	fixed	1	207	2.054	0.108	**19.04*****
Ploidy × Mg × Ni	fixed	1	207	0.111	0.108	1.03
Substrate at origin × Mg × Ni	fixed	1	207	0.338	0.108	3.13
Ploidy × Substrate at origin × Mg	fixed	1	207	0.025	0.108	0.23
Ploidy × Substrate at origin × Ni	fixed	1	207	0.022	0.108	0.2
Ploidy × Substrate at origin × Mg × Ni	fixed	1	207	0.014	0.108	0.13
Error		207		0.108		

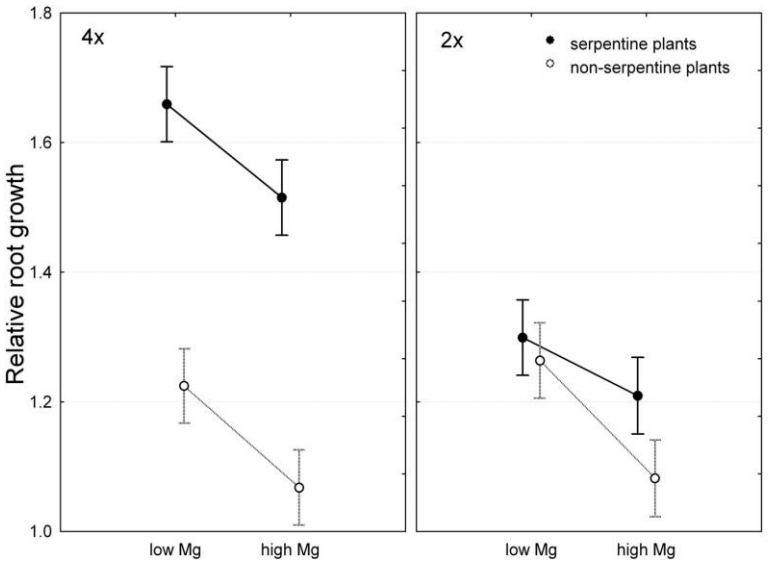


Fig. 5. – Differences recorded in the growth of the root system of diploid and tetraploid seedlings of *Galium valdepilosum* originating from serpentine vs non-serpentine soil when grown in low and high concentrations of Mg. Symbols and vertical bars denote unweighted means and standard errors of the mean, respectively.

Cytogeography of Galium valdepilosum and the underlying mechanisms

The overall cytogeographic pattern inferred from the FCM analysis of nearly 100 populations spread across the entire distribution of *G.valdepilosum* corresponds well with the incidence of different ploidy levels based on the conventional karyological counts of Krendl (1993) and Krahulcová & Štěpánková (1998). This is slightly different from that in the review of Ehrendorfer (1962), partly because he includes chromosomal data of M. Piotrowicz (published in Skalińska et al. 1961), which includes diploid populations from Małopolska upland in southern Poland. In contrast, our thorough investigation of the same geographic area (including searches for all the populations reported by M. Piotrowicz) revealed either only tetraploid individuals or failed to confirm the occurrence of the species. We can only speculate about the reasons for this discrepancy, which include species misidentification, incorrect chromosome counting (other chromosome counts from that area detected only tetraploids; Kucowa & Mądalski 1964) or even extinction of diploid cytotypes in situ (the species seems to be strongly declining particularly at localities with xerothermous grassland; see also Zarzycki & Kaźmierczakowa 2001 and Grulich 2012). The map in Ehrendorfer (1962) also shows a few diploid populations in central Bohemia. However, these records cannot be verified and should be treated with caution because neither exact localities nor references are provided in the original work.

Ploidy distribution in *G. valdepilosum* can best be described as parapatric, i.e. with closely adjacent but not overlapping ranges. Despite intensive sampling in the contact zone (the majority of *Galium* tufts was checked for ploidy in large populations while all individuals were examined in smallpopulations), we did not find any mixed 2x-4x populations or a minority cytotype such as a triploid. This suggests very low rates of neopolyploid formation and/or establishment, leaving very little room for inter-ploidy interactions. Consequently, the contact zone seems to be a non-dynamic system, which contrasts with many other recently investigated intraspecific heteroploid systems in central Europe that frequently comprised cytotype-mixed populations and odd ploidies (e.g. *Allium oleraceum*: Duchoslav et al. 2010; *Gymnadenia conopsea*: Trávníček et al. 2011b; and *Hieracium echioides*: Trávníček et al. 2011a). The static character of the contact zone is further underlined by the overall species' preferences for open relict stands, in which populations of such heliophilous and competitively weak plants are spatially isolated, possibly for many generations (in extreme cases since the spread of closed forests in the middle Holocene; Ložek 1973, Lang 1994). Geographic segregation of different cytotypes is widely considered to be the most important prezygotic reproductive barrier, with many examples described in the literature (see Husband & Sabara 2004, Kron et al. 2007, Šafářová & Duchoslav 2010, Husband et al. 2013).

The analysis of environmental conditions recorded at the localities showed that, despite being restricted to relict habitats, *G. valdepilosum* can grow in a wide range of different soils (including acidic, basic and serpentine soils; Electronic Appendix 5) and different types of vegetation (floristic composition of which is also largely determined by soil parameters). Although we found no evidence for strong inter-ploidy niche divergence (either across the entirerangeofthespecies'distributionorinthecontactzone),someecologicaltrendscanbe discerned. In particular, while diploids typically occurred in open low-competitive oak-pine forests on acidic soils, their tetraploid counterparts were also able to survive in open basiphilous grasslands with comparatively high competitive pressure. In general, tetraploids occupied a wider range of habitats and also colonized larger geographic areas.

Serpentines do not play an important role in inter-ploidy niche segregation and serpentine/non-serpentine differentiation merely reflects colonization history (i.e. diploids occur on serpentines in 2x-dominated areas and vice versa). Serpentine and non-serpentine *G. valdepilosum* populations also do not differ morphologically (Štěpánková 1997). In addition, the results of our cultivation experiment (populations responded in a similar way irrespective of the type of soil they normally grow in) indicate that response to serpentine chemical stress seems to be a constitutive trait common for both serpentine and non-serpentine diploid and tetraploid populations of *G. valdepilosum*. Such constitutive tolerance to serpentine stress implies that the species appears to be somehow "preadapted" to the principal chemical challenges of serpentine substrates such as low Ca/Mg ratio and high Ni content. Our hypothesis of serpentine "preadaption" of *G. valdepilosum* is supported by the high number of spatially isolated serpentine localities (almost all large areas of serpentine on the Hercynian massif) inhabited by the species, which most likely were independently colonized from nearby nonserpentine areas.The absence of local adaptation to high heavy metal toxicity is documented for several plant complexes, including *Silene dioica* (Westerbergh1994), *Thlaspi goesingense* (Reeves & Baker 1984) and *Th. montanum* (Boyd & Martens 1998). Moreover, even plants that do not grow on serpentines can tolerate extremely low Ca/Mg ratios, such as *Phacelia dubia* var. *georgiana*, which is restricted to dry and nutrient poor granite outcrops (Taylor & Levy 2002), i.e. similar areas to those inhabited by *G. valdepilosum*. In summary, serpentine sites seem to have served as an easily colonized refugium for *G. valdepilosum*, but had no influence on the ecological sorting of its cytotypes. This is in marked contrast with another thoroughly investigated central European di-tetraploid complex, *Knautia arvensis*, which includes a distinct serpentine-tolerant genetic lineage comprising diploid and local autotetraploid populations (Kolář et al. 2012, 2014).

Taxonomic implications

The taxonomy of the *G. pusillum* species complex in Europe is challenging due to the high number of phenotypically similar taxa and small differences in the diagnostic characters, mainly in their fruit (Ehrendorfer et al. 1976). Misidentifications are common and literature records not accompanied by herbarium vouchers are likely to be unreliable (Štěpánková 2000).

Galium sudeticum described from the Krkonoše Mts (Tausch 1835) is traditionally reported from two other geographic areas in the Czech Republic (Ehrendorfer et al. 1976, Štěpánková 2000): (i) the glacial cirque Velká Kotlina in the Hrubý Jeseník Mts (not recently rediscovered despite repeated intensive searches, including our own), and (ii) serpentine outcrops in the Slavkovský les in western Bohemia (first referred to as *G. sudeticum* by Ehrendorfer 1956). Its peculiar distribution (high-altitude habitats in the Sudeten Mts vs comparatively lower-lying, more than 200 km distant serpentine sites) has been long noted and considered comparable to some other arcto-alpine species that occur in isolated serpentine areas (Krahulcová & Štěpánková 1998). Nevertheless, certain morphological differences between subalpine and serpentine populations of the putative *G. sudeticum* (Štěpánková 2000) require further detailed study.

This paper contributed to clarifying the taxonomic status of isolated western Bohemian populations traditionallyreferred to as *G. sudeticum*. Currently the availableevidence supports the merger of these serpentine populations with *G. valdepilosum*. First, serpentine plants in western Bohemia share the same nuclear DNA C-values with all the other samples determined as *G. valdepilosum* analysed but differ significantly from those of individuals of *G. sudeticum* in the Krkonoše Mts. Genome size is usually stable at low taxonomic levels and intraspecific variation

often indicates taxonomic heterogeneity (Kron et al. 2007, Loureiro et al. 2010). Consequently, genome size has repeatedly proved to be a useful marker for circumscribing species/subspecies and resolving complex low-level taxonomies (Ekrt et al. 2010, Suda et al. 2010). Another clue comes from the study of their ecological preferences. Environmental conditions at serpentine localities in western Bohemia are virtually identical to those at neighbouring Bavarian serpentines, which host plants invariably identified as *G. valdepilosum* (Noack 1983). In addition, recent morphological investigations(F. Ehrendorfer, pers. comm.) also support the placing of western Bohemian serpentine populations in *G. valdepilosum*. Available data thus suggest that the name *G. sudeticum* should be applied only to subalpine populations currently restricted to the Krkonoše Mts and formerly also occurring in the Hrubý Jeseník Mts. Phenotypic and genome size (Kolář et al. 2013) analyses further indicate that the subalpine populations of *G. sudeticum* are closely related to the highly polymorphic *G. anisophyllon* Villars, which inhabits various neutral to basiphilous subalpine areas in the Alps and Carpathians (Ehrendorfer 1958, Ehrendorfer et al. 1976). The precise taxonomic assignment of serpentine *Galium* populations traditionally referred to as *G. sudeticum* should therefore wait for a detailed assessment of their morphological variation and genetic relationships to other high-altitude taxa.

Finally, we found one distinct but taxonomically uncertain population on a limestone outcrop in the Králický Sněžník Mts in Poland. Although these tetraploid plants are geographically close to the historical *G. sudeticum* occurrence in the Hrubý Jeseník Mts they are ecologically closest to the Alpine-Carpathian species *G. anisophyllon* (note that the Carpathian species *Sesleria tatrae* also occurs on the same outcrop; Fabiszewski 1989). Nevertheless, these plants clearly differ from both *G. anisophyllon* and *G. sudeticum* in their relative genome sizes, and their taxonomic status remains to be clarified.

See http://www.preslia.cz for Electronic Appendices 1–5.

Acknowledgements

We thank J. Chrtek, L. Ekrt, M. Hanzl, J. Janáková, Z. Kaplan, T. Koutecký, P. Kwiatkowski, C. Pachschwöll, J. Prančl, R. Sudová, Z. Szeląg, M. Štech, T. Urfus and P. Vít for their help in the field and/or in providing plant samples, and to F. Ehrendorfer for fruitful discussions on the group investigated. We also thank F. Krahulec, J. Štěpánková and two anonymous reviewers for their helpful comments and suggestions on an earlier version of the manuscript, and T. Dixon for improving our English. The authorities of the Krkonoše Mts National Park are thanked for issuing the collection permit (37/2011). This study was supported by the Czech Science Foundation (project P506/10/0704) and partly also by the Grant Agency of Charles University (project 572213). Additional support was provided by the Academy of Science of the Czech Republic (long-term research development project no. RVO 67985939) and institutional resources of Ministry of Education, Youth and Sports of the Czech Republic for the support of science and research.

References

Baack E. J. (2004): Cytotype segregation on regional and microgeographic scales in snow buttercups (*Ranunculus adoneus*: *Ranunculaceae*). – Am. J. Bot. 91: 1783–1788.

Boyd R. & Martens S. (1998): Nickel hyperaccumulation by *Thlaspi montanum* var. *montanum* (*Brassicaceae*): a constitutive trait. – Am. J. Bot. 85: 259–259.

Brady K. U., Kruckeberg A. R. & Bradshaw H. D. jun. (2005): Evolutionary ecology of plant adaptation to serpentine soils. – Annu. Rev. Ecol. Evol. Syst. 36: 243–266.

Braun-Blanquet J. (1964): Pflanzensoziologie. – Springer, Wien.

Burton T. L. & Husband B. C. (1999): Population cytotype structure in the polyploid *Galax urceolata* (*Diapensiaceae*). – Heredity 82: 381–390.

Castro S., Loureiro J., Procházka T. & Münzbergová Z. (2012): Cytotype distribution at a diploid–hexaploid contact zone in *Aster amellus* (*Asteraceae*). – Ann. Bot. 110: 1047–1055.

Chytrý M. (ed.) (2007): Vegetace České republiky 1. Travinná a keříčková vegetace. [Vegetation of the Czech Republic 1. Grassland and heathland vegetation]. – Academia, Praha.

Chytrý M. (ed.) (2013): Vegetace České republiky 4. Lesní a křovinná vegetace. [Vegetation of the Czech Republic 4. Forest and scrub vegetation]. – Academia, Praha.

Dančák M., Duchoslav M. & Trávníček B. (2012): Taxonomy and cytogeography of the *Molinia caerulea* complex in central Europe. – Preslia 84: 351–374.

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

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

Ehrendorfer F. (1956): Struktur, Verbreitung und Geschichte der Sippen von *Lepto-Galium* in Bayern. – Ber. Bayer. Bot. Ges. 31: 5–12.

Ehrendorfer F. (1958): Die geographische und ökologische Entfaltung des europäisch-alpinen Polyploidkomplexes *Galium anisophyllum* Vill. seit Beginn des Quartärs. – Upsala Univ. Arsskrift 6: 176–181.

Ehrendorfer F. (1960): Neufassung der Sektion *Lepto-Galium* Lange und Beschreibung neuer Arten und Kombinationen. – Sitzungsberichten Österreichischen Akad. Wiss. Math.-Naturwissenschaftliche Kl. Abt. 169: 407–421.

Ehrendorfer F. (1962): Cytotaxonomische Beiträge zur Genese der mitteleuropäischen Flora und Vegetation. – Ber. Deutsch. Bot. Ges. 75: 137–152.

Ehrendorfer F., Krendl F. & Puff C. (1976): *Galium* L. – In: Tutin T. G., Heywood V. H., Burges N. A., Moore D. M., Valentine D. H., Walters S. M. & Webb D. A. (eds), Flora Europaea 4: 14–36, Cambridge Univ. Press, Cambridge.

Ekrt L., Holubová R., Trávníček P. & Suda J. (2010): Species boundaries and frequency of hybridization in the *Dryopteris carthusiana* (*Dryopteridaceae*) complex: a taxonomic puzzle resolved using genome size data. – Am. J. Bot. 97: 1208–1219.

Ellenberg H. (1992): Zeigerwerte der Pflanzen in Mitteleuropa. Ed. 3. – E. Goltze, Göttingen.

Fabiszewski J. (1989): Szata roślinna w sąsiedztwie jaskini [Plants in the vicinity of the cave]. – In: Jahn A., Kozłowski S. & Wiszniowska T. (eds), Jaskinia Niedźwiedzia w Kletnie [Niedźwiedzia cave in Kletno], p. 287–305, Wydawnictwo Polskiej Akademii Nauk, Wroclaw.

Grulich V. (2012): Red List of vascular plants of the Czech Republic: 3rd edition. – Preslia 84: 631–645.

Halverson K., Heard S. B., Nason J. D. & Stireman J. O. (2008): Origins, distribution, and local co-occurrence of polyploid cytotypes in *Solidago altissima* (*Asteraceae*). – Am. J. Bot. 95: 50–58.

Husband B. C., Baldwin S. J. & Suda J. (2013): The incidence of polyploidy in natural plant populations: major patterns and evolutionary processes. – In: Greilhuber J., Doležel J. & Wendel J. F. (eds), Plant genome diversity, Vol. 2, p. 255–276, Springer, Vienna.

Husband B. C. & Sabara H. A. (2004): Reproductive isolation between autotetraploids and their diploid progenitors in fireweed, *Chamerion angustifolium* (*Onagraceae*). – New Phytol. 161: 703–713.

Husband B. C. & Schemske D. W. (1998): Cytotype distribution at a diploid-tetraploid contact zone in *Chamerion* (*Epilobium*) *angustifolium* (*Onagraceae*). – Am. J. Bot. 85: 1688–1694.

Huss-Danell K. (1978): Nitrogenase activity measurements in intact plants of *Alnus incana*. – Physiol. Plant. 43: 372–376.

Kazakou E., Dimitrakopoulos P. G., Baker A. J. M., Reeves R. D. & Troumbis A. Y. (2008): Hypotheses, mechanisms and trade-offs of tolerance and adaptation to serpentine soils: from species to ecosystem level. – Biol. Rev. 83: 495–508.

Kolář F., Dortová M., Lepš J., Pouzar M., Krejčová A. & Štech M. (2014): Serpentine ecotypic differentiation in a polyploid plant complex: shared tolerance to Mg and Ni stress among di- and tetraploid serpentine populations of *Knautia arvensis* (*Dipsacaceae*). – Plant and Soil 374: 435–447.

Kolář F., Fér T., Štech M., Trávníček P., Dušková E., Schönswetter P. & Suda J. (2012): Bringing together evolution on serpentine and polyploidy: spatiotemporal history of the diploid-tetraploid complex of *Knautia arvensis* (*Dipsacaceae*). – PLoS ONE 7: e39988.

Kolář F., Lučanová M., Vít P., Urfus T., Chrtek J., Fér T., Ehrendorfer F. & Suda J. (2013): Diversity and endemism in deglaciated areas: ploidy, relative genome size and niche differentiation in the *Galium pusillum* complex (*Rubiaceae*) in Northern and Central Europe. – Ann. Bot. 111: 1095–1108.

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

Koutecký P., Tuleu G., Baďurová T., Košnar J., Štech M. & Těšitel J. (2012): Distribution of cytotypes and seasonal variation in the *Odontites vernus* group in central Europe. – Preslia 84: 887–904.

Krahulcová A. & Štěpánková J. (1998): Serpentine and polyploid differentiation within *Galium pumilum* agg. (*Rubiaceae*) in eastern central Europe. – Folia Geobot. 33: 87–102.

Krejčíková J., Sudová R., Lučanová M., Trávníček P., Urfus T., Vít P., Weiss-Schneeweiss H., Kolano B., Oberlander K., Dreyer L. L. & Suda J. (2013a): High ploidy diversity and distinct patterns of cytotype distribution in a widespread species of *Oxalis* in the Greater Cape floristic region. – Ann. Bot. 111: 641–649.

Krejčíková J., Sudová R., Oberlander K., Dreyer L. L. & Suda J. (2013b): The spatio-ecological segregation of different cytotypes of *Oxalis obtusa* (*Oxalidaceae*) in contact zones. – South Afr. J. Bot. 88: 62–68.

Krendl F. (1993): Chromosomenzahlen und geographische Verbreitung in der Gattung *Galium* (Sect. *Leptogalium* - *Rubiaceae*). – Biosyst. Ecol. Ser. 4: 51–112.

Kron P., Suda J. & Husband B. C. (2007): Applications of flow cytometry to evolutionary and population biology. – Ann. Rev. Ecol. Evol. Syst. 38: 847–876.

Kucowa I. & Mądalski J. (1964): Nowe stanowiska *Galium valdepilosum* H. Braun w Polsce [New localities of *Galium valdepilosum* H. Braun in Poland]. – Fragm. Florist. Geobot. 10: 1–8.

Kúr P., Štech M., Koutecký P. & Trávníček P. (2012): Morphological and cytological variation in *Spergularia echinosperma* and *S.rubra*, and notes on potential hybridization of these two species. – Preslia 84: 905–924.

Lang G. (1994): Quartäre Vegetationsgeschichte Europas. Methoden und Ergebnisse. – Gustav Fischer Verlag,

Stuttgart.

Lepš J. & Šmilauer P. (2003): Multivariate analysis of ecological data using CANOCO. – Cambridge University Press, Cambridge.

Levin D. A. (1975): Minority cytotype exclusion in local plant populations. – Taxon 24: 35–43.

Levin D. A. (2002): The role of chromosomal change in plant evolution. – Oxford Univ. Press, Oxford.

Loureiro J., Trávníček P., Rauchová J., Urfus T., Vít P., Štech M., Castro S. & Suda J. (2010): The use of flow cytometry in the biosystematics, ecology and population biology of homoploid plants. – Preslia 82: 3–21.

Ložek V. (1973): Příroda ve čtvrtohorách [Nature in the Quaternary]. – Academia, Praha.

Mandáková T. & Münzbergová Z. (2006): Distribution and ecology of cytotypes of the *Aster amellus* aggregate in the Czech Republic. – Ann. Bot. 98: 845–856.

Manzaneda A. J., Rey P. J., Bastida J. M., Weiss-Lehman C., Raskin E. & Mitchell-Olds T. (2012): Environmental aridity is associated with cytotype segregation and polyploidy occurrence in *Brachypodium distachyon* (*Poaceae*). – New Phytol. 193: 797–805.

Mráz P., Šingliarová B., Urfus T. & Krahulec F. (2008): Cytogeography of *Pilosella officinarum* (*Compositae*): altitudinal and longitudinal differences in ploidy level, distribution in the Czech Republic and Slovakia and the general pattern in Europe. – Ann. Bot. 101: 59–71.

Noack M. (1983): Untersuchungen an ausseralpinen Sippen von *Galium* Sektion *Lepto-Galium* in Bayern. – MSc. thesis, München, Germany.

Otto S. P. & Whitton J. (2000): Polyploid incidence and evolution. – Ann. Rev. Genet. 34: 401–437.

Reeves R. D. & Baker A. J. M. (1984): Studies on metal uptake by plants from serpentine and non-serpentine populations of *Thlaspi goesingense* Hálácsy (*Cruciferae*). – New Phytol. 98: 191–204.

Šafářová L. & Duchoslav M. (2010): Cytotype distribution in mixed populations of polyploid *Allium oleraceum* measured at a microgeographic scale. – Preslia 82: 107–126.

Šafářová L., Duchoslav M., Jandová M. & Krahulec F. (2011): *Allium oleraceum* in Slovakia: cytotype distribution and ecology. – Preslia 83: 513–527.

Skalińska M., Piotrowicz M. & Sokołowska-Kulczycka A. (1961): Further additions to chromosome numbers of Polish Angiosperms. – Acta Soc. Bot. Poloniae 30: 463–489.

Sonnleitner M., Flatscher R., García P. E., Rauchová J., Suda J., Schneeweiss G. M., Hülber K. & Schönswetter P. (2010): Distribution and habitat segregation on different spatial scales among diploid, tetraploid and hexaploid cytotypes of *Senecio carniolicus* (*Asteraceae*) in the Eastern Alps. – Ann. Bot. 106: 967–977.

StatSoft (2007): STATISTICA: data analysis software system, version 8.0. – StatSoft, Inc., Tulsa, www.statsoft.com.

Stebbins G. L. (1985): Polyploidy, hybridization, and the invasion of new habitats. – Ann. Missouri Bot. Gard. 72: 824–832.

Štěpánková J. (1997): The effect of serpentine on morphological variation in the *Galium pumilum* group (*Rubiaceae*). – Thaiszia 7: 29–40.

Štěpánková J. (2000): *Galium* L. – In: Slavík B. (ed.), Květena České republiky [Flora of the Czech Republic] 6: 122–156, Academia, Praha.

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 Geobot. 36: 303–320.

Suda J., Malcová R., Abazid D., Banaš M., Procházka F., Šída O. & Štech M. (2004): Cytotype distribution in *Empetrum* (*Ericaceae*) at various spatial scales in the Czech Republic. – Folia Geobot. 39: 161–171.

Suda J., Trávníček P., Mandák B. & Berchová-Bímová K. (2010): Genome size as a marker for identifying the invasive alien taxa in *Fallopia* section *Reynoutria*. – Preslia 82: 97–106.

Suda J., Weiss-Schneeweiss H., Tribsch A., Schneeweiss G. M., Trávníček P. & Schönswetter P. (2007): Complex distribution patterns of di-, tetra-, and hexaploid cytotypes in the European high mountain plant *Senecio carniolicus* (*Asteraceae*). – Am. J. Bot. 94: 1391–1401.

Tausch I. F. (1835): Bemerkungen über *Galium* und einige verwandte Gattungen. – Flora 18: 337–351.

Taylor S. I. & Levy F. (2002): Responses to soils and a test for preadaptation to serpentine in *Phacelia dubia* (*Hydrophyllaceae*). – New Phytol. 155: 437–447.

Tichý L. (2002): JUICE, software for vegetation classification. – J. Veg. Sci. 13: 451–453.

Trávníček P., Dočkalová Z., Rosenbaumová R., Kubátová B., Szeląg Z. & Chrtek J. (2011a): Bridging global and microregional scales: ploidy distribution in *Pilosella echioides* (*Asteraceae*) in central Europe. – Ann. Bot. 107: 443–454.

Trávníček P., Eliášová A. & Suda J. (2010): The distribution of cytotypes of *Vicia cracca* in Central Europe: the changes that have occurred over the last four decades. – Preslia 82: 149–163.

Trávníček P., Jersáková J., Kubátová B., Krejčíková J., Bateman R. M., Lučanová M., Krajníková E., Těšitelová T., Štípková Z., Amardeilh J.-P., Brzosko E., Jermakowicz E., Cabanne O., Durka W., Efimov P., Hedrén M., Hermosilla C. E., Kreutz K., Kull T., Tali K., Marchand O., Rey M., Schiestl F. P., Čurn V. & Suda J. (2012): Minority cytotypes in European populations of the *Gymnadenia conopsea* complex (*Orchidaceae*) greatly increase intraspecific and intrapopulation diversity. – Ann. Bot. 110: 977–986.

Trávníček P., Kubátová B., Čurn V., Rauchová J., Krajníková E., Jersáková J. & Suda J. (2011b): Remarkable coexistence of multiple cytotypes of the *Gymnadenia conopsea* aggregate (the fragrant orchid): evidence from flow cytometry. – Ann. Bot. 107: 77–87.

Westerbergh A. (1994): Serpentine and non-serpentine *Silene dioica* plants do not differ in nickel tolerance. – Plant and Soil 167: 297–303.

Zarzycki K. & Kaźmierczakowa R. (2001): Polska czerwona księga roślin. Paprotniki i rośliny kwiatowe [Polish Red Data Book of plants: pteridophytes and flowering plants]. – Instytut Botaniki im. W. Szafera PAN, Kraków.

Case study III

Kolář F., Lučanová M., Záveská E., Fuxová G., Mandáková T., Španiel S., Senko D., Svitok M., Kolník M., Gudžinskas Z., Marhold K. (2016): **Ecological segregation does not drive the intricate parapatric distribution of diploid and tetraploid cytotypes of the *Arabidopsis arenosa* group (Brassicaceae). – Biological Journal of the Linnean Society 119: 673-688**

Ecological segregation does not drive the intricate parapatric distribution of diploid and tetraploid cytotypes of the *Arabidopsis arenosa* group (Brassicaceae)

Filip Kolář[1,2], Magdalena Lučanová[2,1], Eliška Záveská[1], Gabriela Fuxová[1], Terezie Mandáková[3], Stanislav Španiel[1], Dušan Senko[4], Marek Svitok[5,6], Martin Kolník[7], Zigmantas Gudžinskas[8], Karol Marhold[1,4*]

[1]Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, CZ-128 01 Prague, Czech Republic

[2]Institute of Botany, Academy of Sciences of the Czech Republic, Zámek 1, CZ-252 43 Průhonice, Czech Republic

[3]Plant Cytogenomics Research Group, Central European Institute of Technology (CEITEC), Masaryk University, Kamenice 5, CZ-62500 Brno, Czech Republic

[4]Institute of Botany, Slovak Academy of Sciences, Dúbravská cesta 9, SK-845 23 Bratislava, Slovak Republic

[5]Department of Biology and General Ecology, Faculty of Ecology and Environmental Sciences,Technical University in Zvolen, T. G. Masaryka 24, SK-960 53 Zvolen, Slovak Republic

[6]Eawag Swiss Federal Institute of Aquatic Science and Technology, Department of Aquatic Ecology, Centre of Ecology, Evolution and Biogeochemistry, Seestrasse 79, CH-6047 Kastanienbaum, Switzerland

[7]Tematínska 4, SK-91501 Nové Mesto nad Váhom, Slovak Republic

[8]Nature Research Centre, Institute of Botany, Laboratory of Flora and Geobotany, Žaliu̧ju̧ Ežeru̧ Str. 49, LT-08406 Vilnius, Lithuania

*Corresponding author. E-mail: karol.marhold@savba.sk

Abstract

Detailed knowledge of the geographic distribution of cytotypes is a prerequisite for any experimental or molecular study of ploidy-variable plant systems. The *Arabidopsis arenosa* group, an intricate di-tetraploid complex from the plant model genus *Arabidopsis*, has remained largely neglected regarding the distribution and habitat associations of its cytotypes. Using flow cytometry, we conducted a large population-level cytological screen across the *A. arenosa* group range, involving more than 2900 individuals from 194 populations. We characterized a largely parapatric distribution of the diploid (Southeast Europe) and tetraploid (Northwest Europe) cytotypes with two contact zones – a narrow contact zone in the Slovenian Forealps and a diffuse contact zone across the Carpathians. In addition, a previously unknown isolated diploid lineage with distinct ecology was revealed from sandy areas of the southeastern Baltic coast. We also recorded several adult triploid individuals for the first time in wild *Arabidopsis arenosa*. Particularly in the Western Carpathians, the diploid and tetraploid populations are largely intermingled, and both cytotypes are spread along the whole lowland-alpine gradient of habitats, exhibiting no signs of ploidy-linked habitat differentiation. In contrast with the complexity at the landscape scale, the within-population cytological homogeneity and the rare occurrence of triploids indicate that the contact zone is rather stable.

Keywords: contact zones – *Cruciferae* – environmental predictors – flow cytometry – habitat differentiation – polyploidy.

Introduction

The *Arabidopsis arenosa* group, a diploid-tetraploid species complex, represents one of the closest relatives of the prominent plant model *Arabidopsis thaliana* (Clauss & Koch, 2006). Polyploidization is a major diversification force in the complex, generating an intricate mixture of diploid populations and their tetraploid derivatives. Importantly, origin of the tetraploid populations solely from diploid representative(s) of the *A. arenosa* group is suggested by the cytotype distribution pattern, morphological similarities (Měsíček, 1970), close AFLP multilocus phenotypes (Schmickl *et al*., 2012) and overall similarity in genome scans (Et al.). The close relationships among diploid and tetraploid *A. arenosa* cytotypes represent a unique feature within *Arabidopsis*, as other wild polyploid members are of allopolyploid (hybrid) origin, based on more distantly related parents (*A. suecica*, Jakobsson et al., 2006; *A. lyrata* subsp. *petraea*, Schmickl & Koch, 2011; *A. kamchatica*, Shimizu-Inatsugi *et al*., 2009). The *Arabidopsis arenosa* group thus emerges as a highly promising system for addressing general questions on polyploidy in natural plant populations. Indeed, the first studies dealing with general evolutionary questions in this group have emerged recently, addressing the evolution of meiotic stability in polyploids (Hollister et al., 2012; Yant et al., 2013) and speciation processes (Jørgensen et al., 2011; Schmickl & Koch, 2011).

The *A. arenosa* group comprises up to nine taxa (species or subspecies, partly still not formally described) spanning a wide ecological range from coastal sand dunes to high-alpine environments with a principal diversity centre most likely situated in the Carpathian Mountains in eastern Central Europe (Měsíček, 1998; Měsíček & Goliašová, 2002; Schmickl et al., 2012). Available cytological data indicate that the Carpathian mountain arch harbours a complex mixture of diploid and tetraploid populations [chromosome counts by Měsíček (1970), F. Krendl and A. Polatschek (published in Schmickl et al., 2012)]. In particular, the Western Carpathians appear to be a hotspot of ecological and taxonomic diversity of the whole species complex. There, populations of both diploid and tetraploid representatives of the *A. arenosa* group co-occur along the entire altitudinal gradient, from dry and warm steppes in the foothills (150 m a.s.l.) via shady rocks and screes on various substrates to alpine vegetation on the highest summits (2600 m a.s.l., Měsíček & Goliašová, 2002). This extensive cyto- and eco-geographical variation is remarkable both in general and particularly in the Carpathians, where the largest cytotype mixture of the *A. arenosa* group is found.

In the Carpathians, the few large-scale cytotype screens published to date are inconclusive with respect to general cytogeographic patterns. They range from near cytological homogeneity (*Vicia cracca*, Trávníček, Eliášová & Suda, 2010; *Alyssum montanum*, Španiel et al., 2011) through the absence of geographical patterns and extensive intrapopulation cytotype mixture (*Phleum pratense* agg.; Perný et al., 2008) to a relatively strong altitudinal differentiation (*Sesleria calcarea – S. tatrae* species complex, Lysak & Doležel, 1998; *Senecio jacobaea*, Hodálová et al., 2007; *Pilosella officinarum*, Mráz et al., 2008; *Knautia arvensis* agg., Kolář et al., 2009). However, none of these studied species spans the entire altitudinal range of habitats.

A prerequisite for any ecological and/or molecular study of a ploidy-heterogeneous plant system is knowledge of the geographic distribution of cytotypes. Cytogeographic data complement phylogenetic and experimental data and serve as a foundation for addressing questions of frequency

of polyploid formation, ecological differentiation of cytotypes, and the genetic background of polyploid evolution. For comprehensive evaluation of the true extent of diversity and dynamics of ploidy-mixed plant systems (e.g., detection of minority-represented cytotypes such as triploids), a sufficiently large and geographically wide flow cytometric screen is essential (Duchoslav, Šafářová & Krahulec, 2010; Trávníček et al. 2011a,b; Krejčíková et al. 2013, see Kron, Suda & Husband, 2007 for review). Despite an increasing interest in evolutionary, ecological, and genomic studies of the *A. arenosa* group, we still have only fragmentary knowledge on its karyological diversity and habitat associations. Most of the published records on the ploidy distribution are based on traditional lowthroughput chromosome counting (allowing ploidy determination of a few individuals per population) and/or focus on the uniform tetraploid-inhabited regions of Western and Northern Europe (Měsíček, 1970; Schmickl *et al.* 2012).

In this study, we employed a high-throughput technique for ploidy estimation – flow cytometry – complemented with chromosome counts to assess ploidy level and homoploid genome size diversity over the entire distribution range of the *A. arenosa* group. Considering the intricate and still unresolved relationships within this group, our study addressed only general patterns across the whole species complex and did not aim to resolve its internal taxonomic structure. Specifically, we addressed the following questions: (1) What is the pattern of ploidy distribution, especially of the so far undersampled diploids, and where are the cytotype contact zones located? (2) What is the ploidy level variation within populations? Are there any indications of recent polyploidization events and/or inter-ploidy gene flow? (3) What is the level of variation in DNA content at the homoploid level and, if present, is this variation geographically structured? (4) Are there any indications for substantial niche differentiation between the cytotypes along large-scale environmental gradients (altitude, climatic niche, substrate, disturbance levels)? If so, is the differentiation stronger in the areas where both cytotypes co-occur in sympatry (Western Carpathians)?

Material and methods

Field sampling

In total, 2963 individuals from 194 populations were collected across the entire range of the *Arabidopsis arenosa* group from 2011 to 2013. The sampling covered all currently recognised species and subspecies of the complex (except for the geographically, morphologically and ecologically distinct diploid stenoendemic *A. croatica*), namely *Arabidopsis arenosa* (L.) Lawalrée subsp. *arenosa*, *A. arenosa* subsp. *borbasii* (Zapał.) O'Kane & Al-Shehbaz, *A. carpatica* nom. prov., *A. neglecta* (Schult.) O'Kane & Al-Shehbaz subsp. *neglecta* nom. prov., *A. neglecta* subsp. *robusta* nom. prov., *A. nitida* nom. prov., *A. petrogena* (A. Kern.) V.I. Dorof. subsp. *petrogena* nom. prov., *A. petrogena* subsp. *exoleta* nom. prov. The above-mentioned provisional names on the level of species and subspecies were introduced in the genus *Cardaminopsis* by Měsíček (1970, 1998 and unpublished manuscript), but they were never validly published. Valid publication of these names requires further studies, and we are using them solely for a reference to other papers using this nomenclature (corresponding names are also used in the locality list in Table S1). Whenever possible, fresh tissues (preferably parts of stems with flowers) mostly from 1 to 20 (up to 51) individuals per population (15 individuals on average) were collected and placed in cold storage until flow cytometric evaluation. In selected populations, we also collected seeds for direct counts of chromosome numbers. We recorded GPS co-ordinates and altitude and characterized the

environmental conditions of each site using the following parameters: habitat type, geological substrate and natural/anthropogenic character. Localities were considered anthropogenic only in cases of heavily humandisturbed or entirely human-created habitats (wall crevices, railway tracks, gravel deposits, etc.). Nevertheless, these taxa often colonise such sites as a result of accidental spreading from adjacent natural stands (e.g., road bank below a rock). To differentiate between such short-distance spontaneous colonization and long-distance anthropogenic spread, we further divided the anthropogenic stands into those close (less than approximately 1 km) to a natural habitat and those occupying purely anthropogenic habitats far from any potential natural locality (typically road banks and railway tracks). For locality details, see supplementary Table S1.

Flow cytometry

DNA ploidy level (Suda *et al*., 2006) was inferred from nuclear DNA content determined by flow cytometry following the simplified two-step protocol (Doležel, Greilhuber & Suda, 2007). Approximately 10 square millimetres of fresh leaf tissue or one fresh petal from each plant to be analysed was chopped together with an appropriate volume of the internal reference standard (*Solanum pseudocapsicum*, 2C = 2.59 pg, Temsch, Greilhuber & Krisai, 2010; the same individual was used for all measurements) using a sharp razor-blade in a Petri dish containing 0.5 mL of icecold Otto I buffer (0.1 M citric acid, 0.5% Tween 20). The suspension was filtered through 42-µm nylon mesh and incubated for 10 min at room temperature. Isolated nuclei were stained with 1 mL of Otto II buffer (0.4 M $Na_2HPO_4 \cdot 12H_2O$) supplemented with 4,6-diamino-2-phenylindole (DAPI) at 4 µg mL^{-1} and β-mercaptoethanol at 2 µg mL^{-1}. After 1 min of incubation, the sample was run for 3000 particles in a Cyflow ML flow cytometer (Partec, Münster, Germany) equipped with the UV-led lamp. The histograms were evaluated with FloMax FCS 2.0 software (Partec, Münster, Germany). Fresh petals were preferred over vegetative parts for these analyses due to the absence of endopolyploidy (Galbraith, Harkins & Knapp, 1991). For petal samples, we analysed up to five individuals in a pooled sample to reduce the analysis costs and time demand. Our previous experiments showed that such practice enables reliable detection of minority cytotypes present even at a low proportion (20%). Nevertheless, each plant was separately re-analysed if mixed samples were suspected, peaks were asymmetrical, or the coefficient of variance of the *Arabidopsis* peak exceeded 5%. The same approach was applied for pooled leaf samples of tetraploids (a potential diploid or triploid individual would be clearly identified as an additional peak with lower fluorescence intensity); however, vegetative parts from diploid individuals were analysed individually in any case due to the presence of the tetraploid endopolyploid peak. In ten (5%) populations where fresh tissue was not available, we used samples dried with silica gel for ploidy estimation using the same protocol (see Table S1).

For genome size estimation, one individual per selected population (see Table S1) was run on a CyFlow SL flow cytometer (Partec, Münster, Germany) equipped with a green (532 nm) solid-state laser. The sample preparation followed the methodology described above, with the only modification being that the stain solution consisted of Otto II buffer enriched with propidium iodide and RNase (both at 50 µg mL^{-1}) and β-mercaptoethanol at 2 µg mL^{-1}. The analyses were run for 5000 particles. We applied the following stringent criteria to obtain precise and stable estimates of genome size: (i) only analyses with the coefficient of variation of the sample peak below 3% were taken into account, (ii) each sample was measured at least three times on different days to minimise potential random instrumental drift (Doležel & Bartoš, 2005), and (iii) the between-day variation was defined to not exceed 3%; otherwise, the most remote value was discarded and the

sample was re-analysed. The reliability of flow cytometric measurements (i.e., between-plant differences) was repeatedly confirmed in simultaneous runs of *Arabidopsis* accessions with distinct genome sizes (Greilhuber, 2005).

Chromosome preparations

Plants for chromosome counts were selected such that they covered the entire sampling area. Plants were grown from seeds in plastic Petri dishes on sieved potting soil in a phytotron with long day illumination (16 h light at 20 °C, 8 h dark at 15 °C). Young inflorescences were fixed in ethanol/acetic acid (3 : 1, v/v) fixative for 24 h at 4 °C. The fixative was replaced with 70% ethanol, and the material was stored at −20 °C until further use. Chromosome spreads were prepared as described by Mandáková, Marhold & Lysak, (2014). Slides were examined under phase contrast for the presence of suitable mitotic metaphase spreads. Selected preparations were stained with 2 mg mL−1 DAPI in Vectashield anti-fade mounting medium (Vector Laboratories, Burlingame, CA, USA) and photographed using an Olympus BX-61 epifluorescence microscope and a CoolCube camera (MetaSystems, Altlussheim, Germany). Individual images were processed with Photoshop CS software (Adobe Systems, San Jose, CA, USA).

Data analyses

Spatial segregation of cytotypes across the entire range (except for the spatial outlier Scandinavian populations) and separately within the Western Carpathian contact zone was analysed using the Mantel test implemented in the ade4 R package (Dray & Dufour, 2007). A correlation coefficient (r_M) was calculated for: (i) the matrix of mutual geographic distances among populations; and (ii) the binary matrix of ploidy levels, and it was compared to the distribution of coefficients obtained from matrices generated by random rearrangements (9999 permutations) of the original matrices. Only the majority ploidy level of the population was considered (i.e., rare triploid cytotypes were omitted). In addition, Mantel tests were used for testing the spatial autocorrelation of homoploid genome size by comparing a matrix of geographic distances with genome size distance matrix for a particular cytotype (diploid and tetraploid accessions were analysed separately).

Differences among the cytotypes in associations with anthropogenic stands and geological substrates (assessed only for non-anthropogenic populations) were assessed using the chi-squared test in contingency tables (*P*-values were assessed using 200 replicates). General linear models were used for testing the association of cytotypes with altitude as well as for the relationships among homoploid genome size and the following environmental predictors: (non)anthropogenic character of the original habitat, altitude, and substrate type (the last one only for natural localities). Unless stated otherwise, all analyses were performed in R 2.15.2 (R Development Core Team, 2013).

To capture the interrelationship of environmental predictors and ploidy level in sufficiently detailed scale, it was necessary to use background climatic and landscape data, which are long-term averages and provide seasonal variability. Primary data layers that included air temperature, solar radiation, and terrain (elevation, horizon) were obtained from the SolarGIS data, version 1.9 (the high-resolution climate database operated by GeoModel Solar, Bratislava, Slovakia). Data on air temperature at 2 m (in °C) were derived from the Climate Forecast System Reanalysis and Global Forecast System databases (National Centers for Environmental Prediction,

Suitland, Maryland, USA) for the period from 1994 to 2011, recalculated to 15-minute values. The data were spatially enhanced to 1-km resolution to reflect variability induced by high-resolution (dissected) terrain. Solar radiation was calculated from the satellite and atmospheric data. The sources were: (i) Meteosat First and Second Generation (PRIME and Indian Ocean Data Coverage Regions, European Organisation for the Exploitation of Meteorological Satellites, Darmstadt, Germany) in 15-min or 30-min values, (ii) outputs from the Monitoring Atmospheric Composition and Climate (European Centre for Medium-Range Weather Forecasts, Reading, UK) for the decade from 2003 to 2013, and (iii) atmospheric models from Global Forecast System database (National Oceanic and Atmospheric Administration, Silver Spring, Maryland, USA) for the period from 1994 to 2013. Solar radiation represents annual (total) and monthly long-term averages of global irradiation: (i) without (global horizontal irradiation, GHI), and (ii) with impinging on local terrain accounting for the slope and azimuth of the terrain (GTI) (in $kWhm^{-2}$) and annual (total) and monthly long-term averages of photosynthetically active radiation (PAR) (400–700 nm in $kWhm^{-2}$). Monthly long-term averages of precipitation were obtained from WorldClim, version 1.4 (Hijmans *et al.* 2005). For the purpose of this study, the hourly data on air temperature and solar radiation were integrated into long-term monthly averages. These averages were further spatially enhanced by disaggregation, based on the correlation between terrain altitude and climatic variables. The disaggregated monthly and yearly averages created from this reanalysis were validated against selected ground measurements (from the meteostations flagged with quality codes 2, 3, 6, 7; see list of quality codes from the National Climatic Data Center). Based on disaggregation and validation, which was calculated individually for each pixel (smallest grid unit), these data (rasters) in the GIS (Geographic Information System) environment represent annual trends, seasonality and extremes for particular areas. Morphometry of the terrain (terrain slope, terrain azimuth) was developed via elevation [altitude above sea level; source SRTM3 data (The Shuttle Radar Topography Mission, available at http://srtm.usgs.gov/) up to the latitude 60°N]. We calculated distances from the Equator (northing) and the prime meridian (easting) in kilometres to account for spatial gradients and autocorrelation. For these calculations, we used PostGIS/PostgreSQL, version 1.5.1, released under the GNU/GPL license.

Distribution of the major ploidy levels (diploids and tetraploids) was modelled using generalized linear models (GLM) with binomial error distribution and the logit link function (i.e., logistic regression). A range of GIS-derived data was used as environmental explanatory variables (see Table S2 for a complete list of variables and abbreviations of variable names). Northing and easting were used as spatial predictors to detect possible geographic gradients. Prior to the analyses, distribution of variables and correlations among them were assessed. To avoid a multicollinearity, elevation was excluded from modelling due to its strong correlation with mean annual temperature ($r = –0.94$). Intrinsically strong positive correlations were found among monthly values and annual summary characteristics of temperature (Fisher weighted mean $r = 0.98$), precipitation ($r = 0.82$), GHI ($r = 0.85$), GTI ($r = 0.92$) and PAR ($r = 0.85$); thus, only annual characteristics were pre-selected for further analyses. However, annual GTI, GHI and PAR were highly correlated with each other ($r = 0.98$). Consequently, only PAR was employed as a predictor in the analyses because this quantity is intuitively understandable and is a frequently used measure of radiation. The remaining variables did not show considerable skewness or intercorrelations and were used in the modelling procedure as predictors (see Table S3 for a list of predictors). Separate GLMs were built for the whole dataset and the Western Carpathian contact zone. Initially, full models were fitted to the data, including all spatial and environmental predictors. The full models

were simplified following backward stepwise deletion associated with likelihood-ratio tests. Only those variables for which the conditional effect was significant at $\alpha = 5\%$ were retained in the final models. Spatial correlograms were used to check for autocorrelation in the residuals of the final models. Because the final models showed significant positive autocorrelation at short distances, the data were re-fitted using generalized mixed effect models (GLMM) (Dormann *et al.* 2007) to prevent biased estimates of model coefficients and the inflation of type I errors. GLMMs with Gaussian spatial correlation structure were fitted using penalised quasi-likelihood (Venables & Ripley, 2002). Final GLMMs are presented graphically as a series of effect plots (Fox, 2003). The ability of the final models to discriminate between sites with diploids and those with tetraploids was assessed by means of classification tables (cut-off value: 0.5) and Somers' Dxy rank correlations (Newson, 2006) between observed incidences of cytotypes and predicted probabilities.

Results

Ploidy level variation and cytogeography

Three different DNA ploidy levels (diploid – 2*x*, triploid – 3*x*, and tetraploid – 4*x*) were detected among 2963 individuals from 194 populations belonging to the *A. arenosa* group (Fig. 1). The tetraploid individuals [1588 (54%) individuals in 107 (55%) populations] only slightly prevailed over their diploid counterparts [1369 (46%) individuals in 88 (45%) populations]. The triploid cytotype was extremely rare (six individuals, 0.2%) and it was in all cases represented by a single individual each in otherwise diploid populations. Despite cytotype co-occurrence in several areas and a large within-population sampling (15 individuals per population were sampled on average), the vast majority of the populations (96%) were detected as cytotype uniform, i.e., either diploid or tetraploid. Only a single di-tetraploid mixed-ploidy population was found in the Tatry Mts. (Western Carpathians, AA170) in addition to diploid-triploid mixtures recorded at six sites across the diploid cytotype range (see Table S1, for locality details). Chromosome counts confirmed the estimated ploidy levels and revealed $2n = 2x = 16$ in 17 accessions from the Carpathians (AA018, AA023, AA070, AA084, AA090, AA091, AA123, AA157), Dinaric Alps (AA054, AA124, AA125, AA126, AA127, AA128), Pannonian lowland (AA110), and southern Baltic coast (AA153, AA200) and $2n = 4x = 32$ in 10 accessions from the Carpathians (AA015, AA067, AA082, AA087, AA088), southern and eastern Alps (AA049, AA149), southern Poland (AA059), Scandinavia (AA181) and Luxembourg (AA190) (Fig. 2, Table S1). Neither dysploidy, aneuploidy nor accessory chromosomes were observed in the karyologically investigated accessions.

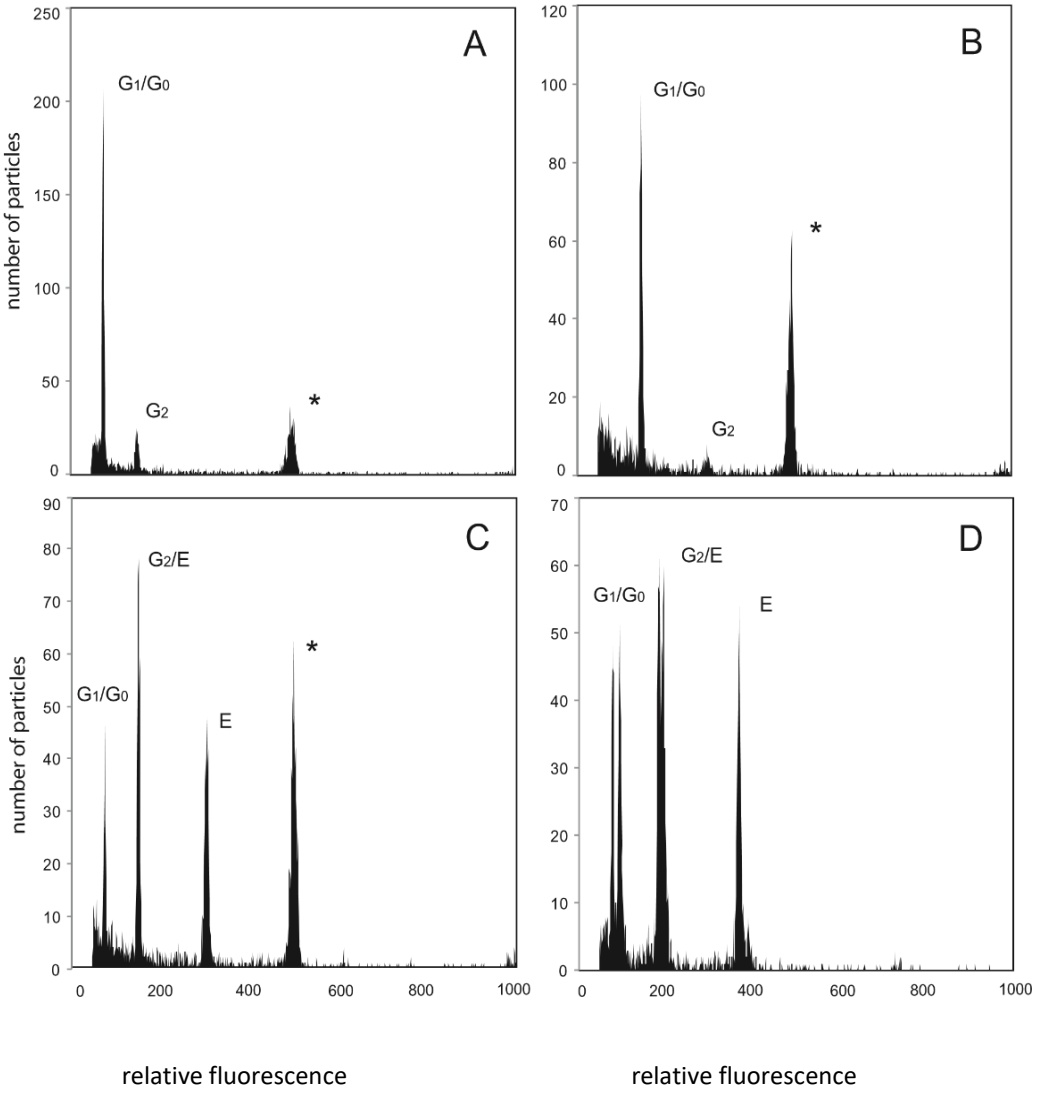


Figure 1. Flow cytometric histograms of suspensions of DAPI-stained nuclei isolated from diploid (A, C, D) and tetraploid (B) accessions of the *Arabidopsis arenosa* group. A + C, Analysis of nuclei of identical diploid individuals (pop. AA084) isolated from either fresh petal (A) or stem leaf (C). B, Pooled sample of five tetraploid individuals (pop. AA117, nuclei isolated from fresh petal tissue). D, Simultaneous analysis of two diploid accessions from pop. AA090 documenting within-population divergence in nuclear DNA contents (difference in fluorescence intensity, 14%; nuclei from both samples were simultaneously isolated, stained, and analysed). Letters denote peaks of nuclei corresponding to different phases of the cell cycle (G_0–G_2) and/or levels of endopolyploidy (E); the internal standard *Solanum pseudocapsicum* used in analyses A-C is marked by an asterisk

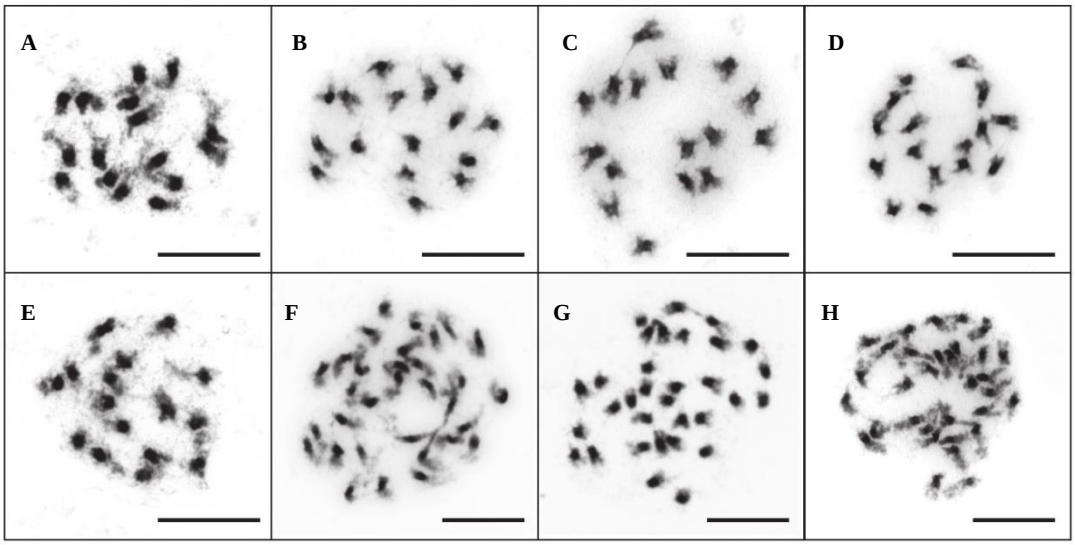


Figure 2. DAPI-stained mitotic chromosome spreads from flower bud tissue of the *Arabidopsis arenosa* group. A, *A. arenosa* s.l. AA200 (Lithuania, coastal sands; 2*n* = 2*x* = 16). B, *A. arenosa* s.l. AA124 (Serbia, dry rocks; 2*n* = 2*x* = 16). C, *A. arenosa* s.l. AA070 (Romania, dry rocks; 2*n* = 2*x* = 16). D, *A. neglecta* AA084 (Slovakia, alpine scree; 2*n* = 2*x* = 16). E, *A. carpatica* AA023 (Slovakia, limestone outcrop in middle altitudes; 2*n* = 2*x* = 16). F, *A. arenosa* AA181 (Norway, secondary gravel; 2*n* = 4*x* = 32). G, *A. neglecta* subsp. *robusta* AA087 (Slovakia, alpine rocks; 2*n* = 4*x* = 32). H, *A. petrogena* subsp. *exoleta* AA082 (Romania, limestone rocks; 2*n* = 4*x* = 32). See Table S1 for locality details. Scale bars = 10 μm.

Diploid and tetraploid cytotypes exhibited a largely parapatric distribution; a weak but significantly nonrandom spatial differentiation of cytotypes was also supported by the Mantel test (r_M = 0.06, P<0.001). Tetraploids dominate in the northwestern half of the A. arenosa group range (Scandinavia, Germany, Alps, Hercynian massif) whereas diploids occupy mainly southeastern areas (most of the Carpathians, Pannonian basin, Dinaric Alps, Fig. 3). In addition, four spatially isolated diploid populations were found along southern shores of the Baltic Sea. They grew exclusively in coastal sand dunes and in adjacent open forests and thus occupied distinct environments from their spatially closest tetraploid counterparts that were found exclusively in human-disturbed habitats (Table S1). Natural populations of both cytotypes meet at the landscape scale in two contact zones, a smaller and rather abrupt one situated in Slovenia (less than 100 km wide) and a large and diffuse zone across the Carpathian mountain arch (Fig. 4). In the Romanian Carpathians, the tetraploids occupy the northern half of the Eastern Carpathians and the Apuseni Mts., whereas diploids dominate in Southern Carpathians and in the southern half of the Eastern Carpathians (the only exceptions in this area are two tetraploid populations, AA065 and AA067, occupying alpine screes and a limestone canyon, respectively). In contrast, in the Western Carpathians, the diploid and tetraploid populations were largely spatially intermingled throughout the landscape (Fig. 4) although the cytotypes still exhibited weak but significant spatial associations (Mantel test, r_M = 0.06, P = 0.013).

Homoploid differentiation in DNA content

In addition to ploidy variation, the accessions of the *A. arenosa* group also exhibited a considerable variation in DNA content at the homoploid level as the di- and tetraploid accessions varied 1.17-fold and 1.21-fold, respectively. Nevertheless, this range included two diploid individuals and one tetraploid individual with abruptly higher genome sizes (9–13% higher than the average, see Fig. S1). After exclusion of these three individuals, the variation dropped to 1.12-fold and 1.14-fold in diploids and tetraploids, respectively. Homoploid genome size was not spatially structured, as evidenced by non-significant Mantel tests ($r_M = -0.11$, $P = 0.88$ and $r_M = 0.07$, $P = 0.19$, for diploid and tetraploid accessions, respectively). In addition, a comparable 1.14-fold difference was found among five individuals from one exceptionally highly variable population from the alpine zone of the Western Carpathians (pop. AA090, see also Fig. 1D). The genome size variation was correlated neither with (non)-anthropogenic habitat character ($F_{1,31} = 0.92$, $P = 0.34$ in tetraploids; diploids were not tested due to the negligible proportion of anthropogenic populations, see the next section) nor with altitude, substrate type and/or their interaction ($F_{3,30} = 0.61$, $P = 0.613$ and $F_{3,20} = 1.37$, $P = 0.28$ in diploids and tetraploids, respectively). Mean monoploid DNA content (after exclusion of the individuals with exceptionally high values) was similar among all three ploidy levels, though it was not entirely identical (average ratio to internal standard divided by ploidy level was 0.068, 0.070, and 0.073 for 2*x*, 3*x*, and 4*x*, respectively: the tetraploid value was on average 7.6% higher than that of diploids).

Niche differentation

Distribution of ploidy levels through the entire investigated area was significantly correlated to a west/east gradient (easting), total annual PAR and total annual precipitation (Table 1). Probability of tetraploid occurrence decreased toward the east and also with increasing PAR and precipitation (Fig. S2). Considering contact zone data, only a south/north gradient appeared significant (Fig. S3). Generally, the cytotypes occupied somewhat different climatic niches as revealed by the moderately high discriminatory power of the models. Nevertheless, the particular climatic factors strongly reflected by the spatial gradients and only two environmental predictors (total annual PAR and total annual precipitation) significantly improved the GLMM with incorporated geographical predictor in the entire A. arenosa group area. No environmental predictors were shown to be significant in the contact zone (Table 1).

Almost no significant differences in substrate requirements and/or altitudinal ranges of the diploids or tetraploids were detected, whether across the entire area or in the densely sampled zone of sympatry in the Western Carpathians. The only exception was a significant association of tetraploids with anthropogenic stands (Table 2). Although both cytotypes were able to grow in habitats created or disturbed by man in close proximity to the natural stands (14 vs. 10 localities for tetraploids and diploids, respectively), the tetraploids were significantly more frequent (22 vs. 4 localities) in anthropogenic stands distant from natural localities, i.e., showing stronger potential for anthropogenic spread. Nevertheless, this difference was not apparent within the Western Carpathian contact zone because tetraploids occupied the anthropogenic stands in other parts of the distributional range (mainly in the northern part, i.e., Scandinavia, Poland, northern Germany, and northern Czech Republic).

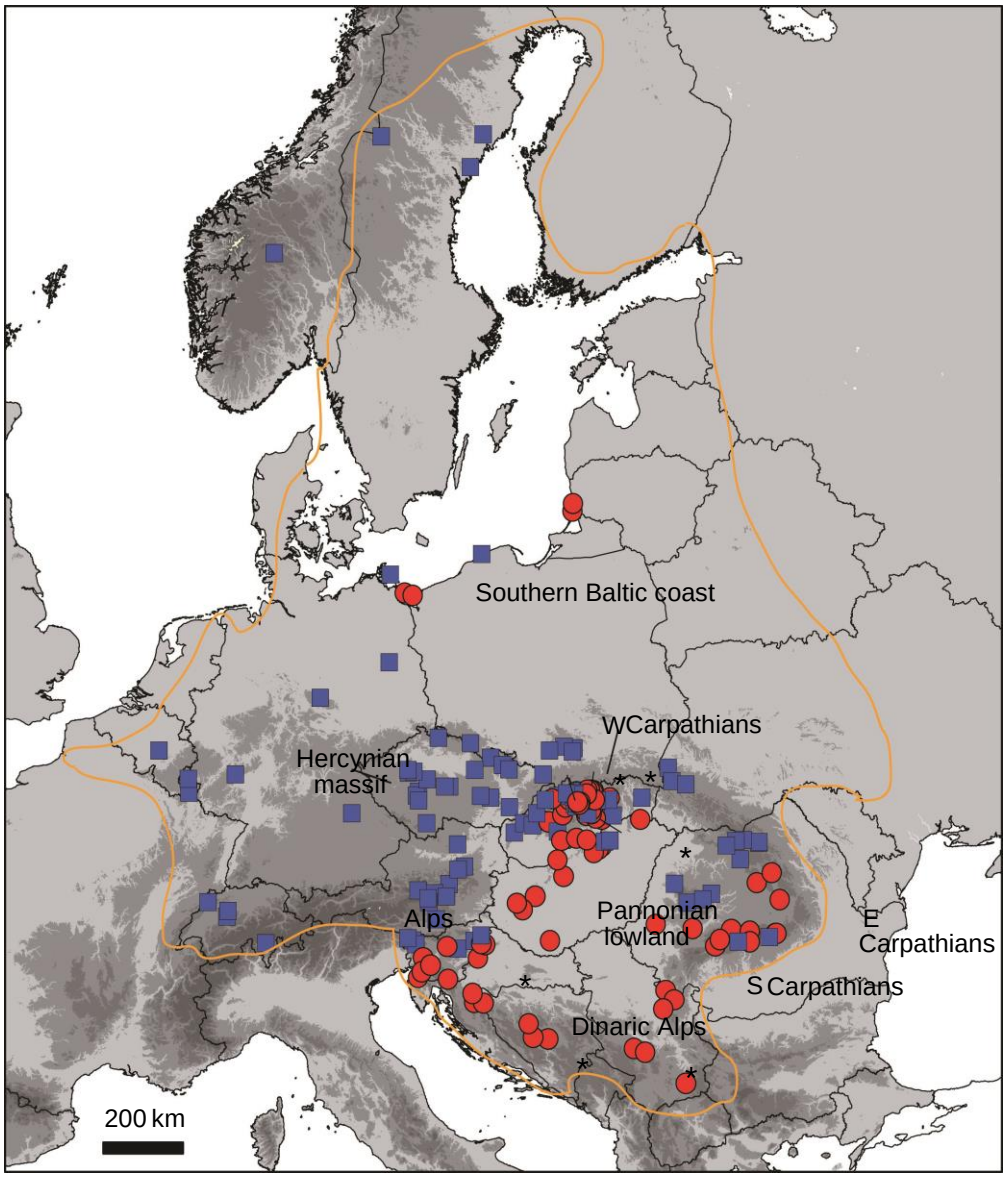


Figure 3. Distribution and ploidy level of the 194 studied populations of the *Arabidopsis arenosa* group in Europe (red – diploid, blue – tetraploid, asterisk – triploid, 2963 individuals investigated in total). The continuous distribution range of the whole species complex is marked by the orange outline (following Hoffmann, 2005).

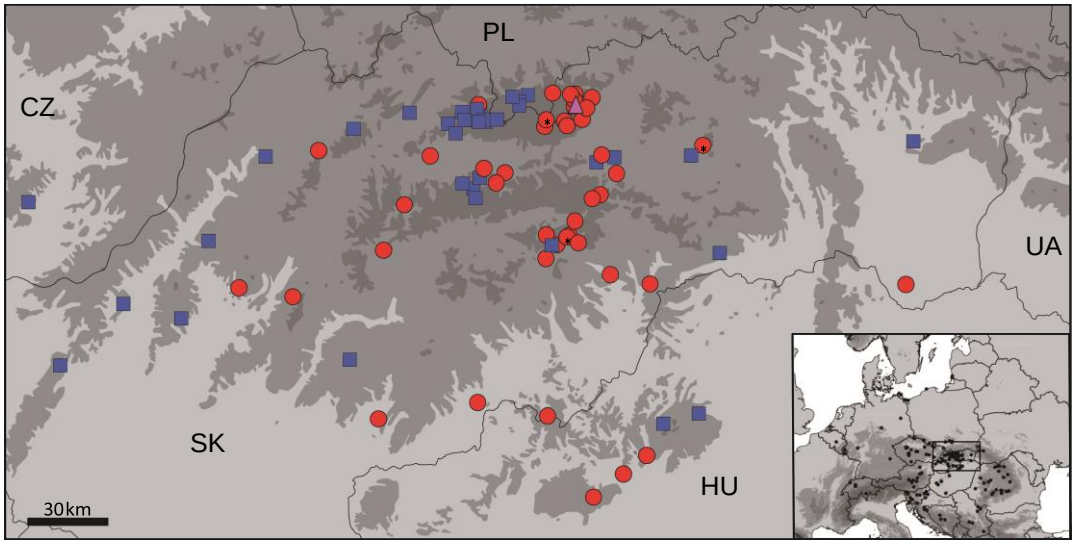


Figure 4. Detail of the contact zone of cytotypes of the *Arabidopsis arenosa* group in the Western Carpathians (red circle – diploid, blue square – tetraploid, pink triangle – mixed di-tetraploid population, asterisk – triploid; based on 1374 individuals from 79 populations).

Table 1. Summary of the final logistic generalized mixed effect models (GLMMs) testing the effects of spatial gradients (northing, easting) and the environmental correlates (total annual PAR, total annual precipitation; only those with the conditional effect significant at ⟨ = 5% are presented) on the distribution of diploid and tetraploid populations of the *Arabidopsis arenosa* group in the entire range of the group and in the Western Carpathian contact zone

	Whole model				Model parameters				
Data set	$\chi 2$	*P*	class (%)	Somers'Dxy (95% CL)	Predictor (unit)	B	SE	$\chi 2_{(1)}$	*P*
Whole data set	36.7	<0.001	71.2	0.59 (0.45, 0.71)	Easting (km)	–0.0036	0.0007	23.7	<0.001
					Total annual PAR (kWh.m−2)	–0.0196	0.0038	27.6	<0.001
					Total annual precipitation (mm)	–0.0014	0.0007	4.47	0.035
Contact zone	4.17	0.041	59	0.29 (0.05, 0.54)	Northing (km)	0.0123	0.0061	4.17	0.041

Characteristics of the final models: χ^2, test statistics; *P*, probabilities; class and Somers' Dxy, classification success. Characteristics of particular parameters: B, estimates of model coefficients; SE, standard error of estimates; $\chi^2_{(1)}$ and *P*, results of likelihood-ratio tests.

Table 2. Differences among diploid and tetraploid populations of the *Arabidopsis arenosa* group from the entire distribution area and from the contact zone in the Western Carpathians in the investigated habitat characteristics (significant results are in bold)

	2*x*	4*x*	Test
Anthropogenic stands+			
Whole range	74/14	36/70	χ^2 = **8.19, *P** = 0.003**
W Carpathians only	37/8	26/8	χ^2 = 0.40, *P** = 0.579
Anthropogenic spread†+			
Whole range	84/4	84/22	χ^2 = **10.88, *P** = 0.001**
W Carpathians only	42/2	34/0	χ^2 = 1.59, *P** = 0.510
Geology (calcareous vs. siliceous-neutral + volcanic)			
Whole range	46/37	46/35	χ^2 = 0.01, *P** = 1
W Carpathians only	25/17	23/11	χ^2 = 0.53, *P** = 0.490
Geology (calcareous vs. siliceous-neutral vs. volcanic)			
Whole range	46/25/12	45/31/4	χ^2 = 4.6, *P** = 0.111
W Carpathians only	25/9/8	23/8/3	χ^2 = 1.59, *P** = 0.505
Altitude			
Whole range	m (1–1950 m)	m (1–2269 m)	F(1,192) = 0.008, *P* = 0.927
W Carpathians only	m (161–1950 m)	m (251–2031 m)	F(1,77) = 1.14, *P* = 0.289

*The *P*-value was estimated using 2000 simulations.

†Only purely anthropogenic habitats far from any potential natural locality were considered as a distinct factor level in this analysis (see Methods).

+Number of positive and negative cases are before and after slash, respectively.

Discussion

We present the first large-scale evaluation of withinand among-population cytotype diversity of the *Arabidopsis arenosa* group, an important ploidyvariable species complex from the plant model genus.

Our study extends the knowledge of cytotype distribution across the range of this group particularly by: (i) expanding sampling efforts to mostly neglected regions (the Balkans, Carpathians, Baltic coast), (ii) a thorough sampling in the zone of spatial contact between cytotypes (Carpathians), and (iii) by a substantial increase of total sample size [over 2900 individuals in total, on average 15 per population in the current dataset vs. 730 and 273 individuals, on average five and two per population, in the previous surveys by Měsíček (1970) and Schmickl *et al*. (2012), respectively]. In addition, we present an overview of genome size variation within each ploidy level of the *A. arenosa* group.

Geography correlates with ploidy level distribution, but not with homoploid DNA content variation

Globally, the diploid and tetraploid populations of the *A. arenosa* group exhibit a parapatric distribution with two zones of cytotype spatial overlap, in the Slovenian Forealps and in the Carpathians. Although tetraploids were the prevailing cytotype, the diploid cytotype spans through more than one third of the total area, which is much larger than previously assumed (see

Fig. 3). In addition, the diploid populations are relatively common in some areas, occupying a variety of habitats, and in certain regions such as the Pannonian basin and the Dinaric Alps, they represent the only cytotype. This is in strong contrast to another di-tetraploid member of the genus in Central Europe, *Arabidopsis lyrata* subsp. *petraea*, which is represented by a few diploid populations isolated in cryptic Holocene refugia and by the locally more common hybridogenous tetraploid cytotype (Polatschek, 1966; Schmickl & Koch, 2011). The differentiation into diploid-dominated southern vs. tetraploiddominated northern (partly even formerly glaciated) regions represents a common cytogeographic pattern in the European flora that most likely reflects environmental changes during past climatic oscillations (Ehrendorfer, 1980; van Dijk & Bakx-Schotman, 1997; Weiss-Schneeweiss *et al.* 2013).

In addition, a previously unknown and ecologically distinct group of diploid populations has been found along the southern Baltic Sea coast, in the previously glaciated region at least 600 km from the closest diploid populations in the Western Carpathians. The recent introduction of these populations is not likely because the *A. arenosa* diploids generally do not show long-distance spreading in man-made habitats (see Table 2) and because the Baltic diploids exclusively occupy natural coastal sandy areas (searches in PR, PRC, W, and WU herbaria, plants from such habitats were found likely to occur from eastern Denmark to Estonia, F. Kolář, unpublished). Considering the large areas currently unfavourable for *A. arenosa* survival in northern Central Europe (forested or cultivated flatlands), a long-term isolation of the Baltic diploids from the main diploid range is probable, at least since the earlier phases of the Holocene. The presence of several geographically distinct and ecologically variable groups of diploid populations (at least two disjunct areas, with a wide range of habitats along a 0–2600 m altitudinal gradient) implies that their tetraploid derivatives, possibly combining several of the distinct diploid gene pools, should show considerable levels of genetic variation. Schmickl *et al*. (2012) did, indeed, detect large genetic variation among tetraploid populations of this group (even in the previously glaciated areas) and attributed it to the combined effects of several periglacial refugia, the absence of large bottlenecks and possible introgression from other sympatric *Arabidopsis* species. We hypothesise that recurrent origins of tetraploids from distinct gene pools and/or subsequent 2x → 4x introgression might have added another level of complexity to the *A. arenosa* group. In conclusion, the complicated cytogeographic pattern together with the most likely intricate internal sub-structuring of the species complex requires careful consideration in any ecological, genetic or genomic study employing taxa of the *A. arenosa* group as a model.

In addition to distinct ploidy levels, the plants studied here also exhibited a small but still considerable variation in genome size within each cytotype (up to 1.21-fold). The observed differences in DNA content might represent a combination of several causes of both biological and methodological origin. First, aneuploidy is usually responsible for larger abrupt differences in genome size (Roux *et al.* 2003; Šmarda & Bureš, 2006), and it also appears to be a plausible explanation for the exceptionally high DNA content values detected in both diploid and tetraploid accessions of the *A. arenosa* group (Fig. S1). Both aneuploidy and dysploidy is not rare in Brassicaceae and may be almost a rule in certain polyploid complexes such as those of the genus *Cardamine* [*Cardamine pratensis* group, Urbanska-Worytkiewicz & Landolt, (1974), Marhold (1994), Mandáková *et al*. (2013); or *C. yezoensis* and related taxa, Marhold *et al.* (2010)]. In addition, high levels of aneuploidy were also observed in karyological analyses of *Arabidopsis* seedlings (Měsíček, 1970; M. Kolník and K. Marhold, unpublished). Second, different intensity of genomic processes, such as non-coding repetitive DNA proliferation, unequal crossing-over and

illegitimate recombination, are considered major causes for gradual homoploid variation in DNA content within a species (Devos, Brown & Bennetzen, 2002; Bennetzen, Ma & Devos, 2005; Leitch & Leitch, 2013). Finally, methodological bias resulting from instrumental shifts and the influence of secondary metabolites could not be ruled out as we analysed different tissues (leaf, stem or petal) of plants that originated from ecologically distinct sites, collected in different parts of the season. Recent investigations have shown that, for instance, seasonal variation, choice of particular instrument or isolation buffer could result in up to 10% variation in fluorescence intensities (Bainard et al. 2011). However, we checked for artificial shifts by performing repeated analyses of the same accession on at least three different days, keeping the between-day variation below 3%. In addition, we demonstrated the genuine basis of the larger genome size differences by the presence of double peaks in simultaneous analyses of the individuals with distinct genome size values (which is considered to be the best evidence for true genome size differentiation, Greilhuber, 2005; Fig. 1D).

Small genuine differences in DNA contents are usually explained either as a result of neutral processes (random within- and across-population fluctuations or random accumulation of changes in spatially isolated areas/genetic lineages, Šmarda & Bureš, 2010; Oliver, McComb & Greene, 2013) or as an evolutionary constraint imposed by the surrounding environment and/or biological traits of the organism (e.g., rapid lifecycle and various traits relate to invasiveness; Greilhuber & Leitch, 2013). Our data favour the first, neutral scenario because we found no correlation of genome size in the entire *A. arenosa* group with any major geographic, altitudinal or environmental gradient. In contrast, a geographycorrelated > 10% variation in genome size has been recently found among Swedish genome-sequenced accessions of *A. thaliana*, but the selective background for such variation remains unconfirmed (Long et al. 2013).

High cytogeographical complexity in the Carpathians contrasts with intrapopulation cytotypes uniformity

Spatial relationships between cytotypes within species can be categorised as sympatric, parapatric or allopatric, depending on whether they are geographically intermixed, adjacent or disjunct, respectively. When polyploids first arise, they necessarily occur in sympatry with their diploid progenitors. Subsequent cytotype expansion or retreat results in parapatric or allopatric distributions. Two types of ploidy contact are recognised depending on their evolutionary history (Petit, Bretagnolle & Felber, 1999): (i) primary, when polyploids arise *de novo* from local diploids/ lower polyploids, and (ii) secondary, when different cytotypes regain contact after a phase of spatial separation. The *Arabidopsis arenosa* group most likely combines both scenarios at different spatio-temporal scales. The mixed diploid-triploid populations could be regarded as the primary cytotype contacts in which triploids originated recurrently via union of reduced (*n*) and unreduced (2*n*) gametes of the diploid. The alternative scenario, of triploid origin via inter-ploidal hybridization (favoured by Měsíček, 1970), seems improbable in light of our cytogeographic data. In all cases, only a single triploid plant was found in otherwise purely diploid populations; moreover, such populations were mostly found in exclusively diploidinhabited areas (e.g., in the Dinaric Alps). Since the advance of large-scale ploidy screening studies enhanced by flow cytometry, the occurrence of odd cytotypes within multiple ploidy species is more the rule than the exception (Husband, Baldwin & Suda, 2013), and rare (auto)triploids have been found even in otherwise purely diploid species (Slovák et al. 2009; Dušková et al. 2010). Our records represent

the first adult triploid individuals of *A. arenosa* detected in the wild. The extremely low frequency of adult triploids in our dataset (0.2%) in contrast with rather frequent triploid incidence in karyologically investigated seedlings (M. Kolník, unpublished) indicate strong yet still incomplete selection against the triploid progeny. Formation of viable triploid individuals in natural populations is an important prerequisite for incipient autopolyploid speciation (via triploid bridge, Husband, 2004) and thus shows important evolutionary potential for recurrent polyploidization within the *A. arenosa* group.

The two large areas of the diploid and tetraploid cytotype contact in the Carpathians and the Slovenian Forealps most likely represent secondary contact zones. This is indicated by the prevailing separate distribution of the cytotypes in the remaining areas and the intrapopulation cytotype uniformity (only one di-tetraploid population was found throughout the area studied). We will further discuss the origin and dynamics only of the sufficiently sampled zone in the Western Carpathians. This area hosts a complex landscape mosaic of spatially intermingled diploid and tetraploid populations that is in striking contrast with the within-population ploidy uniformity. Interestingly, both cytotypes occupy various substrates and climatic niches, and they occur from the lowland steppes up to high-alpine habitats. The absence of altitudinal differentiation is particularly interesting because it has been the only trend found repeatedly among the other investigated Carpathian taxa to date (Lysak & Doležel, 1998; Hodálová *et al*., 2007; Mráz *et al*., 2008). In addition, no general trend in cytotype-specific associations with geological substrates has been detected, although substrate specificity represents a major driver of plant spatial distributions and is also the principal speciation trigger among European mountain plants (Alvarez *et al*., 2009; Moore & Kadereit, 2013) as well as in *Arabidopsis* (Hunter & Bomblies, 2010; Schmickl & Koch, 2011). Collectively, we argue that ecological factors appear to play a minor role in the cytotype segregation; instead, random processes such as colonization history and genetic drift should be taken into account.

The marked prevalence of the cytotype-pure populations even within the Carpathian contact zone could be attributed to the demographic processes in the presumably strongly isolated populations. Both diploid and tetraploid populations of the *A. arenosa* group prefer open primary habitats with low competition, such as rocks, screes, sparse grasslands, and subalpine stands (Holocene cryptic refugia, Birks & Willis, 2008, see Table S1 for details on occupied habitats). In such sites isolated from each other, the processes of neutral evolution (random fluctuations in cytotype frequencies) complemented with frequencydependent selection against the rare cytotype (i.e., minority cytotype exclusion; Levin, 1975) could have occurred, ultimately leading to cytologically pure populations even from the hypothetical ploidy-mixed populations. Such a scenario involving dynamic changes in cytotype frequencies is supported by the short lifespan of the studied plants, which have no special adaptations for long-distance dispersal and very limited clonal growth and vegetative persistence (F. Kolář, M. Lučanová, personal observation). In contrast with *Arabidopsis*, other plant systems in the Western Carpathians investigated at comparable detail exhibit frequent within-population cytotype mixtures. Nevertheless, in both cases, the plants are long-living clonal perennials either with frequent asexual reproduction (*Pilosella officinarum*, Mráz *et al*., 2008) or preferring sites under strong human impact (*Phleum pratense* agg., Perný *et al*., 2008). However, another example of the almost complete absence of cytotype-mixed populations comes from the Brassicaceae family; although diploid, tetraploid and rare hexaploid populations of perennial *Alyssum montanum* are spatially intermingled on a large scale in Central Europe, they are cytotype uniform (Španiel *et al*., 2011, 2012).

It should be noted that other evolutionary processes such as recurrent *in situ* polyploidization and/or local adaptation may also have contributed to the observed pattern in certain areas, and further detailed molecular investigations are needed. For example, the spatially isolated occurrence of tetraploids (admixed in the only ploidy-mixed population AA170) among purely diploid populations suggests a local autopolyploid origin. In summary, current evidence suggests that areas with co-occurring diploid-tetraploid *A. arenosa* represent a rather stabilized secondary contact zone, at least on a coarse spatial scale.

Large niche overlap among cytotypes

Polyploidy can have a profound effect on various morphological, anatomical and physiological plant traits that further translate into distinct ecological requirements of cytotypes (reviewed in Levin, 2002). However, the general validity of shifts in climatic niche of diploids and their polyploid relatives has been recently questioned because no correlation was found in the majority of the thoroughly investigated closely related diploid–(auto)polyploid species groups (Glennon, Ritchie & Segraves, 2014). Our results further support the latter opinion because we found mostly no association or only a weak association between ploidy level and the environment in the *Arabidopsis arenosa* group. With the exception of higher tendency of tetraploids for spreading across anthropogenic stands, both cytotypes occur virtually along the entire range of habitats occupied by the species complex. Both cytotypes could be found on calcareous and acidic substrates, and both span from lowlands to alpine habitats. The climatic niche of the cytotypes is also largely similar, with the only differences caused by spatially correlated factors, reflecting the prevailing non-overlapping distribution ranges of the cytotypes. The absence of polyploidy-linked extension of realized climatic niches has previously been suggested for *Arabidopsis*, although dramatic changes in the realized climatic niche contributed to the evolution of the whole genus (Hoffmann, 2005). In addition, no traces of selection towards the ecological separation have been found: the levels of ecological differentiation were comparable in the areas where the cytotypes co-occur (Western Carpathians) and throughout the distribution area.

Nevertheless, it should be emphasised that our study focused on the *Arabidopsis arenosa* group as a whole, and some genetic lineages with distinct ecological and/or geographical associations may be found *within* each cytotype. For example, the ecologically and partly also morphologically distinct populations on railway tracks and other secondary habitats that prevail in northern Europe (but reach as far as southern Germany and Switzerland) might represent such distinct lineages, thus explaining the observed overall preference of tetraploids for anthropogenous stands.

See additional Supporting Information in the online version of this article at the publisher's web-site.

Acknowledgements

We are indebted to Jana Bayerová, Kateřina Černá, Ondřej Černý, Martin Hanzl, Jindřich Chrtek, Michael Jutzi, Klára Kabátová, Zuzana Khodlová, Eva Kolářová, Petr Koutecký, Jaromír Kučera, Radek Lučan, Lenka Macková, Jana Malinská, Pavol Mereďa, Clemens Pachschwöll, Jana Smatanová, Jan Suda, Milan Štech, Pavel Trávníček, Tomáš Urfus, and Judita Zozomová for help with sampling. The research was supported by the Czech Science Foundation (grant no.

P506/12/0668), the Slovak Research and Development Agency (APVV; grant no. APVV0139-12), and by the European Social Fund and the state budget of the Czech Republic (CZ.1.07/2.3.00/30.0022 to S.Š., CZ.1.07/2.3.00/30.0037 to T.M.). Part of the calculations were performed in the Computing Centre of the Slovak Academy of Sciences using the supercomputing infrastructure acquired in projects ITMS 26230120002 and 26210120002 (Slovak infrastructure for high-performance computing) supported by the Research & Development Operational Programme funded by the ERDF. This paper was presented at the interdisciplinary symposium 'Biogeography of the Carpathians: Evolution of Biodiversity in a Spatiotemporal Context' held in September 2013 in Kraków, Poland. The authors thank the organizers of the symposium for the invitation.

References

Alvarez N., Thiel-Egenter C., Tribsch A., Holderegger R., Manel S., Schönswetter P., Taberlet P., Brodbeck S., Gaudeul M., Gielly L., Küpfer P., Mansion G., Negrini R., Paun O., Pellecchia M., Rioux D., Schüpfer F., Van Loo M., Winkler M., Gugerli F., IntraBioDiv Consortium (2009): History or ecology? Substrate type as a major driver of spatial genetic structure in Alpine plants. – Ecology Letters 12: 632–640.

Bainard J. D., Husband B. C., Baldwin S. J., Fazekas A. J., Gregory T. R., Newmaster S. G., Kron P. (2011): The effects of rapid desiccation on estimates of plant genome size. – Chromosome Research 19: 825–842.

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

Birks H. J. B., Willis K. J. (2008): Alpines, trees, and refugia in Europe. – Plant Ecology & Diversity 1: 147–160.

Clauss M. J., Koch M. A. (2006): Poorly known relatives of *Arabidopsis thaliana*. – Trends in Plant Science 11: 449–459.

Devos K. M., Brown J. K. M., Bennetzen J. L. (2002): Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*. – Genome Research 12: 1075–1079.

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

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

Dormann C. F., McPherson J. M., Araújo M. B., Bivand R., Bolliger J., Carl G., Davies R. G., Hirzel A., Jetz W., Kissling W. D., Kühn I., Ohlemüller R., Peres-Neto P. R., Reineking B., Schröder B., Schurr F. M., Wilson R. (2007): Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. – Ecography 30: 609–628.

Dray S., Dufour A. B. (2007): The ade4 package: implementing the duality diagram for ecologists. – Journal of Statistical Software 22: 1–20.

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.

Dušková E., Kolář F., Sklenář P., Rauchová J., Kubešová M., Fér T., Suda J., Marhold K. (2010): Genome size correlates with growth form, habitat and phylogeny in the Andean genus *Lasiocephalus* (Asteraceae). – Preslia 82: 127–148.

Ehrendorfer F. (1980): Polyploidy and distribution. In: Lewis WH, ed. Polyploidy. Biological relevance. New York: Plenum Press, 45–60.

Fox J. (2003): Effect displays in R for generalised linear models. – Journal of Statistical Software 8: 1–27.

Galbraith D. W., Harkins K. R., Knapp S. (1991): Systemic endopolyploidy in *Arabidopsis thaliana*. – Plant Physiology 96: 985–989.

Glennon K. L., Ritchie M. E., Segraves K. A. (2014): Evidence for shared broad-scale climatic niches of diploid and polyploid plants. – Ecology Letters 17: 574–582.

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

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

Hijmans R. J., Cameron S. E., Parra J. L., Jones P. G., Jarvis A. (2005): Very high resolution interpolated climate surfaces for global land areas. – International Journal of Climatology 25: 1965–1978.

Hodálová I., Grulich V., Horová L., Valachovič M., Marhold K. (2007): Occurrence of tetraploid and octoploid cytotypes in *Senecio jacobaea* ssp. *jacobaea* (Asteraceae) in Pannonia and the Carpathians. – Botanical Journal of the Linnean Society 153: 231–242.

Hoffmann M. H. (2005): Evolution of the realized climatic niche in the genus *Arabidopsis* (Brassicaceae). – Evolution 59: 1425–1436.

Hollister J. D., Arnold B. J., Svedin E, Xue K. S., Dilkes B. P., Bomblies K. (2012): Genetic adaptation associated with genome-doubling in autotetraploid *Arabidopsis arenosa*. – PLoS Genetics 8: e1003093.

Hunter B., Bomblies K. (2010): Progress and promise in using *Arabidopsis* to study adaptation, divergence, and speciation. The Arabidopsis Book / American Society of Plant Biologists 8: e0138.

Husband B. C. (2004): The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. – Biological Journal of the Linnean Society 82: 537–546.

Husband B. C., Baldwin S. J., Suda P. J. (2013): The incidence of polyploidy in natural plant populations: major patterns and evolutionary processes. In: Greilhuber J., Doležel J., Wendel J. F., eds. Plant genome diversity, vol. 2. Vienna: Springer, 255–276.

Jakobsson M., Hagenblad J., Tavaré S., Säll T., Halldén C., Lind-Halldén C., Nordborg M. (2006): A unique recent origin of the allotetraploid species *Arabidopsis suecica*: evidence from nuclear DNA markers. – Molecular Biology and Evolution 23: 1217–1231.

Jørgensen M. H., Ehrich D.,Schmickl R., Koch M. A., Brysting A. K. (2011): Interspecific and interploidal gene flow in Central European *Arabidopsis* (Brassicaceae). – BMC Evolutionary Biology 11: 346.

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.

Krejčíková J., Sudová R., Lučanová M., Trávníček P., Urfus T., Vít P., Weiss-Schneeweiss H., Kolano B., Oberlander K., Dreyer L. L., Suda J. (2013): High ploidy diversity and distinct patterns of cytotype distribution in a widespread species of *Oxalis* in the Greater Cape Floristic Region. – Annals of Botany 111: 641–649.

Kron P., Suda J., Husband B. C. (2007): Applications of flow cytometry to evolutionary and population biology. – Annual Review of Ecology, Evolution, and Systematics 38: 847–876.

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

Levin D. A. (1975): Minority cytotype exclusion in local plant populations. – Taxon 24: 35–43.

Levin D. A. (2002): The role of chromosomal change in plant evolution. Oxford: Oxford University Press.

Long Q., Rabanal F. A., Meng D., Huber C. D., Farlow A., Platzer A., Zhang Q., Vilhjálmsson B. J., Korte A., Nizhynska V., Voronin V., Korte P., Sedman L., Mandáková T., Lysak M. A., Seren Ü., Hellmann I., Nordborg M. (2013): Massive genomic variation and strong selection in *Arabidopsis thaliana* lines from Sweden. – Nature Genetics 45: 884–890.

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

Mandáková T., Kovařík A., Zozomová-Lihová J., Shimizu-Inatsugi R., Shimizu K. K., Mummenhoff K., Marhold K., Lysak M. A. (2013): The more the merrier: recent hybridization and polyploidy in *Cardamine*. – The Plant Cell 25: 3280–3295.

Mandáková T., Marhold K., Lysak M. A. (2014): The widespread crucifer species *Cardamine flexuosa* is an allotetraploid with a conserved subgenomic structure. – New Phytologist 201: 982–992.

Marhold K. (1994): Chromosome numbers of the genus *Cardamine* L. (Cruciferae) in the Carpathians and Pannonia. – Phyton (Horn) 34: 19–41.

Marhold K., Kudoh H., Pak J. H., Watanabe K., Španiel S., Lihová J. (2010): Cytotype diversity and genome size variation in eastern Asian polyploid *Cardamine* (Brassicaceae) species. – Annals of Botany 105: 249–264.

Měsíček J. (1970): Chromosome counts in *Cardaminopsis arenosa* agg.(Cruciferae). – Preslia 42: 225–248.

Měsíček J. (1998): *Cardaminopsis*. In: Marhold K, Hindák F, eds. Zoznam nižších a vyšších rastlin Slovenska – checklist of non-vascular and vascular plants of Slovakia. Bratislava: Veda, 395–396.

Měsíček J., Goliašová K. (2002): *Cardaminopsis* (C. A. Mey.) Hayek. In: Goliašová K., Šípošová H., eds. Flóra Slovenska. Bratislava: Veda, 388–415.

Moore A. J., Kadereit J. W. (2013): The evolution of substrate differentiation in *Minuartia* series *Laricifoliae* (Caryophyllaceae) in the European Alps: in situ origin or repeated colonization? – American Journal of Botany 100: 2412–2425.

Mráz P., Šingliarová B., Urfus T., Krahulec F. (2008): Cytogeography of *Pilosella officinarum* (Compositae): altitudinal and longitudinal differences in ploidy level distribution in the Czech Republic and Slovakia and the general pattern in Europe. – Annals of Botany 101: 59–71.

Newson R. (2006): Confidence intervals for rank statistics: Somers' D and extensions. – Stata Journal 6: 309–334.

Oliver K. R., McComb J. A., Greene W. K. (2013): Transposable elements: powerful contributors to angiosperm evolution and diversity. – Genome Biology and Evolution 5: 1886– 1901.

Perný M., Kolarčik V., Majeský L., Mártonfi P. (2008): Cytogeography of the *Phleum pratense* group (Poaceae) in the Carpathians and Pannonia. – Botanical Journal of the Linnean Society 157: 475–485.

Petit C., Bretagnolle F., Felber F. (1999): Evolutionary consequences of diploid–polyploid hybrid zones in wild species. – Trends in Ecology & Evolution 14: 306–311.

Polatschek A. (1966): Cytotaxonomische Beiträge zur Flora der Ostalpenländer, I. Österreichischen Botanischen Zeitschrift 113: 1–46.

R Development Core Team. (2013): R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Roux N., Toloza A., Radecki Z., Zapata-Arias F. J., Doležel J. 2003. Rapid detection of aneuploidy in *Musa* using flow cytometry. – Plant Cell Reports 21: 483–490.

Schmickl R., Koch M. A. 2011. *Arabidopsis* hybrid speciation processes. – Proceedings of the National Academy of Sciences USA 108: 14192–14197.

Schmickl R., Paule J., Klein J., Marhold K., Koch M. A. 2012. The evolutionary history of the *Arabidopsis arenosa* complex: diverse tetraploids mask the western Carpathian center of species and genetic diversity. – PLoS ONE 7: e42691.

Shimizu-Inatsugi R., Lihová J., Iwanaga H., Kudoh H., Marhold K., Savolainen O., Watanabe K., Yakubov V. V., Shimizu K. K. 2009. The allopolyploid *Arabidopsis kamchatica* originated from multiple individuals of *Arabidopsis lyrata* and *Arabidopsis halleri*. – Molecular Ecology 18: 4024–4048.

Slovák M., Vít P., Urfus T., Suda J. 2009. Complex pattern of genome size variation in a polymorphic member of the Asteraceae. – Journal of Biogeography 36: 372–384.

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.

Španiel S., Marhold K., Filová B., Zozomová-Lihová J. 2011. Genetic and morphological variation in the diploid-polyploid *Alyssum montanum* in Central Europe: taxonomic and evolutionary considerations. – Plant Systematics and Evolution 294: 1–25.

Španiel S., Marhold K., Thiv M., Zozomová-Lihová J. 2012. A new circumscription of *Alyssum montanum* ssp. *montanum* and *A. montanum* ssp. *gmelinii* (Brassicaceae) in Central Europe: molecular and morphological evidence. – Botanical Journal of the Linnean Society 169: 378–402.

Suda J., Krahulcová A., Trávníček P., Krahulec F. 2006. Ploidy level versus DNA ploidy level: an appeal for consistent terminology. – Taxon 55: 447–450.

Temsch E. M., Greilhuber J., Krisai R. 2010. Genome size in liverworts. – Preslia 82: 63–80.

Trávníček P., Dočkalová Z., Rosenbaumová R., Kubátová B., Szeląg Z., Chrtek J. 2011b. Bridging global and microregional scales: ploidy distribution in *Pilosella echioides* (Asteraceae) in central Europe. – Annals of Botany 107: 443–454.

Trávníček P., Eliášová A., Suda J. 2010. The distribution of cytotypes of *Vicia cracca* in Central Europe: the changes that have occurred over the last four decades. – Preslia 82: 149–163.

Trávníček P., Kubátová B., Čurn V., Rauchová J., Krajníková E., Jersáková J., Suda J. 2011a. Remarkable coexistence of multiple cytotypes of the *Gymnadenia conopsea* aggregate (the fragrant orchid): evidence from flow cytometry. – Annals of Botany 107: 77–87.

Urbanska-Worytkiewicz K.,Landolt E. 1974. Biosystematic investigations in *Cardamine pratensis* L. s.l. 1. Diploid taxa from Central Europe and their fertility relationships. – Berichte des Geobotanischen Institutes der Eidgenössischen Technischen Hochschule, Stiftung Rübel 42: 42–139.

Van Dijk P., Bakx-Schotman T. 1997. Chloroplast DNA phylogeography and cytotype geography in autopolyploid *Plantago media*. – Molecular Ecology 6: 345–352.

Venables W. N., Ripley B. D. 2002. Modern applied statistics with S, 4th edn. New York: Springer.

Weiss-Schneeweiss H., Emadzade K., Jang T. S., Schneeweiss G. M. 2013. Evolutionary consequences, constraints and potential of polyploidy in plants. – Cytogenetic and Genome Research 140: 137–150.

Yant L., Hollister J. D., Wright K. M., Arnold B. J., Higgins J. D., Franklin F. C. H., Bomblies K. 2013. Meiotic adaptation to genome duplication in *Arabidopsis arenosa*. – Current Biology 23: 2151–2156.

Case study IV

Dušková E., Kolář F., Sklenář P., Rauchová J., Kubešová M., Fér T., Suda J. & Marhold K. (2010): **Genome size correlates with growth form, habitat and phylogeny in the Andean genus *Lasiocephalus* (Asteraceae). – Preslia 82: 127-148**

Genome size correlates with growth form, habitat and phylogeny in the Andean genus *Lasiocephalus* (Asteraceae)

Eva Dušková[1], Filip Kolář[1,2], Petr Sklenář[1], Jana Rauchová[2,1], Magdalena Kubešová[1,2], Tomáš Fér[1], Jan Suda[1,2] & Karol Marhold[3,1]

[1] Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, CZ-128 01 Prague, Czech Republic, e-mail: jezanek@centrum.cz, filip.kolar@gmail.com, petr@natur.cuni.cz, kubesovm@gmail.com, tomas.fer@centrum.cz, suda@natur.cuni.cz;
[2] Institute of Botany, Academy of Sciences of the Czech Republic, Zámek 1, CZ-252 43 Průhonice, Czech Republic, e-mail: jrauchova@post.cz;
[3] Institute of Botany, Slovak Academy of Sciences, Dúbravská cesta 9, SK-845 23 Bratislava, Slovak Republic, e-mail: Karol.Marhold@savba.sk

Abstract

Variation in genome size in a particular taxonomic group can reflect different evolutionary processes including polyploidy, hybridization and natural selection but also neutral evolution. Using flow cytometry, karyology, ITS sequencingand field surveys, the causes of variation in genomesize in the ecologically and morphologically diverse high-Andean genus *Lasiocephalus* (*Asteraceae*, *Senecioneae*) were examined. There was a 1.64-fold variation in holoploid genome size (C-values) among 189 samples belonging to 20 taxa. The most distinct was a group of plants with large genomes corresponding to DNA triploids. Disregarding the DNA triploids, the remaining samples exhibited a pronounced (up to 1.32-fold) and rather continuous variation. Plants with the smallest genomes most likely represent intergeneric hybrids with the closely related and sympatric *Culcitium nivale,* which has a smaller genome than *Lasiocephalus*. The variation in genome size in samples of diploid *Lasiocephalus* was strongly correlated with several environmental and life history traits (altitude, habitat and growth form). However, all these factors, as well as genome size itself, were correlated with phylogeny (main split into the so-called 'forest' and 'páramo' clades), which most probably represents the true cause of the differentiation in intrageneric genome size. In contrast, relationships between genome size and phylogeny were not apparent at lower divergence levels. Instead, here we suggest that ecological conditions have played a role in driving shifts in genome size between closely related species inhabiting different environments. Collectively, this study demonstrates that various evolutionary forces and processes have shaped the variation in genome size and indicates that there is a need for multi-approach analyses when searching for the causes and consequences of changes in genome size.

Keywords: Andes, *Compositae*, flow cytometry, genome size, growth form, habitat preference, hybridization, *Lasiocephalus*, nuclear DNA amount, nucleotypic effect, phylogeny, polyploidy

Introduction

One of the current challenging questions in plant evolutionary biology is the functional significance of the approximately 2000-fold variation in genome size among angiosperms (Leitch & Bennett 2007). This tremendous variation is largely caused by different proportions of non-coding repetitive DNA, such as transposable elements, satellite DNA, introns, and pseudogenes (Bennett & Leitch 2005). Several mechanisms, both at the chromosomal and molecular levels, are thought to be capable of generating genome size changes (Bennetzen 2005). Polyploidy, chromosome gain, presence of supernumerary or sex chromosomes, amplification of retrotransposons and genomic duplications are included among the processes that cause genome expansion. Counterbalancing mechanisms involve the loss of entire chromosomes, illegitimate recombination, unequal intrastrand homologous recombination or a higher overall rate of nucleotide deletion over insertion (reviewed in Bennett & Leitch 2005). While chromosome-level events typically cause sudden and marked changes in genome size, molecular mechanisms are more gradual with each event producing only slight modifications in genome size.

Variation in genome size has been interpreted in relation to various phenotypic and/or life-history traits. Classical observations documenting a positive relationship between genome size and cell size and duration of cell division (reviewed by Leitch & Bennett 2007) were followed by studies that revealed correlations between genome size and seed mass, leaf mass per unit area, growth rate and/or photosynthetic rate (Knight & Beaulieu 2008). In addition, genome size can also have significant consequences at the ecological level, influencing the range of environmental conditions a plant can tolerate (the so-called "large genome constraint hypothesis"; Knight et al. 2005). Over the years, relationships between genome size and temperature, water availability, latitude and altitude of the habitats have been reported (reviewed by Knight & Ackerly 2002). However, variation in genome size may not always be adaptive as documented for instance in a highly variable population of *Festuca pallens* (Šmarda et al. 2007).

Recent studies have also emphasized the importance of phylogeny in variation in genome size. For example, genome size in *Hieracium* subg. *Hieracium* was strongly correlated with the two major phylogenetic groups (Chrtek et al. 2009) and congruency between genome size and evolutionary lineages is also found in the holoparasitic *Orobanche* (Weiss-Schneeweiss et al. 2005). If phylogeny is not considered, the role of other factors (e.g. environmental conditions) in shaping genome size may be overemphasized. Therefore, when searching for the causes of variation in genome size, both adaptive and non-adaptive components need to be taken into account.

Knowledge of the patterns and dynamics of variation in genome size is largely based on the analysis of European and North American floras, and species from the world's main centres of plant diversity are neglected. One of the geographic regions that hosts an extremely-rich plant biota is the Andes of South America, which rank among the most conspicuous physiographic features on Earth (Vuilleumier & Monasterio 1986, Luteyn 1999, Young et al. 2002). Andean plant diversity largely resulted from the adaptive radiation of numerous plant groups, such as *Espeletia*, *Huperzia*, *Lupinus* and *Valeriana* (Cuatrecasas 1986, Wilkström et al. 1999, Bell & Donoghue 2005, Hughes & Eastwood 2006). Most researchers relate the adaptive radiation to the final uplift of the mountains and emergence of high-altitude non-forest habitats (páramo) about 5–3 mya and the PlioPleistocene climatic oscillations. Although there is an increasing number of phylogenetic studies on Andean plants (e.g. Young et al. 2002, Struwe et al. 2009) the role of genomewide

evolutionary processes (such as polyploidization and genome size change) in species radiation is largely unknown.

Lasiocephalus Willd. ex Schltdl. (*Asteraceae*, *Senecioneae*) is a neotropical highAndean genus of ca 30 species distributed from Venezuela to Bolivia, with the highest species richness in Ecuador (Cuatrecasas 1978). Two main growth forms are recognized in the genus: (i) the broad-leaved suffrutescent climbers of montane forests and the tree-line ecoton, and (ii) erect and ascending narrow-leaved herbaceous plants to subshrubs of the high-altitude páramo. These growth forms plausibly reflect adaptive changes that occurred during the colonization of the páramo habitats (Cuatrecasas 1978). It is likely that *Lasiocephalus* is a monophyletic group nested within *Senecio* s. str. with the high-Andean genus *Culcitium* as a sister taxon (Pelser et al. 2007). The high number of species along with the variation in growth forms and their altitudinal distribution make *Lasiocephalus* a suitable candidate for studying the evolutionary outcomes associated with adaptive radiation.

We use *Lasiocephalus* as a representative example of a group that has radiated in the Andes to address general questions concerning the extent and possible sources of the variation in DNA content. Flow cytometry, karyology and ITS sequencing were employed to address the following specific questions: (i) Does genome duplication play a role in the diversification of *Lasiocephalus*? Which cytotypes are found in this genus? Are there any mixed-ploidy populations? (ii) What is the variation in genome size (C- and Cx-values)? (iii) How does the observed pattern in DNA ploidy and variation in genome size correlate with growth forms and habitat preferences? (iv) How does the observed pattern in DNA ploidy and variation in genome size correlate with the phylogeny of the group?

Materials and methods

Plant material

Plants were collected between 2007 and 2009 in Bolivia, Ecuador and Venezuela (see Electronic Appendix 1 for locality details). In total, 189 individuals from 101 populations corresponding to 20 species of *Lasiocephalus* sensu Cuatrecasas (1978) were sampled (Table 1). Species were identified according to Cuatrecasas (1978) and by comparison with herbarium vouchers deposited at AAU, COL, G, MERF, NY, P, PRC, QCA, S and VEN. Seven distinct morphotypes of *Lasiocephalus*, which could not be assigned to any hitherto described species, were treated as separate morphospecies and referred to as *Lasiocephalus* sp. 1 to 7 (Table 1). GPS coordinates and altitude (WGS 84 system), habitat type and growth form of the plants were recorded at each locality. Young intact leaves were collected and quickly desiccated for DNA analyses using silica gel. The following plant material was sampled for flow cytometry: (i) fresh tissue (stems and leaves) of adult plants stored in plastic bags at 58 °C until analyses (up to 14 days), (ii) silica-dried leaf tissue of adult plants, and (iii) achenes from which seedlings were raised. In addition, three samples of closely related and often sympatric *Culcitium nivale* were also collected and analyzed for comparative purposes. Herbarium vouchers are deposited in PRC, QCA, QCNE and VEN.

Table 1. – Nuclear DNA contents of 13 species and seven unidentified morphospecies of *Lasiocephalus* and the closely related *Culcitium nivale*. Samples analysed = no. of populations/no. of individuals. Values are mean±SD. *Vicia faba* 'Inovec' (2C = 26.9 pg) served as an internal reference standard. In the Grouping columns, different letters indicate groups of taxa that are significantly

Species	Genome size ('fresh' dataset)		
	Samples analysed	2C-value (pg of DNA)	Grouping
Lasiocephalus campanulatus (Sch. Bip. ex Klatt) Cuatrec.	–	–	–
Lasiocephalus cuencanus (Hieron.) Cuatrec.	–	–	–
Lasiocephalus decipiens (Benoist) Cuatrec.	2.2	16.10±0.08	ef
Lasiocephalus heterophyllus (Turcz.) Cuatrec.	–	–	–
Lasiocephalus involucratus (Kunth) Cuatrec.	5.5	15.63±0.10	def
Lasiocephalus lingulatus Schltdl. – 2x	3.3	15.07±0.15	bc
Lasiocephalus lingulatus Schltdl. – 3x	1.1	22.69	–
Lasiocephaluslongipenicillatus(Sch.Bip.exSandw.)Cuatrec.	–	–	–
Lasiocephalus mojandensis (Hieron.) Cuatrec.	–	–	–
Lasiocephalus otophorus (Wedd.) Cuatrec.	3.3	16.10±0.47	f
Lasiocephalus ovatus Schltdl. – 2x	3.3	14.99±0.04	bc
Lasiocephalus ovatus Schltdl. – DNA 3x	1.1	22.45	–
Lasiocephalus patens (Kunth) Cuatrec.	5.6	17.27±0.15	g
Lasiocephalus pichinchensis (Cuatrec.) Cuatrec.	1.1	15.2	bcd
Lasiocephalus sodiroi (Hieron.) Cuatrec.	1.1	14.58	ab
Lasiocephalus L. sp. 1 (liana)	4.6	15.59±0.08	de
Lasiocephalus L. sp. 2 (shrub)	1.2	17.27±0.42	g
Lasiocephalus L. sp. 3 (shrubby liana)	1.1	15.70	cdef
Lasiocephalus L. sp. 4 (cf. *L. ledifolius* (Kunth) C. Jeffrey)	2.2	13.82±0.10	a
Lasiocephalus L. sp. 5 (*L. involucratus*-like)	–	–	–
Lasiocephalus L. sp. 6 (*L. sodiroi*-like) – 2x	–	–	–
Lasiocephalus L. sp. 6 (*L. sodiroi*-like) – DNA 3x	–	–	–
Lasiocephalus L. sp. 7 (*L. sodiroi*-hairy)	–	–	–
Culcitium nivale Kunth	3.3	12.35±0.10	–

different at α = 0.05 using Tukey's HSD multiple comparison test. DNA-triploids were excluded from statistical comparison. ITS clade: CN = '*C. nivale*' clade; F = 'forest' clade; P = 'paramo' clade. Growth form: BH = broad-leaved herb; L = liana; NH = narrow-leaved herb; S = shrub. Habitat type: F = Andean forest; G = grass paramo; SB = subparamo; SP = superparamo.

Relative DNA content ('silica' dataset)			ITS clade	Growth form	Habitat type
Samples analysed	Sample/ standard ratio	Grouping			
2.2	0.667±0.000	fghi	F	L	F
2.4	0.705±0.013	i	F	L	F
1.1	0.646	cefghi	–	L	F
2.8	0.665±0.012	gh	F	L	F
9.18	0.628±0.019	ef	F	L	SB
6.15	0.620±0.018	de	P	NH	G
2.2	0.855±0.031	–	P	NH	G
4.7	0.661±0.021	g	F	NH	SP
2.6	0.617±0.026	ce	–	BH	G
1.1	0.671	efghi	F	L	SB
7.19	0.596±0.011	bc	P	NH	SP
3.4	0.860±0.057	–	P	NH	SP
4.14	0.688±0.023	hi	F	L	F
10.21	0.612±0.016	ce	F + P	NH	G
2.3	0.599±0.003	bcd	P	NH	SP
4.5	0.622±0.011	cef	–	L	SB
–	–	–	–	S	SB
1.2	0.602±0.012	bce	–	L	SB
4.6	0.535±0.007	a	CN	NH	SP
1.6	0.626±0.010	def	F	L	SB
1.2	0.556±0.024	ab	–	NH	G
1.2	0.771±0.009	–	–	NH	G
1.4	0.545±0.007	a	–	NH	SP
1.1	0.494	–	CN	NH	SP

Flow cytometry

DNA ploidy levels (Suda et al. 2006) were determined using flow cytometry and a Partec PA II instrument (Partec GmbH) equipped with a HBO mercury arc lamp. Genome sizes (Cand Cx-values; Greilhuber et al. 2005) were determined using CyFlow SL instrument (Partec GmbH, Münster, Germany) equipped with a green solid-state laser (Cobolt Samba, 532 nm, 100 mW). Sample preparation generally followed the simplified two-step procedure using Otto buffers (Doležel et al. 2007). Briefly, intact leaf tissue of the *Lasiocephalus* plant to be analyzed and an internal reference standard (*Vicia faba* 'Inovec', 2C = 26.9 pg; Doležel et al. 1998) were mixed together and chopped with a sharp razor blade in a Petri dish containing 0.5 ml of ice-cold Otto I buffer (0.1 M citric acid, 0.5% Tween 20; Otto 1990). The suspension was filtered through a 42-μm nylon mesh and incubated for approx. 30 min at room temperature. The staining solution consisted of 1 ml of Otto II buffer (0.4 M $Na_2HPO_4 \cdot 12\ H_2O$), β-mercaptoethanol (final concentration of 2 μl/ml) and a fluorochrome. Propidium iodide (PI) and RNase IIA (both atfinalconcentrations of 50 μg/ml) were used to determine the genome size in absolute values (pg of DNA) in fresh samples, while AT-selective DAPI (4',6-diamidino-2-phenylindole) at a final concentration of 4 μg/ml was used for estimating the DNA content in relative units in silica-dried samples. Samples were stained for 10 min at room temperature and run on the flow cytometer. Isolated nuclei were excited either with a laser (for PI staining) or a mercury arc lamp (for DAPI staining) and the fluorescence intensity of 5000 particles recorded.

Two distinct datasets were produced and treated separately in subsequent statistical analyses. Firstly, relative DNA content was determined for 152 silica-dried *Lasiocephalus* and one *Culcitium nivale* sample (further referred to as the 'silica' dataset) using DAPI flow cytometry. The desiccated samples were not more than 18 months old when analyzed. DAPI was chosen because it is less sensitive to secondary metabolites and chromatin arrangement, which allowed a high-resolution analysis even of dehydrated plant tissues (Suda & Trávníček 2006). Only histograms with coefficients of variation (CVs) of the G_0/G_1 peak of the analyzed sample below 5% were considered. Secondly, absolute values of genome size of a subset of 37 fresh accessions of *Lasiocephalus* and three accessions of *Culcitium nivale* ('fresh' dataset) were determined using PI flow cytometry. More stringent criteria for the quality of the flow analysis were applied in this case: (i) peaks of both the sample and the internal standard of approximately the same height, (ii) CV of G_0/G_1 peak of the analyzed sample below 4%, (iii) three replicates of each sample on different days to minimize potential random instrumental drift, and (iv) the between-day variation in fluorescence intensity below 3%; otherwise, the most extreme value was discarded and the sample re-analyzed. The reliability of FCM measurements (i.e. between-plant differences) was repeatedly confirmed by simultaneous analyses of *Lasiocephalus* samples with distinct genome sizes (Fig. 1).

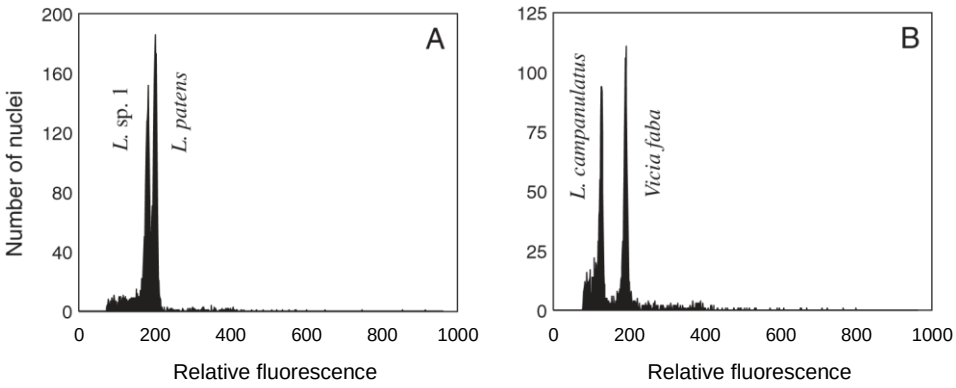


Fig. 1. – Fluorescence histograms for *Lasiocephalus* samples analyzed using flow cytometry. (A) Simultaneous analysis of propidium iodide stained nuclei isolated from fresh tissue of *L.* sp. 1 (2C = 15.59 pg, pop. no. P2) and *L.patens* (2C= 17.27 pg, pop. no. P12),which reveal interspecific differences in genome size. (B) Analysis of relative DNA content of silica gel preserved samples of *L. campanulatus* (pop. no. C1A). Isolated nuclei were stained with DAPI and fresh leaf tissue of *Vicia faba* was used as an internal reference standard. See Electronic Appendix 1 for population details.

The presence of two distinct peaks in analyses of co-processed samples is considered the most convincing evidence for genuine differences in nuclear DNA content (Greilhuber 2005).

To assess relationships between DAPI (AT-selective fluorochrome) and PI (intercalating fluorochrome without base bias) staining, a subset of thirteen fresh DNA diploid *Lasiocephalus* accessions covering the whole range of monoploid genome sizes was analyzed using both PI and DAPI flow cytometry concurrently. The analyses followed the above-defined stringent criteria with three independent replicates for each sample and each stain.

Chromosome counts

To confirm the FCM results, three individuals covering the whole genome size range at the diploid level (Table 1, Electronic Appendix 1) were subjected to conventional karyological counts using rapid squash methods. The apical root meristem of young seedlings was pre-treated with p-dichlorbenzene for 3 h, fixed in ice-cold 3 : 1 ethanol : acetic acid for 12–14 h, macerated for 1 min in 1 : 1 ethanol : hydrochloric acid, and stained with lacto-propionic-orcein (Dyer 1963).

ITS cloning and sequencing

Sequences of the nuclear ribosomal internal transcribed spacer (ITS) region were analyzed for 13 *Lasiocephalus* taxa (Table 1). Total genomic DNA was extracted from silicadried material using the Invisorb Spin Plant Mini Kit (Invitek) according to the manufacturer's instructions. The ITS1-5.8S rDNA-ITS2 region of the nuclear ribosomal DNA was amplified using the primers ITS5 (White et al. 1990) and ITS_Lasio (5'AGTCGRAGCATCGTCATGAGA-3'). PCR amplifications were done in 22 µl reaction containing 0.18 mM of each dNTP (Fermentas), 0.23 mM of each primer (Sigma), 0.5 units of JumpStart REDTaq polymerase (Sigma), 1 × PCR buffer for JumpStart REDTaq (Sigma) and 5 ng of genomic DNA. An initial denaturation step at 94 °C for 1 min was followed by 35 cycles of denaturation (94 °C for 50 s), annealing (53 °C for 50 s) and extension (72 °C for 1 min) steps and a final extension at 72 °C for 10 min. Each sample was

amplified in two separate reactions. Both PCR products were mixed and purified using JETQUICK PCR Product Purification Spin Kit (Genomed). The purified ITS templates were ligatedto the pGEM-T Easy vector (Promega) following the manufacturer's protocol but downscaled to half volume reactions and subsequently transformed into competent cells. After an overnight culture at 37 °C on the LB ampicillin/IPTG/X-gal selective plates, colonies carrying the ITS insert were identified by colour. 10–12 colonies from each reaction were selected for colony PCR using the same conditions as in the first PCR reaction. PCR products were purified using JETQUICK PCR Product Purification Spin Kit (Genomed) and subsequently sequenced (Macrogen, lll) using the ITS5 primer.

Phylogenetic and statistical analyses

Chromatograms of the cloned sequences were inspected by eye using Finch TV sequence viewer. In addition, direct GenBank sequences of nine species identified as putative closest relatives of *Lasiocephalus* by Pelser et al. (2007) were included as an outgroup reference. Initial sequence alignment was done with ClustalW algorithm incorporated in BioEdit v5.0.6 (Hall 1999) and manually improved. The same program was used for further manual editing. Autapomorphies that were found only in one clone within the whole dataset were considered to be polymerase errors and discarded (Popp et al. 2005). Two sequences of a highly ribotype-variable species *L. pichinchensis*, which combined apomorphies of the two main clades (see Results), were identified as chimeras and therefore omitted from analyses. Sequences were submitted to GenBank (GU289931 – GU290036). Bayesian analysis (MrBayes v3.1.2; Huelsenbeck & Ronquist 2001, Ronquist & Huelsenbeck 2003) was run on the complete dataset with GTR model of nucleotide substitutions, gamma model of rate heterogeneity and variable proportion of invariable sites. This model was found in hierarchical likelihood ratio tests as the model of molecular evolution best fitting the data as implemented in Modeltest v3.5 (Posada & Crandall 1998). Two parallel runs with four chains each were used, sampling every 1000th tree for 3 million generations until convergence. The first 750 trees per run were discarded as burn-in and the remaining 2251 trees per run were summarized.

Correlation between DAPI and PI measurements of the same accessions was examined in Statistica 8. Differences in nuclear DNA contents were tested using an analysis of variance (ANOVA) and further assessed by Tukey's HSD multiple comparisons incorporated in the R software 2.9.2 (R Development Core Team 2009). The differences between growth forms and preferred habitat types were assessed by ANOVA with species identity considered as a factor with random effect (nested in the particular factor) in order to avoid pseudoreplication. The relationship between DNA content and altitude was tested by linear mixed-effect model (species identity treated as a factor with random effect) incorporated in the *nlme* package in R (Lindstrom & Bates1990). In all analyses the random effect was modelled only for the intercept, as was suggested by the likelihood ratio test (only non-significant improvement when a more complex model taking into account also the effect of the random factor for the slope was considered).

The relationships between genome size and phylogeny were examined by testing the association of nuclear DNA content with the two main clades detected (i.e. the 'páramo' and the 'forest' clade; see Results). Due to non-homogeneity of variances (indicated by highly significant Bartlett test: $P < 0.002$ in both datasets), differences in DNA content (mean values per species)

between clades were tested by the non-parametric Mann-Whitney test implemented in Statistica 8 (StatSoft 2007). Species for which there was no sequence information were omitted from this kind of analysis. Within the 'forest' clade, differences in relative genome size (the 'silica' dataset) among the three main subclades (Table 1) were tested using the Kruskal-Wallis test implemented in Statistica 8. The relationship between DNA content and altitude was assessed separately for each of the two main clades (Table 1) using the same procedure as in the whole dataset (see above).
In order to quantify the correlation between environmental factors (habitat and growth form) and phylogeny (split into two main clades), variation partitioning based on three-way ANOVA with sequential (Type I) sum of squares was done in Statistica 8. Mean DNA content values of each species were subjected to six three-way ANOVAs with different input sequence of factors and model sums of squares used as variance component measures.

Results

DAPI versus PI staining

A highly linear DAPI/PI correlation was observed across the whole range of genome sizes of the 13 selected fresh diploid *Lasiocephalus* samples (the relationship was: DAPI fluorescence = 1.126 × PI fluorescence, $R^2 = 0.99$, $P < 0.001$). This indicates a constant AT/GC content within the genus and enables a reliable conversion and interpretation of the estimated values.

Chromosome counts

Karyological analyses of three putatively diploid plants with markedly different genome sizes [*Lasiocephalus* sp. 4 (cf. *L. ledifolius*)], pop. G18, 2C = 13.82 pg; *Lasiocephalus* sp. 3, pop. no. 11567, 2C = 15.70 pg; and shrubby *Lasiocephalus* sp. 2, pop. no. 11584, 2C = 17.27 pg) confirmed the diploid number of chromosomes (2n = 2x = 40) in all samples (Fig. 2).

DNA ploidy levels

The analysis of 37 fresh and 152 silica-dried *Lasiocephalus* samples revealed a large variation in nuclear DNA content within this genus, spanning 1.64-fold range (2C = 13.82–22.69 pg) and 1.60-fold range (sample/*Vicia* ratio = 0.535–0.860), respectively (Table 1). In each dataset, two well-defined genome size groups (separated by a distinct gap) were detected. The fluorescence values of the two groups differed by ~1.5-fold, indicating that DNA diploids and DNA triploids were involved (Table 1). DNA triploids were represented by ten samples belonging to three species (*L. lingulatus*, *L. ovatus* and *L.* sp. 6) and occurred in eight populations (~8% of all sampled populations). In most cases, DNA triploids grew in sympatry with more abundant diploid plants.

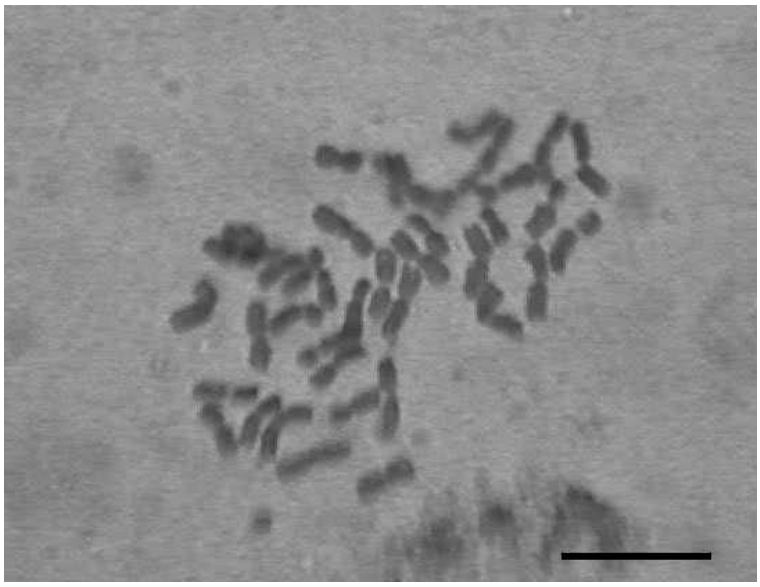


Fig. 2. – Mitotic chromosomes of *Lasiocephalus* sp. 3 (shrubby liana, pop. no. 11567), 2n = 40. Scale bar = 10 μm (photo V. Jarolímová).

Genome size variation

Disregarding the DNA triploids, the remaining *Lasiocephalus* accessions still exhibit a pronounced variation in genome size. Mean 2C-values of fresh accessions ranged from 13.82 pg in *Lasiocephalus* sp. 4 (cf. *L. ledifolius*) to 17.27 pg in shrubby *Lasiocephalus* sp. 2 and *L. patens*, (1.25-fold variation). The variation in silica-dried samples was 1.32-fold, with *L.* sp. 4 (cf. *L. ledifolius*) possessing the smallest and *L. cuencanus* the largest genomes (Table 1, Fig. 3). The closely related *Culcitium nivale* possessed a smaller genome than any species of *Lasiocephalus* (mean 2C = 12.35 pg; Table 1).

In both datasets, samples with the smallest genome sizes formed a discrete group (Fig. 3). This group consisted of two taxa of putative hybrid origin – *Lasiocephalus* sp. 4 (cf. *L. ledifolius*) (a putative intergeneric hybrid between *L. ovatus* and *Culcitium nivale*) and *Lasiocephalus* sp. 7 (a putative hybrid between *L. sodiroi* and *C. nivale*). DNA content values of suspected hybrids were intermediate between values of their putative parents.

A continuous variation in nuclear DNA content was observed in the remaining samples. Nevertheless, this variation was sorted according to the growth form and preferred habitat (Fig. 3). Analysis of variance supported significant differences in genome size among the different habitats ($F_{3,23} = 7.16$, $P = 0.01$ and $F_{3,125} = 9.38$, $P = 0.001$ for 'fresh'

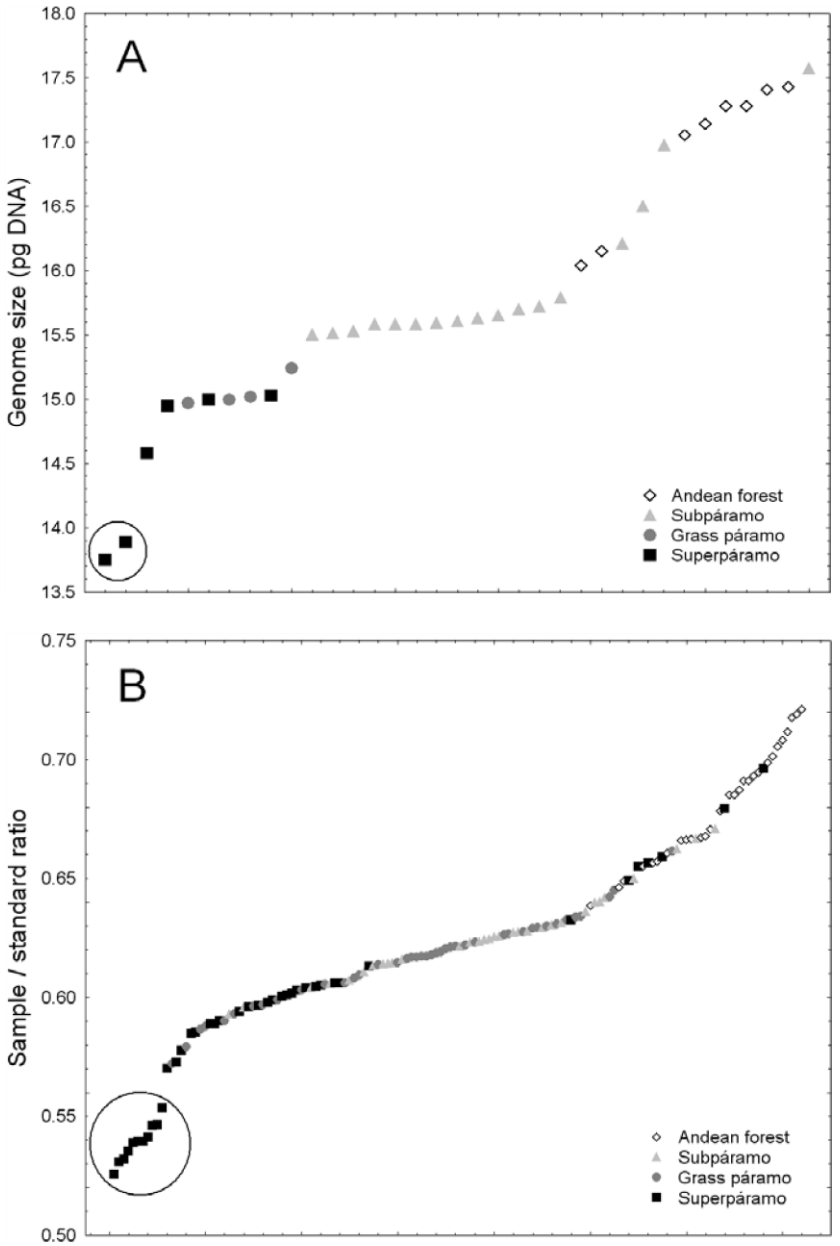


Fig. 3. – Variation in holoploid genome size of putatively diploid *Lasiocephalus*. (A) Absolute genome sizes (2C values) in pgDNA of35accessions of13*Lasiocephalus*species ('fresh' dataset, stainingwith propidiumiodide). (B) Relative fluorescence intensity (expressed as a ratio of the sample and the internal reference standard, *Vicia faba*) in 144 accessions of 19 *Lasiocephalus*species ('silica' dataset, stained with DAPI). Different symbols refer to different habitat types (growth forms correlate with the habitat types, see Table 1). Circles depict putative hybrids between *Lasiocephalus* spp. and *Culcitium nivale*. Note the generally increasing DNA content in species inhabiting different habitats in the sequence: superpáramo < grass páramo < subpáramo < Andean forest.

and ‘silica’ datasets, respectively) and growth forms ($F_{2,23}$ = 6.23, P = 0.02 and $F_{2,125}$ = 5.10, P = 0.02 for ‘fresh’ and ‘silica’ datasets, respectively). Generally, the genome size of the *Lasiocephalus* species analyzed increased with decreasing altitude as documented by the following habitat sequence: 1) superpáramo plants (mean 2C = 14.53 pg, mean sample/standard ratio = 0.591), 2) grass páramo plants (mean 2C = 15.06 pg DNA, mean sample/standard ratio = 0.616), 3) subpáramo plants (mean 2C = 15.71 pg, mean sample/standard ratio = 0.626) and 4) forest plants (mean 2C = 16.97 pg DNA, mean sample/standard ratio = 0.681). Similarly, a progressive increase in DNA content was observed among different growth forms. The narrow-leaved herbaceous plants had the lowest mean DNA values (2C = 14.74 pg, sample/standard ratio = 0.603), followed by broad-leaved herbaceous plants (sample/standard ratio = 0.617), lianas (2C = 16.15 pg, sample/standard ratio = 0.652) and a shrub (2C = 17.27 pg). The only exception was *L. longipenicillatus*, a narrow-leaved superpáramo herbaceous plant, with a distinctly higher genome size than other species of the same growth form and habitat (see Fig. 3). It should, however, be noted that the negative correlation between genome size and altitude was statistically supported only by the ‘silica’ dataset ($F_{1,124}$ = 7.611, P = 0.007) but not by the less representative dataset of fresh samples ($F_{1,22}$ = 0.001, P = 0.996).

ITS phylogeny and genome size

Bayesian phylogenetic analysis of ITS sequences run on a subset of 13 *Lasiocephalus* species revealed two well-supported lineages (Fig. 4), referred to as ‘forest’ and ‘páramo’ clades because of the markedly different habitat preferences of their members.

The ‘forest’ clade consists almost entirely of lianas inhabiting Andean forest and subpáramo habitats (see Table 1). The only exception was *L. longipenicillatus*, a superpáramo herbaceous plant from Venezuela. The ‘forest’ clade appears to be further subdivided into three sub-clades: (i) forest + superpáramo taxa (sub-clade ‘a’ at Fig. 4), (ii) entirely forest species (sub-clade ‘b’), and (iii) forest + subpáramo taxa (sub-clade ‘c’).

The ‘páramo’ clade is composed of four exclusively grass páramo and superpáramo narrow-leaved herbaceous plants with unresolved relationships. The closely related *Culcitium nivale* and one putatively hybrid taxon *Lasiocephalus* sp. 4 (cf. *L. ledifolius*) with similargrowth form and habitatpreferences, are a sistergroup to the ‘páramo’ clade.

One species (*L. pichinchensis*) possessed ITS ribotypes characteristic of both major clades (Fig. 4); sequences were similar to *L. otophorus* and *L. ovatus* from the ‘forest’ and ‘páramo’ clades, respectively, and were isolated from a single cloned individual. For this reason, *L. pichinchensis* was omitted from subsequent tests of correlation between genome size and phylogeny. Our data do not support the monophyly of the genus *Lasiocephalus* sensu Cuatrecasas (1978) because of the polytomy at the base of the whole group, which also includes several species of the closely related genus *Culcitium* (Fig. 4).

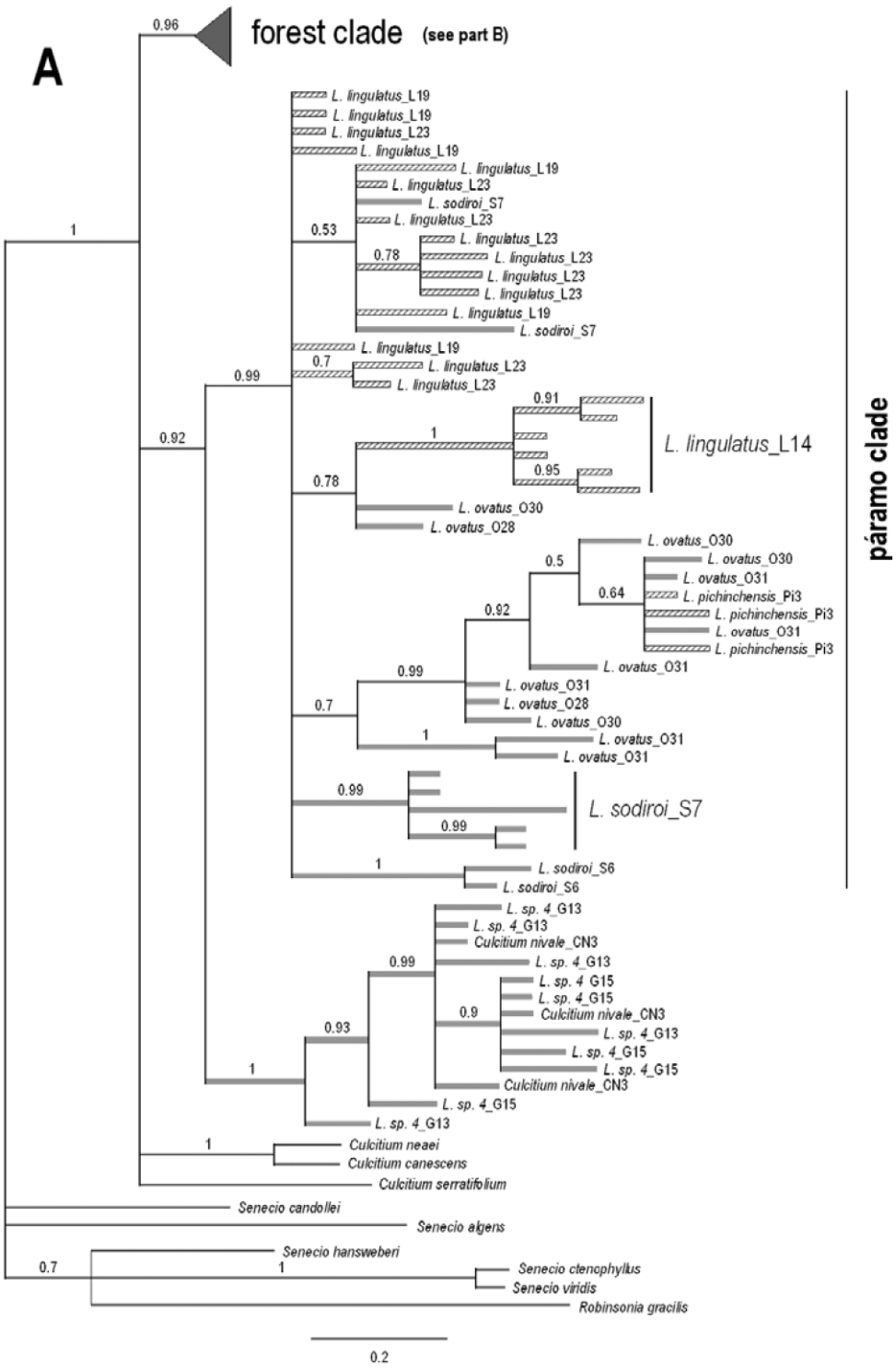
The two main clades differ markedly in DNA ploidy and DNA content variation. DNA triploids were entirely confined to the ‘páramo’ clade (polyploids accounted for 17% of all samples in this clade) whereas the ‘forest’ clade was ploidy-uniform. In addition, diploid species from the ‘páramo’ clade have significantly smaller genome sizes than their ‘forest’ counterparts based on both the ‘fresh’ (Mann-Whitney Z = 2.121, P = 0.034; mean 2C = 14.97 and 16.43 pg, respectively)

and ‘silica’ datasets (Z = 2.717, P = 0.007; mean sample/standard ratio = 0.606 and 0.658, respectively).

Relationships between DNA content and growth form, habitat preferences and altitudinal range within each of the two main clades were further analyzed in the more representative ‘silica’ dataset. Species occupying different habitats also differed significantly in their DNA content values both within the ‘forest’ clade ($F_{2,52}$ = 12.681, P = 0.011) and within the ‘páramo’ clade ($F_{1,34}$ = 170.16, P = 0.048). On the other hand, C-values in the ‘forest’ clade were not associated with different growth forms ($F_{1,52}$ = 0.013, P = 0.915); this test was not performed for the ‘páramo’ clade because all species were narrow-leaved herbaceous plants. The DNA content of ‘forest’ clade taxa increased with decrease in altitude ($F_{1,51}$ = 4.342, P = 0.042) while no such association was found within the ‘páramo’ clade ($F_{1,33}$ = 0.048, P = 0.828). In addition, a clear phylogenetic structure within the ‘forest’ clade was subjected to further analysis. However, only non-significant differences in fluorescence values among the three subclades (Fig. 4) were detected (Kruskal-Wallis H = 0.125, P = 0.939).

Relations among the explanatory variables

Variation partitioning provided a better insight into the relations among habitat, growth form and phylogeny that were selected as possible explanatory variables for the intrageneric variation in genome size in *Lasiocephalus*. Strong correlations among the variables were obtained (Fig. 5). A large proportion of the variation was explained by a combination of all three factors (71% and 48% for ‘fresh’ and ‘silica’ datasets, respectively). Habitat seemed to be the single most important explanatory factor, although this result may be biased due to an unequal number of degrees of freedom for different variables (3 for habitat type versus 1 for ITS clade and 2 for growth form).


A
forest clade (see part B)
páramo clade
L. lingulatus_L14
L. sodiroi_S7
L. lingulatus_L19
L. lingulatus_L23
L. sodiroi_S7
L. ovatus_O30
L. ovatus_O28
L. ovatus_O31
L. pichinchensis_Pi3
L. sodiroi_S6
L. sp. 4_G13
L. sp. 4_G15
Culcitium nivale_CN3
Culcitium neaei
Culcitium canescens
Culcitium serratifolium
Senecio candollei
Senecio algens
Senecio hansweberi
Senecio ctenophyllus
Senecio viridis
Robinsonia gracilis
0.2

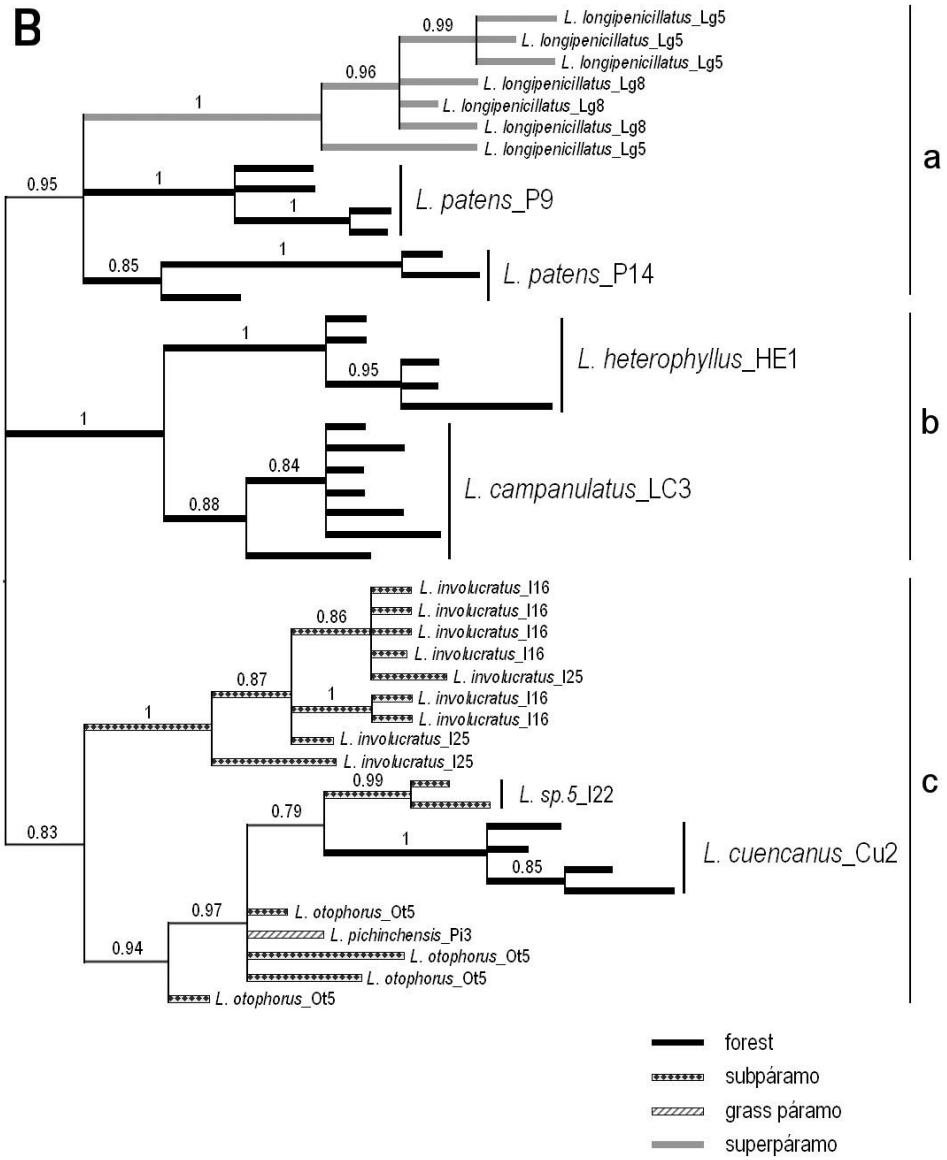


Fig. 4. – Phylogenetic analysis of ITS sequences of thirteen *Lasiocephalus* species and their closest relatives in the *Senecio* s.l. group.A Bayesian 50% probabilityconsensus tree of 4502trees is shown with posterior probabilities indicated above the branches. (A) Basal part of the tree and 'páramo' clade. (B) 'forest' clade further divided into three subclades (designated by a, b and c letters). Mean 2C-values for 'forest' and 'páramo' clades were 16.43pg and 14.97pg DNA, respectively. Different patterns of the branches indicate different habitat preferences of the respective accession/lineage.

Discussion

This study aimed at addressing the role of different evolutionary forces in shaping genome sizein a diverse plant group. We estimatedrelative or absolute genome sizevalues in a representative sample set of the high-Andean genus *Lasiocephalus* (20 taxa, 189 samples) and interpreted the results in the light of the phylogenetic relationships, growth form and ecological preferences.

DNA ploidy level variation

Polyploidy is widely acknowledged as the key force in angiosperm evolution with recent or ancient genome duplication being detected in virtually all angiosperms (Soltis et al. 2009). Although no polyploids have previously been reported in *Lasiocephalus*, our FCM data provide clear evidence for intrageneric cytotype polymorphism. The most parsimonious explanation for the existence of two discrete groups of genome sizes (differing by ~1.5-fold) is the incidence of diploid and triploid plants. Because exact chromosome counts for putatively triploid individuals are not available we refer to these large-genome accessions as DNA triploids (Suda et al. 2006).

DNA triploids are most probably of autopolyploid origin as suggested by their: (i) close morphological similarity with diploid individuals of the corresponding species, and (ii) regular sympatric growth of 2x and 3x cytotypes in the same populations. Indeed, no population for which several samples were analyzed was composed entirely of DNA triploids (but triploids were detected in two populations of *L. lingulatus* for which only one individual was analyzed). Most likely, triploids originated recurrently in each population as a result of fusion of reduced and

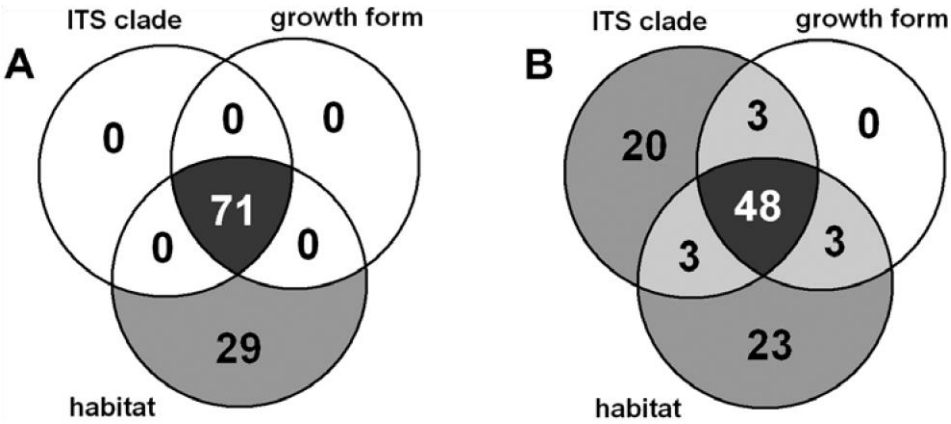


Fig. 5. – Partitioning of variance in genome size of *Lasiocephalus* samples explained by habitat, growth form and phylogeny. (A) 'fresh' dataset, (B) 'silica' dataset. The numbers represent percentages of explained variance (which was 96% and 83% of the total variance in A and B, respectively). The relative importance of a particular factor or factor combination is distinguished by different shades of grey. Note the strong correlation of all three factors. High proportion of variance explained solely by the habitat may be related to the higher number of degrees of freedom for this variable. Differences in the percentages of explained variance between both datasets reflect unequal sample sizes.

unreduced gametes of diploids, such as in *Picris* (Slovák et al. 2009) and *Quercus* (Dzialuk et al. 2007).

Interestingly, all ten DNA triploids belonged to only three narrow-leaved *Lasiocephalus* taxa, namely *L. lingulatus*, *L. ovatus* and *L.* sp. 6 (resembling *L. sodiroi*). All of them grow in the high-altitude grass páramo or superpáramo and are a part of the 'páramo' clade (*Lasiocephalus* sp. 6 was not sequenced but is believed to belong to this clade because of close phenotypic similarities to another clade member, *L. sodiroi*). Genome duplication thus seems to be restricted to a small group of closely related taxa inhabiting the highest and most stressful environments. This pattern can be explained by the higher viability of polyploids in harsh conditions and/or by the higher frequency of unreduced gamete formation in a stressful and fluctuating climate(regular frost, high solar radiation, etc.; Sarmiento 1986). Relatively higher incidence of triploid plants at high altitudes is also recorded for instance in the *Cardamine pratensis* group (*C. rivularis*; Marhold 1994).

Genome size variation at diploid level

The three chromosome counts (2n = 40) determined in this study confirm the previous numbers recorded for Venezuelan *Lasiocephalus longipenicillatus* and Ecuadorian *L. involucratus*, *L. patens* and *Culcitium nivale* (Powell & Cuatrecasas 1970, Robinson et al. 1997). Despite quite a high number of somatic chromosomes, allthese counts are regarded as diploid based on x = 20. This number seems to be the base chromosome number also in some other representatives of the widely conceived genus *Senecio* (see also Suda et al. 2007), although some authors suggest x = 10 (e.g. Hodálová et al. 2007 and references therein). Because the karyologically-confirmed diploid plants are distributed across the entire range of genome sizesthatremain after the exclusion of DNA triploids (Table 1), the 1.32-fold difference was interpreted as a genuine variation in monoploid genome size. It is likely that amplification or deletion of non-genic repetitive DNA (transposable elements in particular), among other mechanisms, drive this variation (Bennetzen et al. 2005, Cavalier-Smith 2005, Piegu et al. 2006).

Genome size and homoploid hybridization

Genome size is mostly stable at the species level, while it often shows considerable differences even among closely related species (Greilhuber et al. 2005, Leong-Škorničková et al. 2007, Ekrt et al. 2009, Loureiro et al. 2010). As such, genome size can be used as a supportive species-specific characteristic and is also a helpful marker for investigating interspecific hybridization. The value of genome size data lies in the fact that both heteroploid and homoploid hybrids can be identified based on intermediate DNA amounts (e.g. Šiško et al. 2003, Mahelka et al. 2005, Yahata et al. 2006, Kron et al. 2007, but see Bureš et al. 2004).

In this study, two different intergeneric hybrids between *Lasiocephalus* spp. and closely related *Culcitium nivale* were assumed on the basis of field observations and plant phenotypes. In both cases, FCM data confirmed our expectation. The average genome size of *Lasiocephalus* sp. 4 (cf. *L. ledifolius*) (2C = 13.82 pg), a putative hybrid between *L. ovatus* (2C = 14.99 pg) and *C. nivale* (2C = 12.35 pg), closely matched the theoretical DNA amount of the F1 hybrid (13.67 pg). Putative crosses showed intermediate morphological characteristics with respect to the size and shape of leaves, type of indumentum, capitula diameter, etc. Further support for their hybrid origin comes from the fact that only a small number of often non-flowering individuals scattered among

abundant and spatially intermingled putative parents was observed at all the localities sampled. An analogous situation exists in a mixed population of *C. nivale* and another páramo species, *L. sodiroi*. The population of plants provisionally called *Lasiocephalus* sp. 7 (no. 11088 in Electronic Appendix 1) is comprised of morphologically transient individuals with fluorescence values (mean sample/standard ratio of 0.545) intermediate between the values of putative parents (0.494 and 0.599, respectively).

In contrast to the FCM results, the ITS sequences did not provide direct evidence supporting the hybrid origin of *Lasiocephalus* sp. 4 (cf. *L. ledifolius*). All clones isolated from two *L.* sp. 4 individuals originating from two different populations were clearly assigned to the *Culcitium nivale* lineage (most of them fell directly among the *C. nivale* accessions). The absence of *L. ovatus*-type ribotypes can be explained by at least two non-exclusive hypotheses: (i) the *L. ovatus*-belonging paralogues were not sampled due to either the limited number of clones (10 clones in total) or because of biased PCR amplification (Wagner et al. 1994, Kanagawa 2003), (ii) concerted evolution could have homogenized the paralogues in putative hybrid accessions towards a *C. nivale* copy. Rapid homogenization of rDNA in several generations is reported for instance in allopolyploid *Nicotiana* (Kovarik et al. 2004) and *Tragopogon* (Kovarik et al. 2005). Ongoing molecular work aims to shed more light on the evolutionary history of *Lasiocephalus* sp. 4 (cf. *L. ledifolius*).

Genome size and phylogeny

The recent burst of phylogeny-based studies on genome size variation revealed a strong phylogenetic component in this trait at various taxonomic levels, from seed plants as a whole (Leitch et al. 2005), through family (Leitch et al. 2007) and (sub)genus (Weiss-Schneeweiss et al. 2005, Garcia et al. 2008, Chrtek et al. 2009) up to the intraspecific level (Popp et al. 2008). Neutral evolution (i.e. random and passive changes in genome size) may often be as important as selective processes in shaping genome size (Lysak et al. 2009).

Two main clades were identified in the *Lasiocephalus* accessions analyzed (Fig. 4) – one containing lianas from the Andean forest and subpáramo habitats (plus one superpáramo species, *L. longipenicillatus*) ('forest' clade) and the other exclusively grass páramo and superpáramo narrow-leaved herbaceous plants ('páramo' clade). Interestingly, both clades differed significantly in genome size, with smaller genomes recorded in the 'páramo' clade. On the other hand, genome size in *Lasiocephalus* did not appear to be affected so much by phylogeny at the lower divergence levels as documented by the non-significant differences in genome size between the three subclades within the 'forest' clade.

Any further considerations of the ancestral state and possible direction of genome size evolution in the two main clades are, however, largely obscured by a polytomy at the base of the whole group that involves several *Culcitium* species. Moreover, *C. nivale* appears to be more closely related to the *Lasiocephalus* 'páramo' clade than to other members of its own genus (Fig. 4). Thus the data indicate that in order to re-establish the group as monophyletic *C. nivale*, at least, has to be included in *Lasiocephalus*. Nevertheless, at this stage of the investigation it is not possible to exclude that the closeness of the relationship between *Lasiocephalus* and *C. nivale* is overestimated due to the incidence of hybridization and introgression. In addition, only direct non-cloned sequences of other *Culcitium* species were obtained from the GenBank (Pelser et al. 2007). As noted by Soltis et al. (2008), the use of direct ITS sequences could mislead the phylogenetic

inference especially in paralogue-rich groups. We also admit that the explanatory value of the current phylogenetic tree may be limited by the incomplete taxon sampling (only 13 out of 20 FCM-analyzed *Lasiocephalus* taxa were included). On the other hand, the sequenced dataset is sufficiently representative. It covers the whole range of estimated genome sizes and all major types of growth forms and habitats (perhaps with a single exception of the broad-leaved herbaceous *L. mojandensis*). A multi-marker molecular analysis applied to a larger dataset is currently in progress in order to draw more robust conclusions regarding the phylogenetic relationships between *Culcitium* and *Lasiocephalus*.

Genome size and environmental correlates

The influence of environmental factors on genome size evolution has been debated for years and is still a matter of controversy. Significant correlations of genome size with various ecological factors (see Leitch & Bennett 2007 for a review) differ when different model systems are compared. For example, of the correlations of genome size with altitude for 24 genera, approximately one third are positive, one third negative and one third not significant (Knight et al. 2005). Similarly, while Suda et al. (2003) report correlations between genome size and altitude, average annual temperature, rainfall and/or humidity in a number of Macaronesian genera, these relationships disappear when the sampling is more representative (Suda et al. 2005). In addition, rigorous studies on variation in genome size at the lowest divergence levels (within species or even within population) often fail to detect any link between genome size and environmental conditions (Šmarda et al. 2007).

Significant correlations between genome size and growth form, habitat and altitude were also found in *Lasiocephalus*. On average, species from high-altitude harsh environments (superpáramo) have the smallest genomes while their counterparts from lower altitudes and more favourable conditions (Andean forest) have the largest genomes, which accords with the 'large genome constraint' theory (Knight et al. 2005). Different growth forms are associated with different altitudes: narrow-leaved herbaceous plants on the superpáramo and lianas in forests. However, itshould be noted thatallthese variables were strongly correlated with the major phylogenetic split in *Lasiocephalus* (i.e. the division into two lineages referred to as 'páramo' and 'forest' clades; Fig. 4, see also Fig. 5). Although we do not apriori reject that environmental factors may have somehow shaped the genome size, we are convinced that phylogenetic relationships were the key determinant of the observed divergence in genom size a tthe highest divergence levels. This assumption is strongly supported by the fact that the only superpáramo herbaceous plant, *L. longipenicillatus* with a large genome, was included in the 'forest' clade. These findings emphasize that for the correct interpretation of the environmental correlates of genome size it is crucially important to incorporate phylogenetic information into analyses (Grotkopp et al. 2004, Jakob et al. 2004, Chase et al. 2005, Beaulieu et al. 2007, Chrtek et al. 2009).

Environmental factors seem to be more important in shaping the genome size of *Lasiocephalus* at lower divergence levels. While the three main sub-clades within the 'forest' clade do not significantly differ from each other in terms of genome size, the variation is structured according to the habitat preferences of the respective species (as also reflected by the negative correlation between genome sizeand altitude). Even within the sub-clades, high-altitude species have smaller DNA C-values than their low-altitude counterparts. Similar patterns are observed, for

example, in the genus *Hordeum* where the environmental variable (length of vegetation period) is correlated with genome size only at the lowest taxonomic levels (Jakob et al. 2004).

Conclusions

This study is the first attempt to assess genome size variation in a radiating plant group in the tropics and interpret the results in the light of ecological conditions, life history traits and phylogeny. It indicates that the genome size in high-Andean *Lasiocephalus* has been shaped by different evolutionary processes, including polyploidy, hybridization and gradual changes at the molecular level. The variation in genome size at the higher divergence level primarily reflects phylogenetic effects, while ecological factors are more important in shaping genome size at the lower divergence levels. This highlights the need to always evaluate ecological and phenotypic correlates of genome size within a phylogenetic framework to avoid misinterpretation (e.g. overestimating the significance of environmental factors). We believe that this study is likely to stimulate further research into the promising but still largely unexplored field of tropical polyploid evolutionary biology and cytogeography.

See http://www.preslia.cz for Electronic Appendix 1.

Acknowledgements

This research was supported by the Czech Science Foundation (project no. 206/07/0273), Grant Agency of the Academy of Sciences of the Czech Republic (project no. KJB601110710) and partly also by projects MSM 0021620828 (Ministry of Education, Youth and Sports of the Czech Republic), GAUK 261211 (Charles University Prague) and AV0Z60050516 (Academy of Sciences of the Czech Republic). We thank V. Jarolímová and E. Záveská for their assistance with the karyological and molecular analyses, respectively, P. Ubiergo, K. Romoleroux, D. Vázquez and P. Macek for their help with the field work, and J. Lepš and Z. Münzbergová for advice with statistical analyses.

References

Beaulieu J. M., Leitch I. J. & Knight C. A. (2007): Genome size evolution in relation to leaf strategy and metabolic rates revisited. – Ann. Bot. 99: 495–505.

Bell C. D. & Donoghue M. J. (2005): Phylogeny and biogeography of *Valerianaceae* (*Dipsacales*) with special reference to the South American valerians. – Org. Divers. Evol. 5: 147–159.

Bennett M. D. & Leitch I. J. (2005): Genome size evolution in plants. – In: Gregory T. R. (ed.), The evolution of the genome, p. 89–162, Elsevier, San Diego.

Bennetzen J. L. (2005): Transposable elements, gene creation and genome rearrangement in flowering plants. – Curr. Opin. Genet. Dev. 15: 621–627.

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

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

Cavalier-Smith T. (2005):Economy, speed and size matter: evolutionary forces driving nuclear genome miniaturization and expansion. – Ann. Bot. 95: 147–175.

Chase M. W., Hanson L., Albert V. A., Whitten W. M. & Williams N. H. (2005): Life history evolution and genome size in subtribe *Oncidiinae* (*Orchidaceae*). – Ann. Bot. 95: 191–199.

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

Cuatrecasas J. (1978): Studies in Neotropical *Senecioneae*, *Compositae* I. Reinstatement of genus *Lasiocephalus*. – Phytologia 40: 307–312.

Cuatrecasas J. (1986): Speciation and radiation of the *Espeletiinae* in the Andes. – In: Vuilleumier F. & Monasterio M. (eds), High altitude tropical biogeography, p. 267–303, Oxford University Press, New York.

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

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

Dyer A. F. (1963): The use of lacto-propionic orcein in rapid squash methods for chromosome preparations. – Stain Technol. 38: 85–90.

Dzialuk A., Chybicki I., Welc M., Śliwińska E. & Burczyk J. (2007): Presence of triploids among oak species. – Ann. Bot. 99: 959–964.

Ekrt L., Trávníček P., Jarolímová V., Vít P. & Urfus T. (2009): Genome size and morphology of the *Dryopteris affinis* group in Central Europe. – Preslia 81: 261–280.

Garcia S., Canela M. A., Garnatje T., McArthur E. D., Pellicer J., Sanderson S. C. & Vallès J. (2008): Evolutionary and ecological implications of genome size in the North American endemic sagebrushes and allies (*Artemisia*, *Asteraceae*). – Biol. J. Linn. Soc. 94: 631–649.

Greilhuber J. (2005): Intraspecific variation in genome size in angiosperms: identifying its existence. – Ann. Bot. 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. – Ann. Bot. 95: 255–260.

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

Hall T. A. (1999): BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. – Nucl. Acids Symp. 41: 95–98.

Hodálová I., Grulich V., Horová L., Valachovič M. & Marhold K. (2007): Occurence of tetraploid and octoploid cytotypes in *Senecio jacobaea* ssp. *jacobaea* (*Asteraceae*) in Pannonia and the Carpathians. – Bot. J. Linn. Soc. 153: 231–242.

Huelsenbeck J. P. & Ronquist F. (2001): MRBAYES: Bayesian inference of phylogeny. – Bioinformatics 17: 754–755.

Hughes C. & Eastwood F. (2006): Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. – Proc. Natl. Acad. Sci. USA 103: 10334–10339.

Jakob A., Meister A. & Blattner F. R. (2004): The considerable genome size variation of *Hordeum* species (*Poaceae*) is linked to phylogeny life form, ecology, and speciation rates. – Mol. Biol. Evol. 21: 860–869.

Kanagawa T. (2003): Bias and artifacts in multitemplate polymerase chain reactions (PCR). – J. Biosc. Bioeng. 96: 317–323.

Knight C. A. & Ackerly D. D. (2002): Genome size variation across environmental gradients in the California flora. – Ecol. Lett. 5: 66–76.

Knight C. A. & Beaulieu J. (2008): Genome size scaling in phenotype space. – Ann. Bot. 101: 759–766.

Knight C. A., Molinari N. & Petrov D. (2005). The large genome constraint hypothesis: evolution, ecology, and phenotype. – Ann. Bot. 5: 177–190.

Kovarik A., Matyasek R., Lim K. Y., Skalická K., Koukalová B., Knapp S., Chase M. & Leitch A. R. (2004): Concerted evolution of 18-5.8-26S r DNA repeats in *Nicotiana* allotetraploids. – Biol. J. Linn. Soc. 82: 615–625.

Kovarik A., Pires J. C., Leitch A. R., Lim K. Y., Sherwood A. M., Matyasek R., Rocca J., Soltis D. E. & Soltis P. S. (2005): Rapid concerted evolution of nuclear ribosomal DNA in two *Tragopogon* allopolyploids of recent and recurrent origin. – Genetics 169: 931–944.

Kron P., Suda J. & Husband B. C. (2007): Applications of flow cytometry to population biology. – Annu. Rev. Ecol. Evol. Syst. 38: 847–876.

Leitch I. J., Beaulieu J. M., CheungK., HansonL., Lysak M. A. & Fay M. F. (2007):Punctuated genome size evolution in *Liliaceae*. – J. Evol. Biol. 20: 2296–2308.

Leitch I. J. & Bennett M. D. (2007) Genome size and its uses: the impact of flow cytometry. – In: Doležel J., Greilhuber J. & Suda J. (eds), Flow cytometry with plant cells, p. 153–176, Wiley-VCH, Weinheim.

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

Leong-Škorničková J., Šída O., Jarolímová V., Sabu M., Fér T., Trávníček P. & Suda J. (2007): Chromosome numbers and genome size variation in Indian species of *Curcuma* L. (*Zingiberaceae*). – Ann. Bot. 100: 500–526.

Lindstrom M. J. & Bates D. M. (1990): Nonlinear mixed effects models for repeated measures data. – Biometrics 46: 673–687.

Loureiro J., Trávníček P., Rauchová J., Urfus T., Vít P., Štech M., Castro S. & Suda J. (2010): The use of flow cytometry in the biosystematics, ecology and population biology of homoploid plants. – Preslia 82: 3–21.

Luteyn J. L. (1999): Páramos: a checklist of plant diversity, geographical distribution, and botanical literature. – Mem. New York Bot. Gard. 84: 1–278.

Lysak M. A., Koch M., Beaulieu J. M., Meister A. & Leitch I. J. (2009): The dynamic ups and downs of genome size evolution in *Brassicaceae*. – Mol. Biol. Evol. 26: 85–98.

Mahelka V., Suda J., Jarolímová V., Trávníček P. & Krahulec F. (2005): Genome size discriminates between closely related taxa *Elytrigia repens* and *E. intermedia* (*Poaceae: Triticeae*) and their hybrid. – Folia Geobot. 40: 367–384.

Marhold K. (1994): Chromosome numbers of the genus *Cardamine* L. (*Cruciferae*) in the Carpathians and in Pannonia. – Phyton (Horn) 34: 19–34.

Otto F. (1990): DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. – In: Crissman H. A. & Darzynkiewicz Z. (eds), Methods in cell biology 33, p. 105–110, Academic Press, New York.

Pelser P. B., Nordensatm B., Kadereit J. W. & Watson L. E. (2007): An ITS phylogeny of tribe *Senecioneae* (*Asteraceae*) and a new delimitation of *Senecio* L. – Taxon 56: 1077–1104.

Piegu B., Guyot R., Picault N., Roulin A., Saniyal 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 Res. 16: 1262–1269.

Popp M., Erixon P., Eggens F. & Oxelman B. (2005): Origin and evolution of a circumpolar polyploid species complex in *Silene* (*Caryophyllaceae*) inferred from low copy nuclear RNA polymerase introns, rDNA, and chloroplast DNA. – Syst. Bot. 30: 302–313.

PoppM., Gizaw A., Nemomissa S., Suda J. & BrochmannC. (2008):Colonizationand diversification in the African "sky islands" by Eurasian *Lychnis* L. (*Caryophyllaceae*). – J. Biogeogr. 35: 1016–1029.

Posada D. & Crandall K. A. (1998): MODELTEST: testing the model of DNA substitution. – Bioinformatics 14: 817–818.

Powel A. M. & Cuatrecasas J. (1970): Chromosome numbers in *Compositae*: Colombian and Venezuelan species. – Ann. Mo. Bot. Garden 57: 374–379.

R Development Core Team (2009): R: a language and environment for statistical computing. – R Foundation for Statistical Computing, Vienna, URL: [http://www.R-project.org].

Robinson H., Carr G. D., King R. M. & Powell A. M. (1997): Chromosome numbers in *Compositae*, XVII. *Senecioneae* III. – Ann. Mo. Bot. Gard. 84: 893–906.

Ronquist F. &. Huelsenbeck J. P. (2003): MRBAYES 3: Bayesian phylogenetic inference under mixed models. – Bioinformatics 19: 1572–1574.

Sarmiento G. (1986): Ecological features of climate in high tropical mountains. – In: Vuilleumier F. & Monasterio M. (eds), High altitude tropical biogeography, p. 11–45, Oxford University Press, New York.

Šiško M., Ivančič A. & Bohanec B. (2003): Genome size analysis in the genus *Cucurbita* and its use for determination of interspecific hybrids obtained using the embryo-rescue technique. – Plant Sci. 165: 663–669.

Slovák M., Vít P., Urfus T. & Suda J. (2009): Complex pattern of genome size variation in a polymorphic member of the *Asteraceae*. – J. Biogeogr. 36: 372–384.

Šmarda P., Bureš P. & Horová L. (2007): Random distribution pattern and non-adaptivity of genome size in a highly variable population of *Festuca pallens*. – Ann. Bot. 100: 141–150.

Soltis D. E., Albert V. A., Leebens-Mack J., Bell C. D., Paterson A. H., Zheng C., Sankoff D., dePamphilis C. W., Wall P. K. & Soltis P. S. (2009): Polyploidy and angiosperm diversification. – Am. J. Bot. 96: 336–348.

Soltis D. E., Mavrodiev E. V., Doyle J. J., Rauscher J. & Soltis P. S. (2008): ITS and ETS sequence data and phylogeny reconstruction in allopolyploids and hybrids. – Syst. Bot. 33: 7–20.

StatSoft, Inc. (2007): STATISTICA (data analysis software system), version 8.0. – URL: [http://www.statsoft.com].

Struwe L., Haag S., Heiberg E. & Grant J. R. (2009): Andean speciation and vicariance in neotropical *Macrocarpaea* (*Gentianaceae-Helieae*). – Ann. Mo. Bot. Gard. 96: 450–469.

Suda J., Krahulcová A., Trávníček P. & Krahulec F. (2006): Ploidy level versus DNA ploidy level: an appeal for consistent terminology. – Taxon 55: 447–450.

Suda J., Kyncl T. & Jarolímová V. (2005): Nuclear DNA amounts in Macaronesian angiosperms: forty percent of Canarian endemic flora completed. – Pl. Syst. Evol. 252: 215–238.

Suda J., Kyncl T. & Wildová R. (2003): Nuclear DNA amount in Macaronesian angiosperms. – Ann. Bot. 92: 153–164.

Suda J. & Trávníček P. (2006): Reliable DNA ploidy determination in dehydrated tissues of vascular plants by DAPI flow cytometry: new prospects for plant research. – Cytometry Part A 69A: 273–280.

SudaJ., Weiss-Schneeweiss H.,TribschA.,Schneeweiss G. M.,TrávníčekP. &SchönswetterP. (2007):Complex distribution patterns of di-, tetra- and hexaploid cytotypes in the European high mountain plant *Senecio carniolicus* (*Asteraceae*). – Am. J. Bot. 94: 1391–1401.

Vuilleumier F. & Monasterio M. (eds) (1986): High altitude tropical biogeography. – Oxford University Press, New York.

Wagner A., Blackstone N., Cartwright P., Dick M., Misof B., Snow P., Wagner G. P., Bartels J., Murtha M. & Pendleton J. (1994): Surveys of gene families using polymerase chain reaction: PCR selection and PCR drift. – Syst. Biol. 43: 250–261.

Weiss-Schneeweiss H., Greilhuber J. & Schneeweiss G. M. (2005): Genome size evolution in holoparasitic *Orobanche* (*Orobanchaceae*) and related genera. – Am. J. Bot. 93: 148–156.

White T. J., Bruns T., Lee S. & Taylor J. (1990): Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. – In: Innis M., Gelfand D., Sninsky J. & White T. (eds), PCR protocols: a guide to methods and applications, p. 315–322, Academic Press, San Diego.

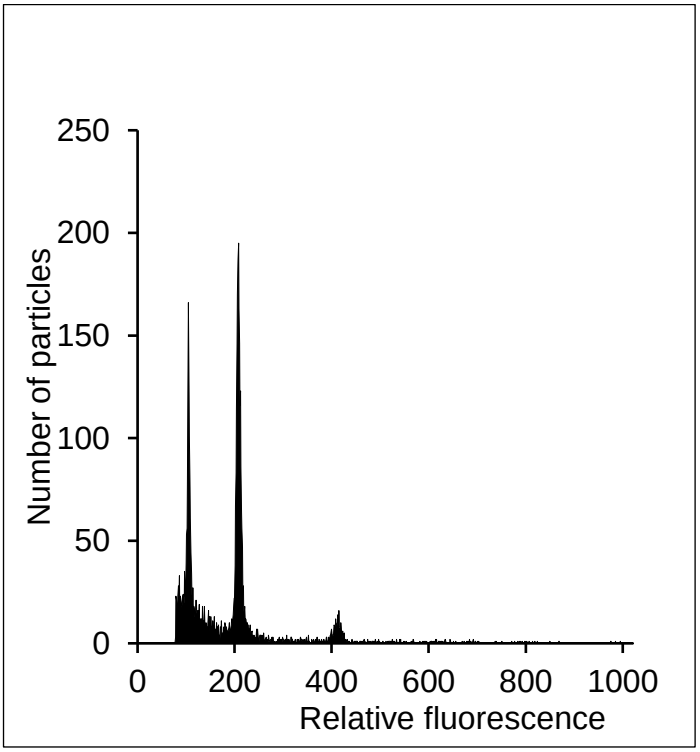
Wilkström N., Kenrick P. & Chase M. (1999): Epiphytism and terrestrialization in tropical *Huperzia* (*Lycopodiaceae*). – Pl. Syst. Evol. 218: 221–243.

Yahata M., Kunitake H., Yasuda K., Yamashita K., Komatsu H. & Matsumoto R (2006): Production of sexual hybrid progenies for clarifying the phylogenic relationship between *Citrus* and *Citropsis* species. – Am. Soc. Hort. Sci. 131: 764–769.

Young K. R., Ulloa C. U., Luteyn L. J. & Knapp S. (2002): Plant evolution and endemism in Andean South America: an introduction. – Bot. Rev. 68: 4–21.

Case study V

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**



Naturalized plants have smaller genomes than their non-invading relatives: a flow cytometric analysis of the Czech alien flora

Magdalena Kubešová[1,2], Lenka Moravcová[1], Jan Suda[2,1], Vojtěch Jarošík[3,1] & Petr Pyšek[1,3]

[1] Institute of Botany, Academy of Sciences of the Czech Republic, CZ-252 43 Průhonice,Czech Republic, e-mail: kubesovm@gmail.com, moravcova@ibot.cas.cz, pysek@ibot.cas.cz;
[2] Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, CZ128 01 Prague, Czech Republic, e-mail: suda@natur.cuni.cz;
[3] Department of Ecology, Faculty of Science, Charles University in Prague, Viničná 7, CZ-128 44 Prague, Czech Republic, e-mail: jarosik@cesnet.cz

Abstract

Genome size has been suggested as one of the traits associated with invasiveness of plant species. To provide a quantitative insight into the role of this trait, we estimated nuclear DNA content in 93 alien species naturalized in the Czech Republic, belonging to 32 families, by using flow cytometry, and compared it with the values reported for non-invading congeneric and confamilial species from the Plant DNA C-values database. Species naturalizedin the Czech Republic have significantly smaller genomes than their congeners not known to be naturalized or invasive in any part of the world. This trend is supported at the family level: alien species naturalized in the Czech flora have on average a smaller genome than is the mean value for non-invading confamilials. Moreover, naturalized and non-invading species clearly differed in the frequency of five genome size categories; this difference was mainly due to very small genomes prevailing and intermediate to very large genomes underrepresented in the former group. Our results provide the first quantitative support for association of genome size with invasiveness, based on a large set of alien species across a number of plant families. However, there was no difference in the genome size of invasive species compared to naturalized but non-invasive. This suggests that small genome size provides alien plants with an advantage already at the stage of naturalization and need not be necessarily associated with the final stage of the process, i.e. invasion.

Keywords: alien plants, confamilials, congeners, C-value, flow cytometry, genome size, invasive species, large genome constraint hypothesis, nuclear DNA content, plant invasions

Introduction

The numbers of invasive species in various parts of the world continue to increase, representing a serious threat to biodiversity worldwide (e.g. Meyerson & Mooney 2007, Blackburn et al. 2009, Hulme et al. 2009b, McGeoch et al. 2010). As a result, biological invasions have been receiving serious attention from both scientists and practitioners and research in invasive plant and animal species has been increasing exponentially (e.g. Crall et al. 2006, Pyšek et al. 2006, 2008, Lambdon et al. 2008, Ricciardi & MacIsaac 2008, Chytrý et al. 2009, DAISIE 2009, Davis 2009).

The knowledge of ecological impacts on native biodiversity and ecosystem functioning improved dramatically in the last decade (e.g. Levine et al. 2003, Gaertner et al. 2009, Hejda et al. 2009a, Nentwig et al. 2010, Vilà et al. 2010), and it is now widely recognized that invasive species incur serious economic costs (Binimelis et al. 2007, Kettunen et al. 2009, Vilà et al. 2010). The awareness of the magnitude of problem is stimulating not only management efforts (Keller et al. 2007, Richardson et al. 2007, Hulme et al. 2008, 2009a, Simberloff 2009, McGeoch et al. 2010), but also research aimed at deeper understanding of underlying processes and determinants of naturalization and invasiveness (e.g. Pyšek et al. 2008, 2009a, Blackburn et al. 2009, Davis 2009, Essl et al. 2009, Štajerová et al. 2009). Within this research realm, an effort to identify biological and ecological traits conferring invasiveness is as well established as the field of invasion biology itself (see Pyšek & Richardson 2007 for a review). Recent developments, including the development of new technologies (Richardson & Pyšek 2008), in particular molecular techniques, now make it possible to include traits that were until not long ago impossible to consider in multispecies studies focusing on determinants of invasiveness (Pyšek & Richardson 2007). The amount of nuclear DNA (genome size) is one of the traits for which knowledge has improved dramatically in the last decade, largely due to the advent and spread of flow cytometry (Kron et al. 2007, Ekrt et al. 2009).

Genome size is a fundamental biological parameter involved in the scaling of both plants and animals (Gregory 2005). DNA is known to play not only a qualitative (i.e. genic) role but also a quantitative one because of its direct and sequence-independent influence on cellular (and by extension, organismal) characteristics. Correlations between genome sizeand plant traitsare many and range from nuclear and cellvolumes through the duration of cell cycle (both meiotic and mitotic) up to seed size and specific leaf area (reviewed by Leitch & Bennett 2007, see also Loureiro et al. 2010). Through concomitant changes in cellular parameters, genome size affects several aspects of a plant's development. Minimum generation time (i.e. time to flowering) and life history (i.e. whether ephemeral, annual or perennial) are illustrative examples of developmental traits constrained by the amount of nuclear DNA. On average, ephemerals (plants completing their life cycle in less then seven weeks) have been shown to possess the smallest genomes, followed by annuals, whereas obligate herbaceous perennials have the highest DNA amounts (Bennett 1972). Whereas species with small genomes can display any developmental life history, their large-genome counterparts are restricted to an obligate perennial life history. Large genomes also impose constraints on ecologicalbehaviour, influencing where a plant may grow and its chances of survival in a changing world (Knight et al. 2005, Vidic et al. 2009). In addition, traits associated with genome size (seed size and mass, and the rate of developmental processes in particular) may co-determine the life strategy adopted by the plant (i.e. whether competitor, stress tolerator or ruderal). In their study on 156 weedy angiosperm species, Bennett et al. (1998) showed that the probability of being recognized as a weed decreases with increasing genome size.

On the same conceptual basis, small genomes have been suggested as a prerequisite for plant invasiveness because species with low nuclear DNA content usually produce many light seeds and their establishment is fast (Rejmánek 1996). In addition, invasions typically occur in disturbed habitats (Davis et al. 2000, Chytrý et al. 2005, 2008) and small genomes have been shown to represent an evolutionary advantage in time-limitedenvironments (Bennett 1987). In his "theory of seed plant invasiveness", Rejmánek listed a low amount of nuclear DNA among the most important factors contributing to the invasiveness of seed plants (Rejmánek 1996, 2000, Rejmánek et al. 2005). Experimental support for this conclusion comes mainly from comprehensive studies on genome size variation in the genus *Pinus* (Wakamiya et al. 1993, Grotkopp et al. 2002, 2004).

Invasiveness of pines, particularly of wind-dispersed species, was shown to be negatively associated with both genome size and seed mass (Grotkopp et al. 2002). Smaller genomes in invasive species as compared to their non-invasive congeners have also been found in some other genera such as *Senecio* (Lawrence 1985) or *Acacia* (Mukherjee & Sharma 1990), although the number of analyzed invasive species was usually quite low. In addition, a negative relationship was observed between the genome size of three *Briza* species and the invaded area (Rejmánek 1996).

Despite the pieces of evidence mentioned above for the role of genome size in plant invasions, a systematic study aimed at comparing genome sizes in invasive plant species and their non-invasive counterparts across different taxonomic groups is still lacking. To fill this gap, we determined nuclear DNA amounts in a representative set of alien species occurring in the Czech Republic and compared their genome size values with those of non-invasive congeners and confamilials. Specifically, we addressed the following questions: (i) What is the distribution of holoploid genome sizes in alien species and how it differs from the general pattern found in angiosperms? (ii) Which factors affect the genome size of alien species? Is genome size related to the invasion status? (iii) Do alien species differ in genome size from their non-invasive congeners and confamilials?

Material and methods

Analyzed species

The species set included 93 neophytes (alien species introduced after 1500 A.D.; see Pyšek et al. 2002, 2004) occurring in the flora of the Czech Republic. They belonged to 70 genera and 32 families according to the Angiosperm phylogeny group classification (Stevens 2001). Seeds were collected in the field during 2005–2007 (see Electronic Appendix 1). Seedlings were germinated in a growth chamber and cultivated in the experimental garden of the Institute of Botany, Academy of Sciences, Průhonice, Czech Republic(49°59'30"N, 14°34'00"E, ca 320 m a.s.l.). Fresh young leaftissuewas used for genome size estimation. Herbarium vouchers are kept at PRA.

Species' invasion status in the Czech Republic (casual; naturalized; invasive) was taken from Pyšek et al. (2002). The vast majority of species were naturalized, only three (*Ambrosia trifida*, *Bidens connata* and *Panicum miliaceum*) were casual; for the sake of simplicity, all the species analyzed are further referred to as 'naturalized'. Of these naturalized species, 41 were invasive and 49 naturalized but not invasive (sensu Richardson et al. 2000, Pyšek et al. 2004). Each species was further characterized (see Table 1) by its life history (annual; monocarpic perennial; polycarpic perennial) and moisture score. The moisture score was calculated by using data from Hejda et al. (2009b); this paper and associated database give, for species alien to the Czech Republic, information on habitats in which they occur in their native range. These habitats were classified using a 5-degree ordinal scale (1-dry, 3-mesic, 5-humid, with 2 and 4 representing transitions) and average value was used as the moisture score. Of our species set, habitat data for 58 species were available in Hejda etal. (2009b);forremainingspeciesweusedtheaverageEllenberg'sindicator value for moisture (Ellenberg et al. 1992), transformed to a 5-degree scale as follows: 1–3=1, 4=2, 5=3, 6–7=4, 7–9=5

Genome size estimation

Holoploid genome sizes (C-values sensu Greilhuber et al. 2005) were determined using propidium iodide flow cytometry following the simplified two-step protocol as described by Doležel et al. (2007). Briefly, young intact leaf tissue of the analyzed plant was chopped together with an appropriate internal reference standard in 0.5 ml of ice-cold Otto I buffer (0.1 M citric acid, 0.5% Tween 20; Otto 1990). The sample was filtered through 42-μm nylon mesh, and incubated 10 min at room temperature. The staining solution consisted of 1 ml of Otto II buffer (0.4 M $Na_2HPO_4 \cdot 12H_2O$) supplemented with propidium iodide and RNase IIA (both at final concentrations of 50 μg/ml) and β-mercaptoethanol (2 μl/ml). The fluorescence intensity of isolated nuclei (5000 particles) was recorded using Partec CyFlow SL cytometer equipped with a diode-pumped solid state laser 532 nm (Cobolt Samba, 100 mW output power). Each sample was analyzed at least three times on different days;only analyses with a between-day fluctuation below 3% were considered.

The following species were used as internal reference standards (Doležel et al. 2007): *Solanum lycopersicum* 'Stupické polní rané' (2C = 1.90 pg), *Glycine max* 'Polanka' (2C = 2.30 pg), *Bellisperennis* (2C = 3.46 pg), *Zea mays* 'CE-777' (2C = 5.47 pg), *Pisum sativum* 'Ctirad' (2C = 8.76 pg) and *Vicia faba* 'Inovec' (2C = 26.92 pg). *Pisum sativum* 'Ctirad' (Doležel et al. 1998) served as a primary reference standard, with 2C-value of 8.76 pg as recommended by Greilhuber et al. (2007). Genome sizes of other reference species were calibrated against *Pisum*, based on three measurements on different days. For each analyzed plant, internal standard was selected so that its genome size was close to but not overlapping with that of the analyzed sample.

Ploidy levels of analyzed naturalized plants were inferred from chromosome numbers taken from various karyological databases and flora handbooks, including Goldblatt & Johnson (1979), Marhold et al. (2007), Flora of the Czech Republic (Hejný & Slavík 1988–1992, Slavík 1995–2000, Slavík & Štěpánková 2004), the database of the flora of the Czech Republic (CzechFlor) and the internal karyological database of plants of the Czech Republic (both held at the Institute of Botany AS CR, Průhonice). Monoploid genome sizes (1Cx-values) were calculated as 2C-values / ploidy level.

Reference genome size data

To compare the genome size of naturalized alien plants in the Czech flora with non-invasive species, 2C-values and ploidy levels for plants from corresponding genera and families that are not reported to be naturalized or invasive were extracted from the Plant DNA C-values database (Bennett & Leitch 2005). Species in this reference data set are referred to as 'non-invading' to reflect not only that they are not invasive (in the sense of Richardson et al. 2000) but neither naturalized, i.e. they do not successfully enter the invasion process. The selection of non-invading congeners and confamilials was made by omitting from the Plant DNA C-values database any species reported as naturalized or invasive in any part of the world, based on the updated database of Weber (2003) and other sources. In some cases, the ploidy level taken from the Plant DNA C-values database was corrected so that the basic chromosome number (x) was the same for both naturalized species and their non-invading congeners. Reference genome size data were available for 45 congeneric and 31 confamilial non-invading counterparts.

Statistical analysis

Comparison of genome size categories between naturalized plants of the Czech flora and non-invading species taken from the Plant DNA C-values database was done by G-test on a contingency table (e.g. Crawley 2002, p. 548–550).

The effect of invasion status (41 invasive vs 51 non-invasive species, the latter including 48 naturalized and three casual; *Oxybaphus nyctagineus* was excluded because of nonavailable moisture data), life history and moisture score on 2C-values was analyzed by general linear model. The most parsimonious model was selectedby a stepwise procedure, beginning with the maximal model (containing all predictors and all their possible interactions) and proceeding by the elimination of non-significant terms, using deletion tests. This was done by an automatic step-wise process of model simplification of deviance tables, based on Akaike Information Criterion (AIC) (program Spotfire S-Plus v. 8.1, TIBCO Software Inc. 2008; e.g. Crawley 2002). Observed power of the chosen model (e.g. Steidel& Thomas 2001) was computed for $\alpha = 0.05$ in SPSS v. 18 (SPSS Inc. 2010).

Paired t-tests (Sokal & Rohlf 1995) were used for comparisons of 2C- and Cx-values of naturalized aliens with corresponding mean values of their non-invading congeners, and of mean 2C-values of naturalized aliens with corresponding mean values of their noninvading confamilials. All 2C-values were ln-transformed to normalize the data, and then checked for homogeneity of variance. The general linear model was checked by plotting standardized residuals against fitted values, and by normal probability plots (Crawley 1993).

Results

Genome size variation in naturalized alien species

Flow cytometric analyses yielded histograms with mean coefficients of variation (CVs) of 3.18% and 2.50% for the sample and internal reference standard, respectively (Fig. 1).

Genome size values were determined in one APG family (*Phrymaceae*) and 66 species for the first time (Table 1). 1C-values of analyzed plants varied from 0.24 pg in *Sisymbrium loeselii* to 15.27 pg in *Rudbeckia laciniata*, spanning ~64-fold range. The majority of naturalized species possessed low nuclear DNA amounts, with mean 1C-value of 1.93 pg and median of 1.17 pg. The distribution of genome sizes was highly skewed (Fig. 2). The most common were plants with 1C-values between 0.4 and 0.6 pg; other classes were much less frequent, and only four species had 1C-values above 4.4 pg. The same pattern was obtained when the species analyzed were attributed to five genome size categories as defined by Leitch et al. (1998). Very small genomes ($1C \leq 1.4$ pg) were represented by 49 species, small genomes ($1.4 \text{ pg} < 1C \leq 3.5$ pg) by 36 species, intermediate genomes ($3.5 \text{ pg} < 1C \leq 14.0$ pg) by seven species, and large genomes ($14.0 \text{ pg} < 1C \leq 35.0$ pg) by one species. Plants with very large genomes ($1C > 35$ pg) were not present in our species set. Compared to non-invading species (Table 2), naturalized species had more often very small and small

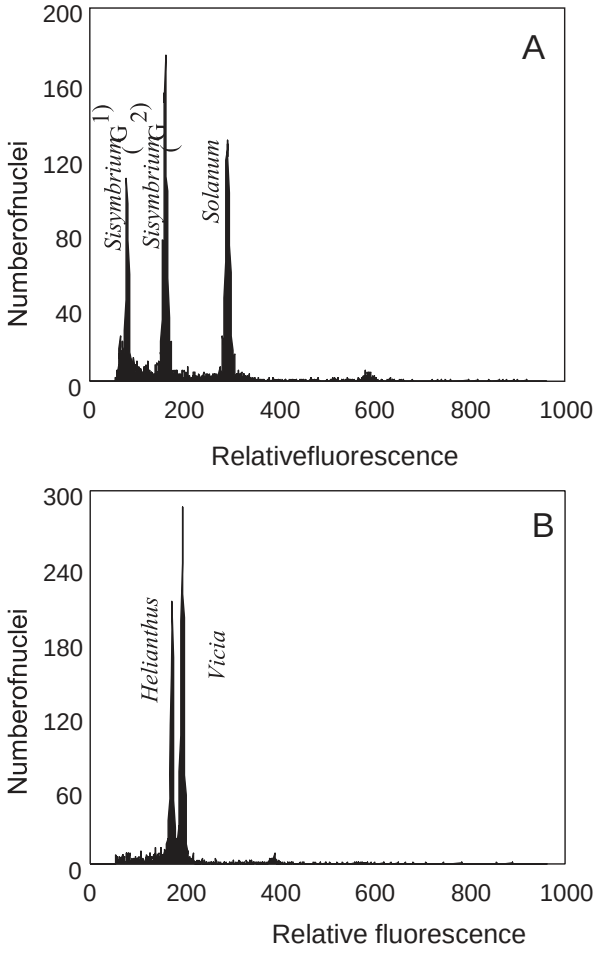


Fig. 1. – Flow cytometric histograms showing genomesize determinationin species with very small (*Sisymbrium altissimum*, 2C = 0.52 pg; panel A) and large (*Helianthus tuberosus*, 2C = 23.89 pg; panel B) genomes. Nuclei of both the analyzed sample and internal reference standard were isolated, stained with propidium iodide and analyzed simultaneously. *Solanum lycopersicum* and *Vicia faba*, respectively, were used as reference standards.

genomes, and less often intermediate to very large genomes than expected by chance (G-test on contingency table: $\chi^2 = 61.15$, df = 2, $P < 0.0001$).

Ploidy levels in the species set analyzed varied from diploid to dodecaploid (Table 1), and monoploid genome sizes (1Cx-values) ranged from 0.12 pg in 8x *Juncus tenuis* to 4.38 in 2x *Virga strigosa* (= 38-fold range).

Factors affecting genome size of naturalized alien species

None of the examined explanatory variables (invasion status, life history, moisture score) had a significant effect on genome size (full model: F = 1.141; df = 11, 80; NS; observed power = 0.486).

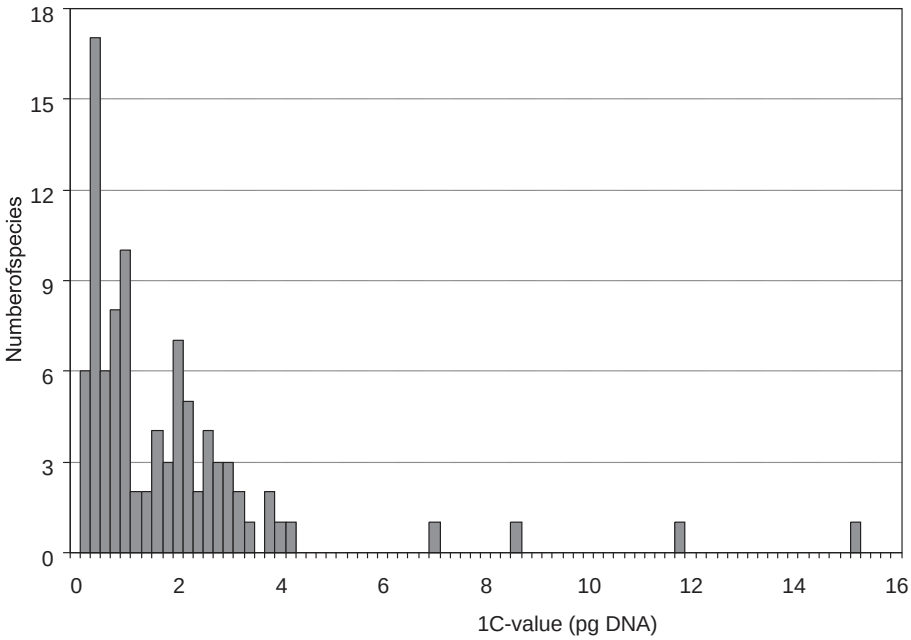


Fig. 2. – Distribution of 1C-values (pg means) in 93 naturalized alien species occurring in the Czech Republic.

Table 2. – Contingency table on observed and expected counts of genome size categories according to Leitch et al. (1998) in naturalized plants of the Czech flora (n = 93) and non-invading species (n = 4148) taken from the Plant DNA C-values database (Bennett & Leitch 2005). Genome size categories "intermediate", "large" and "very large" were pooled together due to small sample sizes in these categories for naturalized species. Statistics are given in the text.

Genome size category	Naturalized species		Non-invading species	
	observed	expected	observed	expected
Very small (1C ≤ 1.40 pg)	49	30.0	1287	1306.7
Small (1C = 1.41–3.50 pg)	36	22.1	958	972.2
Intermediate to very large (1C ≥ 3.51 pg)	8	40.9	1903	1869.1

Genome size in naturalized alien species vs their non-invading relatives

Naturalized aliens had significantly smaller holoploid and monoploid genome sizes than their non-invading congeners (two-sided paired t-tests: 2C-values t = 2.161, df = 45, P = 0.04; Cx-values t = 2.70, df = 44, P = 0.01), and the same held for 2C-values on confamilials(t = 3.161, df = 30, P = 0.004). Visual inspection of data indicatesthat naturalized aliens had smaller genomes in 19 of 31 families considered (Fig. 3).

Table 1. – List of analyzed species, with their family affiliation, life history, LH (an – annual; mono – monocarpic perennial; per – polycarpic perennial), invasion status, moisture score (see Methods for calculation) and genome size characteristics: mean holoploid genome sizes (2C-values with standard deviation, and 1C-values in picograms of DNA and megabase pairs; 1 pg = 978 Mbp), ploidy levels, monoploid genome sizes (1Cx-values), and internal reference standard used (B – *Bellis perennis*; G – *Glycine max*; P – *Pisum sativum*; S – *Solanum lycopersicum*; V – *Vicia faba*; Z – *Zea mays*). Invasion status was taken from Pyšek et al. (2002); species marked with asterisk are classified differently to better reflect situation in sampled localities. Species analyzed for the first time are shown in bold; those used for comparison with non-invading congeners are designated by "+" after species name. Empty cells – data not available.

Species and family	Life history	Invasion status	Moisture score	2C-value (pgDNA)	S.D.	1C-value (pgDNA)	1C-value (Mbp)	Ploidy level	1Cx-value (pgDNA)	Int. st.
Abutilon theophrasti Malvaceae	an	nat.*	4.0	2.17	0.02	1.9	1061.1	6	0.36	B
Amaranthus albus + Amaranthaceae	an	nat.	2.0	1.11	0.00	0.55	540.8	4	0.28	S
Amaranthus blitoides + Amaranthaceae	an	nat.	2.3	1.13	0.00	0.56	552.1	4	0.28	G
Amaranthus powellii + Amaranthaceae	an	invasive	4.0	1.2	0.00	0.51	496.8	4	0.25	B
Amaranthus retroflexus + Amaranthaceae	an	invasive	2.2	1.12	0.01	0.56	546.2	4	0.28	S
Ambrosia artemisiifolia Asteraceae	an	invasive	3.0	2.32	0.03	1.16	1134.5	4	0.58	B
Ambrosia trifida Asteraceae	an	casual	3.7	3.82	0.03	1.91	1865.5	2	1.91	P
Angelica archangelica Apiaceae	mono	invasive	5.0	6.46	0.10	3.23	3158.0	2	3.23	P
Antirrhinum majus Plantaginaceae	mono	nat.	1.0	1.17	0.01	0.58	570.7	2	0.58	S
Arabis alpina Brassicaceae	per	nat.	2.0	0.83	0.02	0.42	406.8	2	0.42	S
Asclepias syriaca Apocynaceae	per	nat.	3.0	0.84	0.03	0.42	409.8	2	0.42	S
Aster lanceolatus + Asteraceae	per	invasive	4.0	5.41	0.01	2.71	2647.0	8	0.68	P
Bidens connata Asteraceae	an	casual	5.0	6.44	0.06	3.22	3150.6	4	1.61	Z
Bidens frondosa Asteraceae	an	invasive	5.0	3.20	0.02	1.60	1566.3	4	0.80	P
Bunias orientalis + Brassicaceae	mono	invasive	2.1	5.40	0.02	2.70	2641.6	2	2.70	P
Cannabis ruderalis + Cannabaceae	an	nat.*	2.5	1.81	0.01	0.91	885.1	2	0.91	S
Cardamine chelidonia + Brassicaceae	mono	nat.	3.0	0.71	0.00	0.36	349.1	4	0.18	S
Claytonia alsinoides Portulacaceae	an	nat.	4.0	6.91	0.05	3.45	3377.0			P

Collomia grandiflora Polemoniaceae	an	nat.	3.0	4.15	0.04	2.8	2030.3	2	2.8	B
Consolida orientalis Ranunculaceae	an	nat.	2.4	4.57	0.04	2.28	2233.3	2	2.28	P
Conyza canadensis Asteraceae	an	invasive	1.8	0.91	0.01	0.45	443.5	2	0.45	S
Corydalis lutea Papaveraceae	per	nat.	1.0	1.26	0.01	0.63	616.1	4	0.32	S
Cuscuta campestris Convolvulaceae	an	invasive	3.0	0.96	0.00	0.48	469.4	8	0.12	S
Datura stramonium + Solanaceae	an	nat.	2.0	4.19	0.05	2.9	2047.4	2	2.9	Z
Digitalis purpurea Plantaginaceae	mono	invasive	2.4	1.86	0.01	0.93	910.0	8	0.23	G
Duchesnea indica Rosaceae	per	nat.	3.0	3.45	0.01	1.72	1686.6	12	0.29	P
Echinocystis lobata Cucurbitaceae	an	invasive	3.0	1.49	0.01	0.74	727.6	4	0.37	S
Echinops sphaerocephalus Asteraceae	per	invasive	2.5	8.2	0.11	4.1	3923.7	4	2.1	Z
Epilobium ciliatum + Onagraceae	per	invasive	4.0	0.82	0.03	0.41	401.0	4	0.21	S
Epilobium dodonaei + Onagraceae	per	nat.	5.0	1.17	0.01	0.59	573.1	4	0.29	S
Erigeron annuus Asteraceae	mono	invasive	2.0	4.33	0.05	2.17	2118.8	3	1.44	Z
Erucastrum gallicum Brassicaceae	mono	nat.	2.0	2.8	0.01	1.4	1018.6	4	0.52	B
Galega officinalis Fabaceae	per	nat.	3.0	4.42	0.04	2.21	2161.4	2	2.21	G
Galinsoga parviflora Asteraceae	an	invasive	3.0	1.47	0.01	0.74	720.3	4	1.1	S
Galinsoga quadriradiata Asteraceae	an	invasive	2.0	4.3	0.04	2.2	1972.6	2	0.74	Z
Geranium pyrenaicum Geraniaceae	per	invasive	3.0	2.87	0.02	1.44	1403.9	2	1.44	G
Helianthus tuberosus + Asteraceae	per	invasive	4.0	23.89	0.09	11.95	11682.7	12	1.99	V
Heracleum mantegazzianum + Apiaceae	mono	invasive	3.0	3.56	0.03	1.78	1740.8	2	1.78	Z
Hesperis matronalis Brassicaceae	per	nat.	3.4	7.61	0.07	3.80	3719.3	4	1.90	P
Hordeum jubatum + Poaceae	an	nat.	2.0	17.38	0.09	8.69	8499.8	4	4.35	P
Chenopodium pumilio + Amaranthaceae	an	nat.	1.0	0.73	0.02	0.37	357.0	2	0.37	P
Chenopodium strictum + Amaranthaceae	an	nat.	3.1	1.60	0.01	0.80	782.4	4	0.40	P
Impatiens glandulifera + Balsaminaceae	an	invasive	3.3	1.90	0.01	0.95	927.6	2	0.95	B

Impatiens parviflora + Balsaminaceae	**an**	invasive	3.0	4.26	0.04	2.13	2083.6	2	2.13	Z
Imperatoria ostruthium Apiaceae	**per**	invasive	3.0	3.89	0.08	1.95	1904.2	2	1.95	P
Inula helenium + Asteraceae	**mono**	nat.	3.0	4.53	0.01	2.26	2214.2	2	2.26	P
Iva xanthiifolia Asteraceae	**an**	nat.	2.0	6.34	0.07	3.17	3098.8	4	1.58	Z
Juncus tenuis + *Juncaceae*	**per**	invasive	3.0	0.92	0.01	0.46	450.9	8	0.12	S
Kochia scoparia Amaranthaceae an	**an**	invasive	2.3	2.23	0.01	1.12	1090.5	2	1.12	G
Lepidium densiflorum + *Brassicaceae* mono	**mono**	nat.	2.0	0.66	0.03	0.33	322.7	4	0.17	S
Lupinus polyphyllus + Fabaceae	per	invasive	3.0	1.60	0.02	0.80	783.4	4	0.40	S
Lychnis coronaria Caryophyllaceae	mono	nat.	2.0	6.30	0.17	3.15	3078.7	2	3.15	Z
Lysimachia punctata Primulaceae	per	nat.	5.0	4.43	0.02	2.21	2165.8	2	2.21	P
Matricaria discoidea + Asteraceae	an	invasive	3.0	4.70	0.01	2.35	2298.3	2	2.35	B
Medicago sativa + Fabaceae	per	nat.*	3.0	3.49	0.03	1.74	1706.1	4	0.87	S
Mimulus guttatus Phrymaceae	per	invasive	5.0	0.73	0.03	0.37	357.0	4	0.18	S
Myrrhis odorata Apiaceae	per	invasive	5.0	1.81	0.01	0.90	883.1	2	0.90	G
Oenothera biennis + Onagraceae	mono	invasive	2.0	2.30	0.02	1.15	1124.7	2	1.15	S
Oenothera glazoviana + Onagraceae	mono	nat.	3.0	2.30	0.01	1.15	1122.7	2	1.15	S
Oxalis corniculata subsp. *repens* + Oxalidaceae	mono	nat.*	2.0	2.11	0.02	1.5	1031.3	8	0.26	S
Oxalis dillenii + Oxalidaceae	mono	nat.	2.0	1.1	0.01	0.50	491.9	4	0.25	G
Oxalis fontana + Oxalidaceae	mono	nat.	3.0	1.22	0.01	0.61	594.1	4	0.30	G
Oxybaphus nyctagineus Nyctaginaceae	per	nat.		1.89	0.01	0.95	924.2	6	0.32	B
Panicum capillare + Poaceae	an	nat.	2.0	0.91	0.01	0.45	443.5	2	0.45	B
Panicum miliaceum + Poaceae	an	casual	4.3	2.9	0.04	1.4	1020.5	4	0.52	B
Phytolacca esculenta + Phytolaccaceae	per	nat.	3.0	5.68	0.10	2.84	2778.5	8	0.71	P
Potentila intermedia + Rosaceae	mono	nat.	1.0	1.80	0.01	0.90	880.2	4	0.45	S
Rudbeckia hirta Asteraceae	per	nat.	2.6	14.33	0.09	7.17	7008.3	4	3.58	P
Rudbeckia laciniata Asteraceae	per	invasive	4.0	30.54	0.12	15.27	14935.5	8	3.82	P

Rumex alpinus + Polygonaceae	per	invasive	3.0	0.96	0.01	0.48	467.0	2	0.48	S
Rumex longifolius + Polygonaceae	per	invasive	2.0	3.99	0.07	2.00	1951.6	6	0.67	B
Rumex patientia + Polygonaceae	per	nat.	2.0	4.87	0.02	2.43	2380.5	6	0.81	Z
Rumex thyrsiflorus + Polygonaceae	per	invasive	1.0	7.81	0.05	3.90	3817.1	2	3.90	P
Scutellaria altissima Lamiaceae	per	nat.	2.0	0.79	0.03	0.40	386.3	4	0.20	S
Sedum hispanicum Crassulaceae	per	invasive	1.0	5.39	0.04	2.70	2636.2			S
Sedum rupestre Crassulaceae	per	nat.	3.0	5.41	0.09	2.70	2643.0	4	1.35	P
Sedum spurium + Crassulaceae	per	nat.	1.0	4.16	0.03	2.8	2034.2	2	2.8	P
Senecio inaequidens + Asteraceae	per	nat.*	3.0	2.90	0.01	1.45	1419.6	4	0.73	B
Senecio vernalis + Asteraceae	an	nat.	1.7	2.33	0.01	1.16	1138.4	2	1.16	B
Setaria faberi + Poaceae	an	nat.	3.0	2.56	0.02	1.28	1253.8	4	0.64	B
Silene dichotoma + Caryophyllaceae	mono	nat.	3.0	5.89	0.02	2.94	2878.7	2	2.94	P
Sisymbrium altissimum Brassicaceae	an	nat.	2.3	0.52	0.00	0.26	255.7	2	0.26	S
Sisymbrium loeselii Brassicaceae	an	invasive	3.5	0.48	0.00	0.24	233.3	2	0.24	S
Sisymbrium strictissimum Brassicaceae	**per**	nat.	4.0	1.39	0.01	0.70	680.2	4	0.35	S
Smyrnium perfoliatum Apiaceae	mono	nat.	3.0	5.64	0.05	2.82	2758.0	2	2.82	P
Solidago canadensis Asteraceae	per	invasive	3.0	2.4	0.01	1.2	999.5	2	1.2	G
Solidago gigantea Asteraceae	per	invasive	4.0	3.65	0.03	1.82	1782.9	4	0.91	Z
Telekia speciosa Asteraceae	per	invasive	3.0	2.57	0.01	1.29	1258.2	2	1.29	S
Trifolium hybridum + Fabaceaea	mono	invasive	2.0	1.9	0.01	0.54	532.0	2	0.54	S
Veronica persica + Plantaginaceae	an	invasive	4.0	1.38	0.01	0.69	672.9	4	0.34	S
Vicia grandiflora + Fabaceae	an	nat.	3.0	6.23	0.07	3.11	3046.0	2	3.11	P
Virga strigosa Dipsacaceae	mono	invasive	3.0	8.76	0.05	4.38	4283.2	2	4.38	Z
Xanthium albinum Asteraceae	an	nat.	5.0	5.18	0.09	2.59	2531.6	4	1.29	P

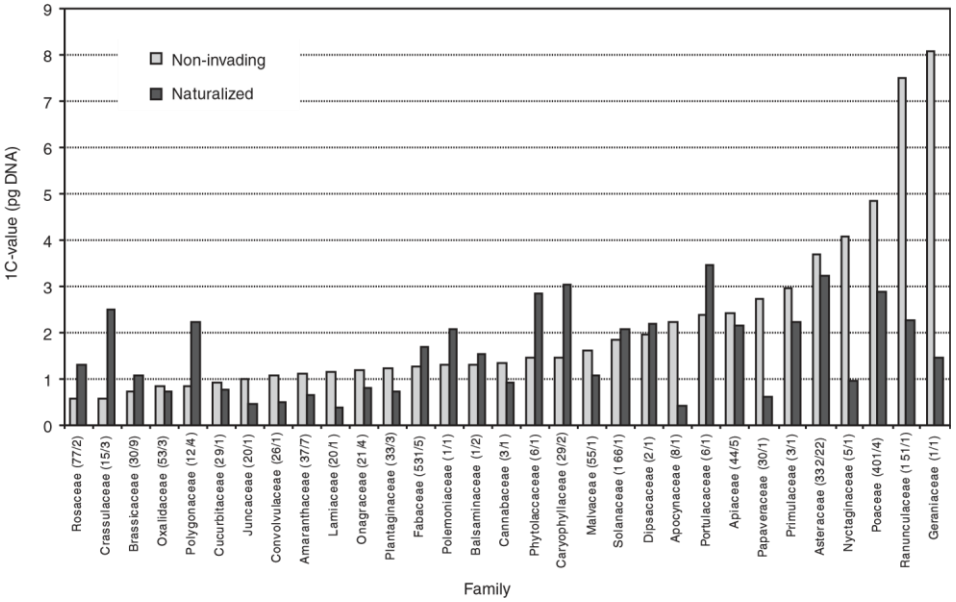


Fig. 3. – Comparison of median genome sizes (1C-values) of naturalized aliens with those of their non-invading confamilials in 31 plant families. Genome sizes for non-invading species were taken from the Plant DNA C-values database (Bennet & Leitch 2005), those known to be naturalized or invasive in any part of the world were excluded. Plant families are sorted according to the genome size of non-invading species. Numbers of species (non-invading/naturalized) are shown in parentheses.

Discussion

Genome size variation

We determined nuclear DNA amounts in a representative setof naturalized plant speciesin the flora of the Czech Republic and compared their values with genome sizes of noninvading species taken from the Plant DNA C-values database (Bennett & Leitch 2005). Out of 93 naturalized species included, 66 (= 71%) were analyzed for the first time. In addition, the first record was obtained for the family *Phrymaceae*, which had a very small genome (*Mimulus guttatus*; 1C = 0.37 pg).

A comparison of genome size values for the same species as determined in our study with those extracted from the database (Bennett & Leitch 2005) revealed some discrepancies. Disregarding variation caused by potential differences in ploidy level, the average absolute difference in C-values was 28%. While about one third of the species (nine out of 26) showed differences below 10% (which is within the acceptable between-laboratory limit as suggested by Doležel et al. 1998), four species differed more than 1.5-fold (*Amaranthus retroflexus*, *Solidago canadensis*, *Galinsoga parviflora* and *Antirrhinum majus*). It should, however, be noted that the difference largely depended on the methodology used. The best congruency between our data and the database values was observed for measurements performed using propidium iodide flow cytometry (absolute difference < 10%, n = 4), which is generally recommended as the most reliable technique for genome size estimation in plants (Greilhuber et al. 2007, Temsch et al. 2010). Species

analyzed either by Feulgen densitometry (which is much more sensitive to working conditions; Greilhuber 2005) or flow cytometry with base-selective fluorochromes (Doležel et al. 1992) showed higher differences (30%, n = 21 and 54%, n = 1, respectively). Whereas small differences in determined genome sizes can be explained by minor variation in adopted protocols between laboratories (use of different buffers, different internal reference standards, etc.), it is possible that more serious methodological flaws were involved in other cases and such data should be treated with caution.

Very small and small genomes (in the sense of the classification of genome sizes defined by Leitch et al. 1998) clearly prevailed in our data set, accounting for 53 and 39% of the species total, respectively. This is a dramatically different frequency of individual genome size categories as compared to non-invading species (Table 2). For example, while intermediate genomes (1C = 3.51–14.0 pg) are as common as the very small ones (1C ≤ 1.4 pg) in non-invading plants, their ratio drops down to only one seventh of the frequency of very small genomes in naturalized aliens. Significantly smaller genomes in naturalized plants as compared to their phylogenetically related non-invading counterparts were also confirmed at both taxonomic levels tested, i.e., the rank of genus and family. The same pattern of genome size variation as in naturalized plants (i.e. predominance of species with low nuclear DNA amounts and the lack of DNA-obese species) was observed, for instance, in weedy plants (Bennett et al. 1998) or in endemics on oceanic islands (Suda et al. 2005). Selection for rapid development, fast growth, and production of many light and easily dispersible seeds are plausible evolutionary forces that constrain the genome size in these plant groups.

The presence of species with different life histories allowed us to test the potential association between genome size and life history (Bennett 1972). Although basic descriptive statistics for 36 polycarpic perennial plants (mean and median 1C-values 2.44 pg and 1.73 pg, respectively) somehow differed from corresponding values for both 36 annuals (mean = 1.58 pg, median = 1.10 pg) and 21 monocarpic species (mean = 1.65 pg, median = 1.15 pg), the differences were not statistically significant.

Relationship between genome size and invasion success

By using the data set analyzed in this study, we were unable to detect the effect of any species traits examined on the genome size. However, it should be noted that the low test power of the model does not allow us to conclude that such effects do not exist. The analysis was carried out with individual species as random independent data points. This can inflate degrees of freedom, because the species can be mutually dependent due to their phylogenetic relatedness (e.g. Harvey & Pagel 1991). However, removing phylogenetic effects from the nonsignificant general linear model would require methods based on eigenvector filtering (Diniz-Filho et al. 1998) and repeating the analysis after this correction could thus only further decrease the statisticalsignificanceof the results. That genome size is associated with invasion success is clearly demonstrated by the comparison of naturalized aliens in the Czech flora with the reference global set of non-invading congeners. In this analysis the effect of phylogenetic relatedness, which was shown to bias the effect of traits on species' invasion success (e.g. Crawley et al. 1996, Pyšek 1997, Goodwin et al. 1999, Grotkopp et al. 2004, Hamilton et al. 2005, Lloret et al. 2005, Cadotte et al. 2006) was suppressed by the congeneric comparison. This is a convenient approach to studying the role of species traits in plant invasions (Pyšek & Richardson 2007, Perglová et al. 2009).

Our results therefore provide robust evidence, based on a large number of species across a wide range of plant families, that alien species that successfully naturalize have smaller genomes than those that do not reach the stage of naturalization. It should be, however, noted that sample size for some families was rather limited (e.g. only one naturalized and/or non-invading species was available for 18 out of 31 families used for comparison), which may limit the generality of our conclusions and calls for further comparative studies. The association of small genome size with invasiveness was previously suggested in a number of studies (Rejmánek 1996, 2000, Grotkopp et al. 2002, Rejmánek et al. 2005, Garcia et al. 2008). Small genome size seems to be a result of selection for short minimum generation time.It isalso associatedwith smallseed size,high leafarea ratio and high relative growth rate of seedlings (Grotkopp et al. 2002), and as such may be an ultimate determinant of plant species invasiveness in disturbed habitats (Rejmánek 1996, Bennett et al. 1998, Grotkopp et al. 1998, Rejmánek 2000).

However, studies that addressed the role of genome size in plant species' invasiveness usually compared invasive and non-invasive species and did not distinguish between species at different stages of the invasion process. In our data there was no difference in the genome size of invasive species compared to naturalized but non-invasive species. This indicates that the small genome size may provide alien plants with an advantage already at the stage of naturalization and need not necessarily play a role during the follow-up step, transition from naturalized to invasive species. It also points to the importance of distinguishing the stages of invasion in such studies since the determinants of invasion success may differ between stages (Williamson 2006, Pyšek et al. 2008, 2009a, b).

See http://www.preslia.cz for Electronic Appendix 1.

Acknowledgements

The study was funded by grants no. 206/05/0323, 206/09/0563 and 206/08/H049 from the Czech Science Foundation, long-term research plans no. AV0Z60050516(from the Academy of Sciences of the Czech Republic) and no. MSM0021620828, project LC06073 (both from the Ministry of Education, Youth and Sports of the Czech Republic), and no. 261 211 (Grant Agency of Charles University). We thank Laura Meyerson and Johann Greilhuberfor helpful comments on the manuscript,Ewald Weber for providingus with an updatedversion of his database of alien plants of the world, Milan Chytrý for consultation on habitat moisture classification, David Richardson for improving our English, Vendula Havlíčková for technical assistance and all colleagues who provided us with locality details and/or helped with seed collection.

References

Bennett M. D. (1972): Nuclear DNA content and minimum generation time in herbaceous plants. – Proc. R. Soc. London B 181: 109–135.

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

Bennett M. D. & Leitch I. J. (2005): Plant DNA *C*-values database (release 4.0, Oct. 2005). – URL: [http://www.kew.org/cvalues/].

Bennett M. D., Leitch I. J. & Hanson L. (1998): DNA amounts in two samples of angiosperm weeds. – Ann. Bot. 82: 121–134.

Binimelis R., Born W., Monterroso I. & Rodríguez-Labajos B. (2007): Socio-economic impacts and assessment of biological invasions. – In: Nentwig N. (ed.), Biological invasions, p. 331–347, Springer, Berlin.

Blackburn T. M., Lockwood J. L. & Cassey P. (2009):Avian invaders: the ecologyand evolutionof exotic birds. – Oxford University Press, Oxford.

Cadotte M. W., Murray B. R. & Lovett-Doust J. (2006): Evolutionary and ecological influences of plant invader success in the flora of Ontario. – Ecoscience 13: 388–395.

Chytrý M., Maskell L. C., Pino J., Pyšek P., Vilŕ M., Font X. & Smart S. M. (2008): Habitat invasions by alien plants: a quantitative comparison among Mediterranean, subcontinental and oceanic regions of Europe. – J. Appl. Ecol. 45: 448–458.

ChytrýM.,PyšekP.,TichýL.,KnollováI.&DanihelkaJ.(2005):InvasionsbyalienplantsintheCzechRepublic: a quantitative assessment across habitats. – Preslia 77: 339–354.

Chytrý M., Wild J., Pyšek P., Tichý L., Danihelka J. & Knollová I. (2009): Maps of the level of invasion of the Czech Republic by alien plants. – Preslia 81: 187–207.

Crall A. W., Meyerson L. A., Stohlgren T. J., Jarnevich C. S., Newman G. J. & Graham J. (2006): Show me the numbers: what data currently exist for non-native species in the USA? – Front. Ecol. Environ. 4: 414–418.

Crawley M. J. (1993): GLIM for Ecologists. – Blackwell, Oxford.

Crawley M. J. (2002): Statistical computing: an introduction to data analysis using S-Plus. – Wiley, Chichester.

Crawley M. J., Harvey P. H. & Purvis A. (1996): Comparative ecology of the native and alien floras of the British Isles. – Biol. Trans. R. Soc. B 351: 1251–1259.

DAISIE (2009): Handbook of alien species in Europe. – Springer, Berlin.

Davis M. A. (2009): Invasion biology. – Oxford University Press, Oxford.

Davis M. A., Grime J. P. & Thompson K. (2000): Fluctuating resources in plant communities: a general theory of invasibility. – J. Ecol. 88: 528–534.

Diniz-Filho J. A. F., De Santana C. E. R. & Bini L. M. (1998): An eigen vecto rmethod for estimating phylogenetic inertia. – Evolution 52: 1247–1262.

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. – Ann. Bot. 82 (Suppl. A): 17–26.

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

Doležel J., Sgorbati S. & Lucretti S. (1992): Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. – Physiol. Plant. 85: 625–631.

Ekrt T., Trávníček P., Jarolímová V., Vít P. & Urfus T. (2009): Genome size and morphology of the *Dryopteris affinis* group in Central Europe. – Preslia 81: 261–280.

Ellenberg H., Weber H. E., Düll R., Wirth W., Werner W. & Paulißen D. (1992): Zeigerwerte von Pflanzen in Mitteleuropa. Ed. 2. – Scripta Geobotanica 18: 1–258.

Essl F., Dullinger D. & Kleinbauer I. (2009): Changes in the spatio-temporal patterns and habitat preferences of *Ambrosia artemisiifolia* during the invasion of Austria. – Preslia 81: 119–133.

Gaertner M., Breeyen A. D., Hui C. & Richardson D. M. (2009): Impacts of alien plant invasions on species richness in Mediterranean-type ecosystems: a meta-analysis. – Progr. Phys. Geogr. 33: 319–338.

Garcia S., Canela M. A., Garnatje T., McArthur E. D., Pellicer J., Sanderson S. C. & Vallés J. (2008): Evolutionary and ecological implications of genome size in the North American endemic sagebrushes and allies (*Artemisia*, *Asteraceae*). – Biol. J. Linn. Soc. 94: 631–649.

Goldblatt P. & Johnson D. E. (eds) (1979 onwards): Index to plant chromosome numbers. – Missouri Botanical Garden, St. Louis, URL: [http://mobot.mobot.org/W3T/Search/ipcn.html].

Goodwin B. J., McAllister A. J. & Fahrig L. (1999): Predicting invasiveness of plant species based on biological information. – Conserv. Biol. 13: 422–426.

Gregory T. R. (2005): The C-value enigma in plants and animals: a review of parallels and an appeal for partnership. – Ann. Bot. 95: 133-146.

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

Greilhuber J., Doležel J., Lysák M. & Bennett M. D. (2005): The origin, evolution and proposed stabilization of the terms 'genome size' and 'C-value' to describe nuclear DNA contents. – Ann. Bot. 95: 255–260.

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, Wiley-VCH, Weinheim.

Grotkopp E., Rejmánek M. & Rost T. L. (2002): Toward a causal explanation of plant invasiveness: seedling growth and life-history strategies of 29 pine (*Pinus*) species. – Am. Nat. 159: 396–419.

Grottkop E., Rejmánek M., Sanderson M. J. & Rost T. L. (2004): Evolution of genome size in pines (*Pinus*) and its life-history correlates: supertree analyses. – Evolution 58: 1705–1729.

GrotkoppE., Stoltenberg R., Rejmánek M. & Rost T. L. (1998):The effect of genomesize on invasiveness. – Am. J. Bot. 85 (Suppl.): 34.

Hamilton M. A., Murray B. R., Cadotte M. W., Hose G. C., Baker A. C., Harris C. J. & Licari D. (2005):Life-history correlates of plant invasiveness at regional and continental scales. – Ecol. Lett. 8: 1066–1074.

Harvey P. H. & Pagel M. D. (1991): The comparative method in evolutionary biology. – Oxford University Press, Oxford.

Hejda M., Pyšek P. & Jarošík V. (2009a):Impactofinvasive plantsonthe species richness,diversity andcomposition of invaded communities. – J. Ecol. 97: 393–403.

Hejda M., Pyšek P., Pergl J., Sádlo J., Chytrý M. & Jarošík V. (2009b): Invasion success of alien plants: do habitats affinities in the native distribution range matter? – Glob. Ecol. Biogeogr. 18: 372–382.

Hejný S. & Slavík B. (eds) (1988–1992): Květena ČR [Flora of the Czech Republic]. Vol. 1 (1988), 2 (1990), 3 (1992). – Academia, Praha.

Hulme P. E., Bacher S., Kenis M., Klotz S., Kühn I., Minchin D., Nentwig W., Olenin S., Panov V., Pergl J., Pyšek P., Roques A., Sol D., Solarz W. & Vilà M. (2008): Grasping at the routes of biological invasions: a framework for integrating pathways into policy. – J. Appl. Ecol. 45: 403–414.

Hulme P. E., Nentwig W., Pyšek P. & Vilà M. (2009a): Common market, shared problems: time for a coordinated response to biological invasions in Europe? – Neobiota 8: 3–19.

Hulme P., Pyšek P., Nentwig W. & Vilà M. (2009b): Will threat of biological invasions unite the European Union? – Science 324: 40–41.

Keller R., LodgeD. M. & Finnoff D. (2007):Risk assessment for invasive species producesnet bioeconomicbenefits. – Proc. Natl Acad. Sci. USA 104: 203–207.

Kettunen M., Genovesi P., Gollasch S., Pagad S., Starfinger U., ten Brink P. & Shine C. (2009):Technical support to EU strategy on invasive species (IAS): assessment of the impacts of IAS in Europe and the EU (final module report for the European Commission). – Institute for EuropeanEnvironmental Policy, Brussels.

Knight C. A., Molinari N. A. & Petrov D. A. (2005): The large genome constraint hypothesis: evolution, ecology and phenotype. – Ann. Bot. 95: 177–190.

Kron P., Suda J. & Husband B. C. (2007): Applications of flow cytometry to evolutionary and population biology. – Annu. Rev. Ecol. Evol. Syst. 38: 847–876.

Lambdon P. W., Pyšek P., Basnou C., Hejda M., Arianoutsou M., Essl F., Jarošík V., Pergl J., Winter M., Anastasiu P., Andriopoulos P., Bazos I., Brundu G., Celesti-Grapow L., Chassot P., Delipetrou P., Josefsson M., Kark S., Klotz S., Kokkoris Y., Kühn I., Marchante H., Perglová I., Pino J., Vilà M., Zikos A., Roy D. & Hulme P. E. (2008): Alien flora of Europe: species diversity, temporal trends, geographical patterns and research needs. – Preslia 80: 101–149.

Lawrence M. E. (1985): *Senecio* L. (*Asteraceae*) in Australia: nuclear DNA amounts. – Aust. J. Bot. 33: 221–232.

Leitch, I. J. & Bennett M. D. (2007). Genome size and its uses: the impact of flow cytometry. – In: Doležel J., Greilhuber J. & Suda J. (eds), Flow cytometry with plant cells, p. 153–176, Wiley-VCH, Weinheim.

Leitch I. J., Chase M. W. & Bennett M. D. (1998): Phylogenetic analysis of DNA C-values provides evidence for a small ancestral genome size in flowering plants. – Ann. Bot. 82 (Suppl. A): 85–94.

Levine J. M., Vilà M., D'Antonio C. M., Dukes J. S., Grigulis K. & Lavorel S. (2003): Mechanisms underlying the impacts of exotic plant invasions. – Proc. R. Soc. London B 270: 775–781.

Lloret F., Médail F., Brundu G., Camarda I., Moragues E., Rita J., Lambdon P. & Hulme P. E. (2005): Species attributes and invasion success by alien plants on Mediterranean islands. – J. Ecol. 93: 512–520.

Loureiro J., Trávníček P., Rauchová J., Urfus T., Vít P., Štech M., Castro S. & Suda J. (2010): The use of flow cytometry in the biosystematics, ecology and population biology of homoploid plants. – Preslia 82: 3–21.

MarholdK., Mártonfi P., Mereďa jun. P., Mráz P., Hodálová I., KolníkM., Kučera J., Lihová J., Mrázová V., Perný M. & Valko I. (2007): Karyologická databáza papraďorastov a semenných rastlín Slovenska [Karyological database of the ferns and flowering plants of Slovakia], ver. 1. – URL: [http://www.chromosomes.sav.sk].

McGeoch M. A., Butchart S. H. M., Spear D., Marais E., Kleynhans E. J., Symes A., Chanson J. & Hoffmann M. (2010): Global indicators of biologica linvasion: species numbers, biodiversity impact and policy responses. – Diversity Distrib. 16: 95–108.

Meyerson L. A. & Mooney H. A. (2007): Invasive alien species in an era of globalization. – Front. Ecol. Environ. 5: 199–208.

Mukherjee S. & Sharma A. K. (1990): Mitotic cycle duration and its relationship with nuclear DNA content, chromosome number and size and rate of growth in species of *Acacia* growing under stress. – Ind. J. Exp. Biol. 28: 508–510.

Nentwig W., Kühnel E. & Bacher S. (2010): A generic impact-scoring system applied to alien mammals in Europe. – Cons. Biol. 204: 302–311.

OttoF. (1990): DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. – In: CrissmanH. A. & Darzynkiewicz Z. (eds), Methods in cell biology 33: 105–110, Academic Press,New York.

Perglová I., Pergl J., Skálová H., Moravcová L., Jarošík V. & Pyšek P. (2009): Differences in germination and seedling establishment of alien and native *Impatiens* species. – Preslia 81: 357–375.

Pyšek P. (1997): Clonality and plant invasions: can a trait make a difference? – In: de Kroon H. & van Groenendael J. (eds), The ekology and evolution of clonal plants, p.405–427, Backhuys Publishers, Leiden.

Pyšek P., Jarošík V., Pergl J., Randall R., Chytrý M., Kühn I., Tichý L., Danihelka J., Chrtek jun. J. & Sádlo J. (2009a): The global invasion success of Central European plants is related to distribution characteristics in their native range and species traits. – Diversity Distrib. 15: 891–903.

Pyšek P., Křivánek M. & Jarošík V. (2009b): Planting intensity, residence time, and species traits determine invasion success of alien woody species. – Ecology 90: 2734–2744.

Pyšek P. & Richardson D. M. (2007):Traits associated with invasiveness in alien plants: where do we stand? – In: Nentwig W. (ed.), Biological invasions, p. 97–125, Springer-Verlag, Berlin & Heidelberg.

Pyšek P., Richardson D. M. & Jarošík V. (2006): Who cites who in the invasion zoo: insights from an analysis of the most highly cited papers in invasion ecology. – Preslia 78: 437–468.

Pyšek P., Richardson D. M., Pergl J., Jarošík V., Sixtová Z. & Weber E. (2008): Geographical and taxonomic biases in invasion ecology. – Trends Ecol. Evol. 23: 237–244.

Pyšek P., Richardson D. M., Rejmánek M., Webster G., Williamson M. & Kirschner J. (2004): Alien plants in checklists and floras: towards better communication between taxonomists and ecologists. – Taxon 53: 131–143.

Pyšek P., Sádlo J. & Mandák B. (2002): Catalogue of alien plants of the Czech Republic. – Preslia 74: 97–186.

Rejmánek M. (1996): A theory of seed plant invasiveness: the first sketch. – Biol. Cons. 78: 171–181.

Rejmánek M. (2000): Invasive plants: approaches and predictions. – Austral Ecol. 25: 497–506.

Rejmánek M., Richardson D. M. & Pyšek P. (2005): Plant invasions and invasibility of plant communities. – In: Van der Maarel E. (ed.), Vegetation ecology, p. 332–355, Blackwell Science, Oxford.

Ricciardi A. & MacIsaac H. J. (2008): The book that began invasion ecology. – Nature 452: 34.

Richardson D. M., Holmes P. M., Esler K. J., Galatowitsch S. M., Stromberg J. C., Kirkman S. P., Pyšek P. & Hobbs R. J. (2007): Riparian vegetation: degradation, alien plant invasions, and restoration prospects. – Diversity Distrib. 13: 126–139.

Richardson D. M. & Pyšek P. (2008): Fifty years of invasion ecology: the legacy of Charles Elton. – Diversity Distrib. 14: 161–168.

Richardson D. M., Pyšek P., Rejmánek M., Barbour M. G., Panetta F. D. & West C. J. (2000): Naturalization and invasion of alien plants: concepts and definitions. – Diversity Distrib. 6: 93–107.

Simberloff D. (2009): We can eliminate invasions or live with them: successful management projects. – Biol. Inv. 11: 149–157.

Slavík B. (ed.) (1995–2000):Květena ČR [Flora of the Czech Republic].Vol. 4 (1995),5 (1997),6 (2000).– Academia, Praha.

Slavík B. & Štěpánková J. (eds) (2004): Květena ČR 8 [Flora of the Czech Republic. Vol. 8]. – Academia, Praha.

Sokal R. R. & Rohlf F. J. (1995): Biometry. Ed. 3. – Freeman, New York.

Štajerová K., Šmilauerová M. & Šmilauer P. (2009): Arbuscular mycorrhizal symbiosis of herbaceous invasive neophytes in the Czech Republic. – Preslia 81: 341–355.

Steidel R. J. & Thomas L. (2001): Power analysis and experimental design. – In: Scheiner S. M. & Gurevitch J. (eds), Design and analysis of ecological experiments, Ed. 2., p. 14–36, Oxford University Press, Oxford.

Stevens P. F. (2001 onwards): Angiosperm Phylogeny Website. Version 9, June 2008. – URL: [http://www.mobot.org/MOBOT/Research/APweb/welcome.html].

Suda J., Kyncl T. & Jarolímová V. (2005): Nuclear DNA amounts in Macaronesian angiosperms: forty percent of Canarian endemic flora completed. – Plant Syst. Evol. 252: 215–238.

Temsch E. M., Greilhuber J. & Krisai R. (2010): Genome size in liverworts. – Preslia 82: 63–80.

Vidic T., Greilhuber J., Vilhar B. & Dermastia M. (2009): Selective significance of genome size in a plant community with heavy metal pollution. – Ecol. Appl. 19: 1515–1521.

Vilà M., Basnou C., Pyšek P., Josefsson M., Genovesi P., Gollasch S., Nentwig W., Olenin S., Roques A., Roy D., Hulme P. E. & DAISIE partners (2010):How well do we understand the impacts of alien species on ecological services? A pan-European cross-taxa assessment. – Front. Ecol. Environ. 8: 135–144.

Wakamiya I., Newton R. J., Johnston J. S. & Price H. J. (1993): Genome size and environmental factors in the genus *Pinus*. – Am. J. Bot. 80:1235–1241.

Weber E. (2003): Invasive plant species of the world: a reference guide to environmental weeds. – CAB International Publishing, Wallingford.

Williamson M. (2006):Explaining and predicting the success of invading species at different stages of invasion. – Biol. Inv. 8: 1561–1568.

Case study VI

Kolář F., Lučanová M., Těšitel J., Loureiro J. & Suda J. (2012): **Glycerol-treated nuclear suspensions – an efficient preservation method for flow cytometric analysis of plant samples. – Chromosome Research 20(2): 303-315 https://doi.org/10.1007/s10577-012-9277-0**

Glycerol-treated nuclear suspensions – an efficient preservation method for flow cytometric analysis of plant samples

Filip Kolář[1,2], Magdalena Lučanová[2,1], Jakub Těšitel[3], João Loureiro[4] & Jan Suda[1,2]

[1] Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, 128 01 Prague, Czech Republic e-mail: filip.kolar@gmail.com
[2] Institute of Botany, Academy of Sciences of the Czech Republic, Zámek 1, 252 43 Průhonice, Czech Republic
[3] Department of Botany, Faculty of Science, University of South Bohemia, Branišovská 31, 370 05 České Budějovice, Czech Republic
[4] Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, P.O. Box 3046, 3001-455 Coimbra, Portugal

Abstract
Flow cytometry (FCM) has been widely used in plant science to determine the amount of nuclear DNA, either in absolute units or in relative terms, as an indicator of ploidy. The requirement for fresh material in some applications, however, limits the value of FCM in field research, including plant biosystematics, ecology and population biology. Dried plant samples have proven to be a suitable alternative in some cases (large-scale ploidy screening) although tissue dehydration is often associated with a decrease in the quality of FCM analysis. The present study tested, using timescale laboratory and in situ field experiments, the applicability of glycerol-treated nuclear suspension for DNA flow cytometry. We demonstrate that plant nuclei preserved in ice-cold buffer + glycerol solution remain intact for at least a few weeks and provide estimates of nuclear DNA content that are highly comparable and of similar quality to those obtained from fresh tissue. The protocol is compatible with both DAPI and propidium iodide staining, and allows not only the determination of ploidy level but also genome size in absolute units. Despite its higher laboriousness, glycerol-preserved nuclei apparently represent the most reliable way of sample preservation for genome size research. We assume that the protocol will provide a vital alternative to other preservation methods, especially when stringent criteria on the quality of FCM analysis are required.

Keywords: DAPI . flow cytometry . genome size . ploidy . propidium iodide . sample preservation

Abbreviations

CV	Coefficient of variation
DAPI	4',6-diamidino-2-phenylindole
FCM	Flow cytometry/flow cytometric
PI	Propidium iodide
SD	Standard deviation

Introduction

Flow cytometry is a high-throughput, cost-effective and accurate cytogenetic method with broad applications in plant sciences (Doležel et al. 2007a). In plant evolutionary and ecological studies, flow cytometry (FCM) has played a prominent role in the estimation of nuclear DNA content, either in absolute units (picograms of DNA or mega base pairs) or in relative terms as an indicator of ploidy (Kron et al. 2007). By enabling the analysis of large population samples over a short time span, FCM has significantly advanced our knowledge of ploidy and genome size variation in natural systems (Kron et al. 2007; Suda et al. 2007a; Loureiro et al. 2010). The global application of FCM is, however, limited by the need to use fresh plant material for reliable measurements of DNA content (Doležel et al. 2007a; Greilhuber et al. 2007). This prerequisite hampers largescale population studies in regions without easily accessible FCM facilities (e.g. most of the tropical regions) and may cause difficulties elsewhere (e.g. when the capacity of a laboratory is saturated or during instrument maintenance). Transport of seeds and their direct FCM measurement or analysis of growing seedlings can partly solve the problem (Sliwinska et al. 2005; Suda et al. 2005). Nevertheless, this approach may be hampered by the need to collect plant material during the seed gathering season, difficulties in seed germination ex situ, potential shifts in genome size values estimated from dry seeds (e.g. Sliwinska et al. 2005) and/or by taxonomic complexity (e.g. hybrid origin of the seeds). Tissue preservation using chemical fixatives (ethanol- or formaldehyde-based), although widely used in animal and human FCM studies, have elicited only little interest from the plant community (Kron et al. 2007). The last years have seen several attempts to substitute fresh plant samples with dry or frozen tissue. Suda and Trávníček (2006a, b) introduced a protocol for reliable ploidy estimation (using 4',6-diamidino-2-phenylindole (DAPI)) in desiccated plant material (either silica- or air-dried) and this method has been successfully applied to a number of plant groups (e.g. Šmarda et al. 2005; Suda et al. 2007b; Popp et al. 2008; Hülber et al. 2009; Košnar and Kolář 2009; Volkova et al. 2010). Reasonable FCM histograms can also be achieved by analysing rapidly frozen plant tissues (Dart et al. 2004; Nsabimana and Van Staden 2006; Halverson et al. 2008; Cires et al. 2009).

Despite these promising steps towards the routine use of non-fresh plant samples, the above-mentioned protocols are still considered inadequate for estimating genome size in absolute units (using intercalating fluorochromes). The quality of measurements from dry or frozen tissue samples only rarely reaches the level achievable for fresh material and further decreases with the ageing of the samples. Usual symptoms accompanying the analysis of non-fresh material are: (1) lower reliability of DNA content estimates as evidenced by pronounced shifts in fluorescence intensity compared to that of fresh samples (e.g. Šmarda 2006; Suda and Trávníček 2006b; Cires et al. 2009; Bainard et al. 2011), and (2) decrease in the uniformity of fluorescence, resulting in higher coefficient of variations (CVs) of the peaks and more prominent background (Suda and Trávníček 2006b). These observations are not compatible with the high standards required for some FCM applications, including the determination of absolute genome size (Doležel and Bartoš 2005; Doležel et al. 2007b; Greilhuber et al. 2007). Consequently, the majority of published studies have used non-fresh material solely to determine DNA ploidy levels, which can tolerate some relaxation of the quality criteria (e.g. Eidesen et al. 2007; Schönswetter et al. 2007a; Popp et al. 2008; Bendiksby et al. 2011), or interpret the results as supplementary to fresh tissue analysis (Dušková et al. 2010). Only very recently, Bainard et al. (2011) conducted a careful experimental study to evaluate the potential of silica-dried plant material for genome size research. The authors

concluded that sample desiccation introduced comparatively minor variation (<10%), a level of which was species-specific and comparable to other sources of artefactual variation. They considered dehydrated plant samples promising for assessing absolute genome size, yet admitted that relaxed demands should be applied to the quality of analysis and caution must be exercised in interpreting the results.

Whereas the effects of physical preservation on FCM estimates of nuclear DNA content have been intensively studied and there are some comparative studies showing advantages and limitations of these approaches (Suda and Trávníček 2006b; Bainard et al. 2011), the potential of chemical fixatives has been largely neglected. This reluctance likely stems from the higher laboriousness of the protocols and potential chemical-induced changes in chromatin condensation, which can affect the stoichiometric staining of DNA using intercalating dyes (Shapiro 2003). The search for alternative modes of preservation is desirable in order to (1) increase the accuracy of nonfresh tissue measurements and (2) extend FCM measurements to species in which other preservation techniques have failed (according to our knowledge, around 15% of plant species do not yield acceptable FCM histograms after dehydration). A promising alternative to physical and chemical preservation of plant tissues is the storage of isolated nuclear suspensions in intact protective solutions such as glycerol (propane-1,2,3-triol). The value of glycerol for the preservation of isolated nuclei for FCM analysis was first mentioned by Chiatante et al. (1990), and a more thorough evaluation of the method was developed by Hopping (1993). This researcher stored isolated nuclei of *Actinidia deliciosa* in 30% glycerol at approximately −20°C and found that storage for 9 months did not compromise FCM analysis, and the estimated values were highly comparable with those obtained from fresh samples (fluorescence decrease of 5% to 7%). Unfortunately, his results were only based on the analysis of a single plant species and were not subjected to a rigorous statistical evaluation.

This study aimed to investigate the applicability of glycerol-preserved plant nuclei for genome size research and assess the capabilities and limitations of this approach. We conducted two complementary experiments: (1) a time-scale laboratory experiment using six model plant species from different families and covering a range of genome sizes, to systematically compare the glycerol-based protocol with other currently used methods of sample preservation, and (2) a multi-species experiment using a set of tropical species collected and preserved in the field, to test the feasibility of the methodology in situ. The effects of fluorescent dyes with different modes of DNA binding (AT-selective DAPI vs. intercalating propidium iodide (PI)) were also investigated.

Methods

Plant material

Six plant species from five plant families and spanning nearly 18-fold range of genome sizes (from 1.52 pg/2C to 26.9 pg/2C) were selected for a time-scale laboratory experiment. They included both frequently used FCM reference standards (*Bellis perennis* – Asteraceae, *Pisum sativum* 'Ctirad' – Fabaceae and *Vicia faba* 'Inovec' – Fabaceae) and representatives of the major tropical families analysed in the second experiment (*Euphorbia milii* – Euphorbiaceae, *Ficus elastica* – Moraceae and *Galium album* – Rubiaceae). This species selection comprised both plants with soft and rapidly decaying leaves (*B. perennis*, *G. album* and *P. sativum*) and plants with rather

tough or even leathery leaves (*E. milii*, *F. elastica* and *V. faba*). Plants of *P. sativum* and *V. faba* were grown from seeds (kindly provided by J. Doležel, Institute of Experimental Botany, Olomouc, the Czech Republic) while the remaining species were available from the living collection of the Botanical Garden, Faculty of Science, Charles University in Prague, Czech Republic.

The in situ experiment involved 21 species from 12 angiosperm families (both monocots and dicots) that were collected in primary and secondary rainforests around the Wannang village (approximately 60 km west of Madang) in northern Papua New Guinea in August 2006 (see Table S1). Representatives of speciesrich tropical genera (e.g. *Ficus* and *Macaranga*) as well as economically important crops (e.g. *Musa*, *Strychnos* and *Syzigium*) were included. Herbarium vouchers are kept in CBFS.

Sample preservation

Four different methods of sample preservation were tested in a time-scale experiment: (1) young healthy leaves stored in a moist plastic bag at 4°C in a refrigerator (further referred to as 'plastic bag'), (2) leaf tissue rapidly dehydrated using silica gel ('silica gel'), (3) isolated nuclei suspended in Otto I buffer + glycerol (see below) and kept at −18°C in a freezer ('ice-cold glycerol'; note that the solution remained liquid at this temperature), and (4) isolated nuclei kept in the same solution at room temperature (23±2°C; 'RT glycerol'). In addition, fresh leaves picked from the cultivated plants were used as 'control'.

The in situ experiment involved the same preservation methods except for the silica gel treatment. For all the used treatments, there was an approximately 24-h interruption in low-temperature storage due to sample transportation from Papua New Guinea to the Czech Republic.

FCM analysis

Sample preparation generally followed the simplified two-step procedure using Otto's buffers (Doležel et al. 2007b). Briefly, ~50 mg of sample leaf tissue and the same amount of the fresh internal reference standard were chopped with a sharp razor blade in a Petri dish containing 0.5 mL of ice-cold Otto I buffer (0.1 M citric acid, 0.5% Tween-20) (Otto 1990). The suspension was filtered through a 42-μm nylon mesh and incubated for approximately 15 min at room temperature. Samples were then stained for 10 min at room temperature. The staining solution consisted of 1 mL of Otto II buffer (0.4 M $Na_2HPO_4 \cdot 12\ H_2O$), β-mercaptoethanol (final concentration of 2 μL mL^{-1}) and a fluorochrome. Two DNA-binding fluorochromes were employed: (1) intercalating PI plus RNase IIA (both at final concentrations of 50 μg mL^{-1}) and (2) AT-selective DAPI at a final concentration of 4 μg mL^{-1}. Stained nuclei were run on a flow cytometer and excited either with (1) a green diode-pumped solid-state laser (Cobolt Samba, 532 nm, 100 mW; Cobolt, Stockholm, Sweden) embedded in a Partec CyFlow SL instrument (Partec GmbH., Münster, Germany) (for PI staining) or (2) a UV mercury arc lamp embedded in a Partec PA II instrument (for DAPI staining). In the time-scale experiment, fluorescence intensity (measured in linear scale) and forward and side scatter (both in logarithmic scale) were recorded in laser-based measurements while only the first parameter was recorded in lamp-based measurements; in both cases, 5,000 particles were analysed. The following instrument settings were kept constant throughout the experiment: (1) the position of the first G_0/G_1 peak on channel 100 (using a 1,024-channel scale), (2) the discriminator for fluorescence (i.e. the lowest recorded

value) on channels 30 and 50 in the time-scale laboratory experiment and in the in situ experiment, respectively and (3) discriminators for forward and side scatter on channels 30 and 10, respectively. In the in situ experiment, only fluorescence intensity was recorded for both laser- and lamp-based instruments and the setting was adjusted independently for each sample to achieve optimal FCM results.

The following modifications were adopted for the analysis of glycerol-preserved samples (see Supplementary file S4 for a summarised procedure): approximately 300 mg of intact fresh leaf tissue of the sample was chopped together with the same amount of the internal reference standard in a Petri dish containing 6 mL of Otto I buffer. The suspension was filtered through a 42-µm nylon mesh, divided into twelve 0.5-mL aliquots and 0.5 mL of 60% glycerol solution was added to each aliquot. Six aliquots were kept in a freezer (−18°C) until FCM analysis ('ice-cold glycerol'), while the other six were left at room temperature ('RT glycerol'). Before FCM analysis, the suspension was centrifuged for 3 min at 3,200 rpm, the supernatant was discarded and 100 µL of ice-cold Otto I buffer was added to resuspend the nuclei. The sample tubes were gently shaken and the nuclear suspension was incubated for 15 min at room temperature. Finally, 1 mL of staining solution (Otto II buffer supplemented with β-mercaptoethanol and a fluorochrome) was added, and after 10 min of incubation, the samples were run on a flow cytometer.

B. perennis (2C03.38 pg; Schönswetter et al. 2007b) served as the internal reference standard for *E. milii*, *F. elastica* and *P. sativum*, while *P. sativum* (2C=8.84 pg; Greilhuber et al. 2007) was used as a reference point for *B. perennis*, *G. album* and *V. faba*. *Zea mays* from a local field in Madang was used as a reference standard for the tropical species included in the in situ experiment.

FCM histograms were evaluated using Partec Flomax 2.4d software. 'Fit Gauss Peaks' function was used to calculate basic descriptive statistics (mean position, CV and number of particles) of G_0/G_1 peaks. Because non-fresh material was measured, we adopted more relaxed quality standards, and as successful considered analyses with CVs of sample G_0/G_1 peaks up to 10%. The proportion of background noise was determined as a ratio between the number of particles outside and inside the area of G_0/G_1 peaks defined by the 'Fit Gauss Peaks' function.

In the time-scale experiment, all except 'silica gel' samples were analysed after 1, 7 and 15 days of storage; the 'silica gel' samples were only analysed after 15 days of storage. Each measurement (including sample preparation) was repeated on three subsequent days to minimise potential artefactual instrumental drift. As a result of this experimental design, each species was analysed 18 times per treatment (three times of storage, three replicates and two fluorochromes), except 'silica gel' for which each species was analysed only six times (three replicates and two fluorochromes). In the in situ experiment, samples were analysed once after 15 days of storage.

Statistical analyses

Parameters describing the quality of FCM histograms (CVs and proportions of background noise) were analysed using a mixed-effect analysis of variance (ANOVA) with species identity as a random factor. Separate ANOVAs were conducted for (1) all treatments (i.e. 'control', 'plastic bag', 'ice-cold glycerol', 'RT glycerol' and 'silica gel'), testing the effects of stain and preservation method, and (2) all except 'silica gel' treatments, testing the effect of storage time. Differences among individual treatments were further analysed by a Fisher's LSD test.

The stability of sample/standard fluorescence ratio across treatments was tested separately for each fluorochrome using a linear mixed-effect model with species identity as a random factor. The effects of the preservation method and storage time (including their interactions) were tested after the exclusion of the 'silica gel' treatment, while the effect of the preservation method alone was tested on the data from the 15th day of storage (i.e. with the 'silica gel' treatment included). The probability of success of FCM analysis for individual treatments was analysed using a binomial generalised linear mixed-effect model fit with the Laplace approximation (Bates and Maechler 2009).

ANOVAs with post hoc comparisons were calculated in Statistica version 8 (StatSoft, Inc. 2008), while R package version 2.9.2 (R Development Core Team 2009) was used to calculate mixed-effect and generalised mixed-effect linear models.

Results

Time-scale experiment

FCM acquisitions resulted in histograms with sample peak CVs ranging from 0.98% to 9.58% (mean 2.4%; Figs. 1 and 2); thus, all analyses were considered successful. The mixed-effect ANOVA on the entire dataset revealed significant effect of the preservation method on the quality of analysis ($F_{4, 20}$=67.62, p<0.001): the 'silica gel' samples generally exhibited the highest CVs (mean 4.2%) while the 'plastic bag' and 'ice-cold glycerol' samples (means for both treatments 2.0%) exhibited CV values comparable to those of fresh 'control' (mean 1.7%). PI-stained samples had generally slightly higher CVs than their DAPI-stained counterparts (means 2.67% and 2.19%, respectively; $F_{1, 5}$=29.28, p=0.003).

There was a significant interaction between the preservation method and the time of storage ($F_{6, 30}$=42.43, p<0.001). After 15 days of storage, 'plastic bag' and 'RT glycerol' samples generally showed higher CVs than freshly collected tissue. By contrast, the quality of 'ice-cold glycerol' samples was highly comparable to that of fresh control (Fig. 2). The effect of fluorochrome remained significant ($F_{1, 5}$=11.43, p=0.020); however, no significant interaction with the time of storage or preservation method was detected.

The proportion of background noise significantly differed among the preservation methods ($F_{4, 20}$=19.3, p<0.001); the highest background levels were generally observed in 'silica gel' samples. However, if the storage time was considered (i.e. with 'silica gel' samples excluded), no significant differences were detected either among the preservation methods ($F_{3, 15}$=2.85, p00.072) or their interactions with the time of storage ($F_{6, 30}$=0.97, p00.462).

With the exception of 'RT glycerol' samples, tissue preservation using any of the methods caused only a negligible shift in fluorescence intensity relative to the standard over the time span tested. Although the effect of both preservation method and its interaction with the time of storage was significant in both DAPI and PI datasets (Table 1), this was largely caused by 'RT glycerol' samples, which showed a significant decrease in sample/standard fluorescence ratio compared to the fresh 'control' (t_{203}=−2.65, p=0.008 and t_{203}=−2.14, p=0.033 for DAPI and PI data, respectively). The most stable were 'ice-cold glycerol' samples (absolute difference between preserved/fresh tissue below 2%), followed by 'plastic bag' (difference 0.2% to 3.1% and 0.3% to 5.7% for DAPI

and PI staining, respectively) and 'silica gel' (difference 0.2% to 2.6% and 1.6% to 12.9% for DAPI and PI staining, respectively) treatments. The difference in 'RT glycerol' samples usually exceeded 2% and reached up to ~18% in some cases (Table 2).

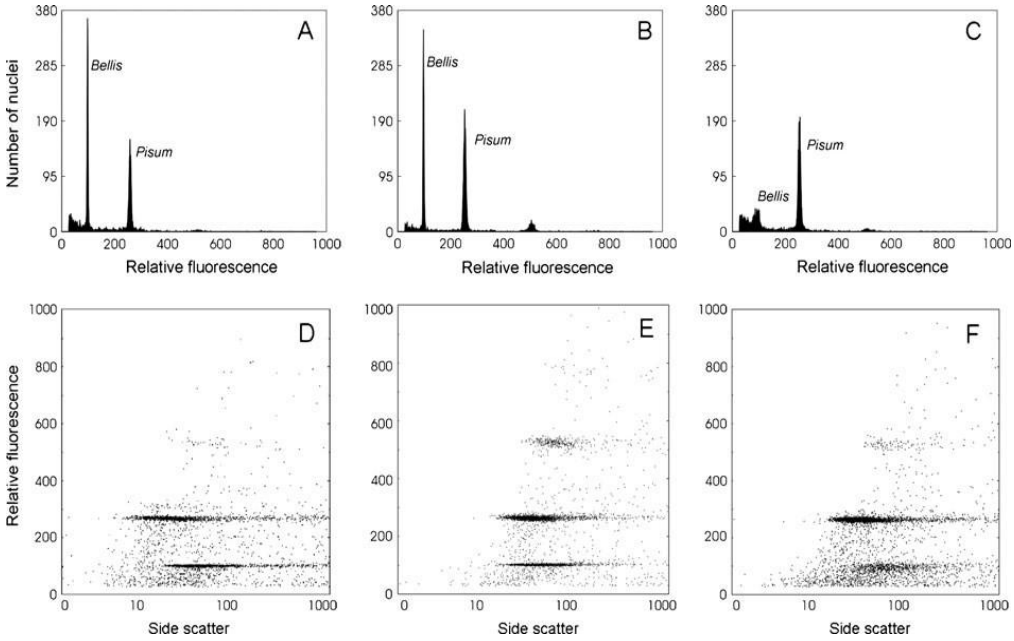


Fig. 1 Illustrative histograms of fluorescence intensities and side scatter/fluorescence scattergraphs of 'control' (A, D), 15-day-old samples of *B. perennis* stored in 'ice-cold glycerol' (B, E) and 'silica gel' (C, F), analysed together with the internal reference standard, *P. sativum* and stained with propidium iodide. CVs (%) of G_0/G_1 peaks of *Bellis/Pisum* are 1.92/1.68, 1.84/1.82 and 9.04/ 1.92 for the 'control', 'ice-cold glycerol' and 'silica gel' samples, respectively

In situ experiment

Eighty-one out of 126 samples transported from Papua New Guinea and stored ~15 days had been successfully analysed (i.e. yielded histograms with distinct peaks and CVs <10%, for illustrative histograms see Fig. S2). Acceptable histograms using at least one preservation technique were obtained in all but one species (Macaranga fallacina, Euphorbiaceae). The probability of success was significantly influenced by the preservation method (binomial generalised linear mixed-effect model $\chi^2=79.99$, $p<10^{-6}$; see also Fig. 3) and marginally also by the DNA fluorochrome ($\chi^2=4.5054$, $p=0.034$; DAPI slightly less successful). Regardless of the fluorochrome, the most successful was the 'ice-cold glycerol' treatment whereas the 'RT glycerol' one was the least successful.

In successful analyses, the type of preservation significantly influenced the sample CV ($F_{2, 24}$=19.51, p<0.001). The highest CVs were recorded in 'RT glycerol' samples (mean 8.20%) while CVs for 'ice-cold glycerol' (mean 4.51%) and 'plastic bag' (mean 5.02%) samples did not differ significantly from each other at α=0.05. Nuclei preserved in 'icecold glycerol' yielded the best FCM results (in terms of the lowest CVs) in more than two-thirds of the analysed species (14 and 13 species in DAPI and PI analyses, respectively), while the remaining species gave the lowest CVs when stored in a plastic bag in a refrigerator. 'Ice-cold glycerol' was the only mode of preservation that allowed successful FCM analysis in four species (Endospermum labios, Macaranga aleuritoides, Osmoxylon novo-guineense and Versteegia sp.; Table S1).

Discussion

The value of glycerol-preserved nuclei

Unlike animal and human biology, fresh samples still dominate plant FCM research, especially when absolute genome size values are required (Doležel et al. 2007b; Kron et al. 2007). Nonetheless, the ever-increasing number of applications in biosystematics, ecology and evolutionary biology has accelerated the search for methods of sample preservation applicable in field conditions and allowing longer-term sample storage. An ideal protocol should be simple and rapid (to be easily performed outside the laboratory), universal (applicable to a wide range of plant species) and reliable (introducing no artefactual shift in fluorescence intensity). The present study adds to previous methodological attempts and describes the advantages, limitations and potential use of glycerol-treated nuclear suspensions.

The suitability of glycerol solution for preserving nuclear suspension was first documented by Hopping (1993), who achieved promising results (i.e. distinct peaks and only a small shift in fluorescence intensity) using nuclei of A. deliciosa stored up to 9 months in the frost and stained by PI. We tested the value of his protocol on a set of plants from different families and covering a range of genome sizes (including popular plant reference standards), using two the most important DNA-selective fluorochromes with different binding modes (DAPI and PI). In addition, we applied the methodology in situ in the tropics to estimate the amount of nuclear DNA in 21 native species.

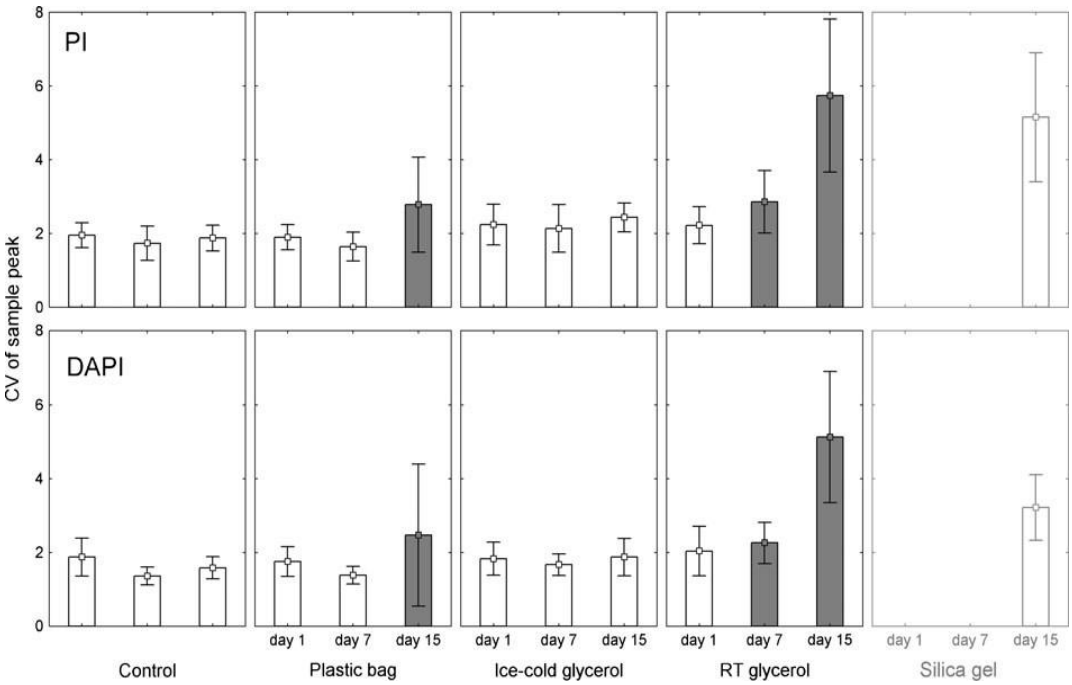


Fig. 2 The effects of preservation method and the time of storage on the quality of FCM analysis of six plant species expressed as the CV (mean±SD) of sample G_0/G_1 peak. Treatments significantly (at α=0.05) different from the control in a particular time point are highlighted in grey. Freshly collected leaves from the cultivated plants were used as control at each time point. Silica gel samples were only analysed after 15 days of storage and not tested for the temporal variation

Table 1 Summary of the linear mixed-effect model testing the effects of preservation method and storage time on sample/standard fluorescence ratio

Effect	DAPI staining			PI staining		
	df	F	p	df	F	p
Variation in fluorescence values across all storage times[a]						
Preservation method	3, 203	3.65	**0.014**	3, 203	5.80	**<0.001**
Time of storage	1, 203	2.98	0.086	1, 203	0.51	0.477
Preservation method×time of storage	3, 203	4.46	**0.004**	3, 203	4.65	**0.003**
Variation in fluorescence values after 15 days of storage						
Preservation method	4, 80	2.92	**0.026**	4, 80	2.33	0.063

Species identity was treated as a random factor. Significant effects are in bold

[a] Samples preserved in silica gel were not included in this test

Table 2 Variation in sample/standard fluorescence ratios of six plant species preserved using four different techniques and stained with two different fluorochromes after 15 days of storage

Species[a]	Preservation method	DAPI staining		PI staining	
		Sample/standard ratio (mean±SD)[b]	Difference from the control (%)	Sample/standard ratio (mean±SD)[b]	Difference from the control (%)
Bellis perennis[P]	Fresh control	0.419±0.008	–	0.381±0.007	–
(2C=3.38 pg)	Ice-cold glycerol	0.420±0.009	0.2	0.384±0.008	0.8
	RT glycerol	0.494±0.012	**17.9**	0.448±0.024	**17.6**
	Plastic bag	0.430±0.004	**2.6**	0.382±0.016	0.3
	Silica gel	0.427±0.003	1.9	0.374±0.012	−1.9
Euphorbia milii[B]	Fresh control	1.264±0.004	–	1.255±0.003	–
(2C=4.27 pg)	Ice-cold glycerol	1.258±0.015	−0.5	1.241±0.016	−1.1
	RT glycerol	1.200±0.029	**−5.3**	1.178±0.027	**−6.5**
	Plastic bag	1.274±0.011	0.8	1.262±0.013	0.6
	Silica gel	1.279±0.005	1.2	1.276±0.048	1.7
Ficus elastica[B]	Fresh control	0.533±0.004	–	0.452±0.004	–
(2C=1.52 pg)	Ice-cold glycerol	0.529±0.003	−0.8	0.444±0.004	−1.8
	RT glycerol	0.519±0.005	**−2.7**	0.442±0.003	**−2.3**
	Plastic bag	0.531±0.001	−0.4	0.450±0.003	−0.4
	Silica gel	0.530±0.003	−0.6	0.445±0.007	−1.6
Galium album[P]	Fresh control	0.447±0.003	–	0.412±0.003	–
(2C=3.62 pg)	Ice-cold glycerol	0.446±0.009	−0.2	0.419±0.002	1.7
	RT glycerol	0.497±0.024	**11.2**	0.467±0.020	**13.3**
	Plastic bag	0.449±0.002	0.4	0.426±0.020	**3.4**
	Silica gel	0.453±0.011	1.3	0.465±0.016	**12.9**
Pisum sativum[B]	Fresh control	2.386±0.046	–	2.625±0.047	–
(2C=8.84 pg)	Ice-cold glycerol	2.383±0.048	−0.1	2.605±0.053	−0.8
	RT glycerol	2.025±0.047	**−17.8**	2.239±0.119	**−17.2**
	Plastic bag	2.461±0.029	**3.1**	2.774±0.154	**5.7**
	Silica gel	2.449±0.025	**2.6**	2.584±0.230	−1.6
Vicia faba[P]	Fresh control	3.092±0.012	–	3.047±0.017	–
(2C=26.9 pg)	Ice-cold glycerol	3.122±0.023	1	3.011±0.039	−1.2
	RT glycerol	3.117±0.050	0.8	3.054±0.067	0.2
	Plastic bag	3.097±0.040	0.2	3.106±0.034	1.9
	Silica gel	3.087±0.037	−0.2	3.139±0.092	**3**

Differences from the fresh control that exceeded the 2% threshold are in bold

[a] Internal reference standards: superscript B – *Bellis perennis*, superscript – *Pisum sativum*

[b] Three independent replicates on three successive days

The preservation of nuclear suspensions in a 30% glycerol–Otto I buffer solution (Otto 1990; Doležel et al. 2007b) and sample storage at −18°C was found to be a reliable method for the FCM estimation of genome size in plants in both absolute and relative units. The quality of analysis (expressed as CVs of G_0/G_1 sample peaks and the proportion of background noise) as well as their reliability (i.e. the stability of fluorescence intensity) appeared to be unaffected by the glycerol treatment, at least during the time period studied (15 days). After 2 weeks of storage in frost, both the quality measures of resulting FCM histograms and genome size estimates were fully comparable to those of fresh control samples (Fig. 2; Table 2).

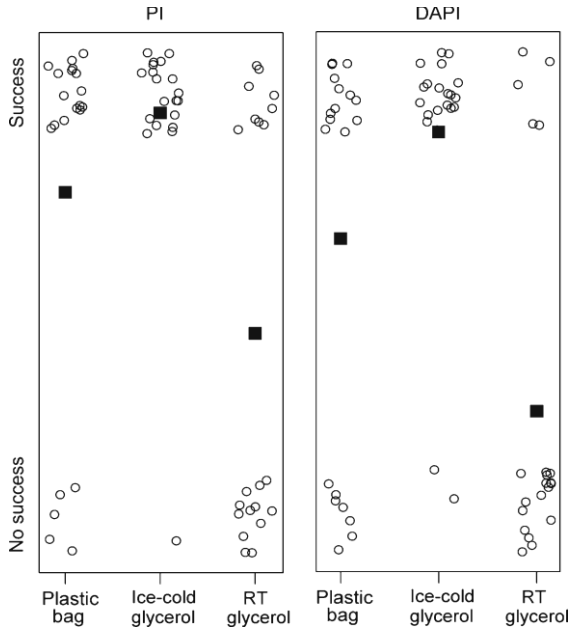


Fig. 3 The probability of achieving successful FCM analysis (i.e. sample CV below 10%) using three different preservation methods in the multi-species in situ experiment. Open circles samples, black squares mean probability success for a particular treatment. To visualise all samples, a small error variance was added to each value representing individual measurements

Furthermore, light scattering properties of the nuclei showed no signs of the so-called tannic acid effect, indicating that the analyses were not negatively affected by interfering secondary metabolites and/or nuclei aggregation (Loureiro et al. 2006; see Fig. 1). In general, the analyses of glycerol-preserved nuclei mostly fulfilled the stringent criteria applied for the genome size estimation in absolute units in freshly collected samples, including the requirements of stable sample/ standard fluorescence ratio and CVs below 5% (e.g. Doležel et al. 2007a). A crucial step of the proposed protocol seems to be the storage of glycerol-preserved nuclei at −18°C, as samples kept at room temperature deteriorated quickly (Fig. 2).

Comparison with other modes of sample preservation

Despite the ongoing debate about the use of non-fresh (preserved) plant material for FCM analysis (Doležel and Bartoš 2005; Doležel et al. 2007a; Bainard et al. 2011), only a handful of studies have ventured to explore the value of fixed plant nuclei, cells and/or tissues to estimate genome size. Rapid tissue desiccation (most conveniently done using silica gel) is the only way of sample preservation that has received wider attention in ecological and evolutionary plant FCM studies. However, the dehydrated samples were mostly stained with AT-selective fluorochrome DAPI, which has favourable staining properties (e.g. comparatively low sensitivity to chromatin condensation and high increase in quantum efficiency after binding to the DNA molecule; Shapiro 2003), but precludes genome size estimation in absolute units. The artefactual shift in fluorescence intensity often observed after tissue dehydration (e.g. Šmarda 2006; Suda and Trávníček 2006b; Cires et al. 2009) and lower quality of FCM analysis also work against its use for absolute genome size estimation. Only recently, the potential of dried plant material for genome size research was thoroughly evaluated (Bainard et al. 2011). The authors considered the fluorescence shift introduced by drying (<10%) to be acceptable, as it fell within the limits introduced by other methodological factors (e.g. seasonal variation, instrument and buffer used, among others), and concluded that PI-stained samples can represent a promising option.

Despite its higher laboriousness in comparison with tissue dehydration, glycerol-preserved nuclei apparently represent the most reliable way of sample preservation for genome size research, at least in a short-term time frame. In the present study, estimates of nuclear DNA in glycerol-treated samples after 2 weeks of storage in frost were highly comparable to those obtained using fresh material. Whereas the ‘silica gel’ samples experienced up to a 12.9% shift in fluorescence intensity and CVs of 4.2%, on average, the ‘ice-cold glycerol’ samples showed very stable fluorescence (maximum difference <1.7%) and much lower CVs (Table 2). Interestingly, 2-week-old glycerol-preserved nuclei yielded better results than living plant tissues kept for the same time in a humid environment in the cold (Fig. 2), which is usually the first choice for short-term sample storage (Suda et al. 2007a). It should be noted that the period for which fresh tissues can be stored in a refrigerator before FCM analysis is considerably influenced by leaf characteristics. Whereas species with small and soft leaves (in our set, for instance, B. perennis and P. sativum) deteriorate quickly, plants with tough and leathery leaves (e.g. E. milii and F. elastica) seem to be generally less sensitive (Fig. S3). In parallel, glycerol preservation represented the only way to analyse several soft-leaved tropical plants in the in situ experiment (Table S1). Although the fluorescence properties of glycerol-preserved nuclei after long-term storage are largely unknown and in need of further study (but note that Hopping (1993) suggests that nuclei are likely to remain intact for at least several months), the presented methodology appears to be a very promising way of sample preservation for genome size studies.

Applicability of the protocol

Of particular importance is the fact that plant nuclei stored in ice-cold glycerol remain intact for at least a few weeks and provide estimates of nuclear DNA content that are highly comparable and of similar quality to those obtained from fresh tissue. Furthermore, the protocol is compatible with both DAPI and PI staining, allowing the determination of not only ploidy level but, more importantly, genome size in absolute units. Moreover, high resolution of resulting FCM histograms opens the possibility of detecting small differences in nuclear DNA content.

Table 3 Advantages, limitations and potantial applications of different approaches used to substitute fresh somatic tissues in plant FCM

Type of material	Advantage(s)	Limitation(s)	Applications(s)
Cold-stored somatic tissue	Reliable FCM measurements (fluorescence stability)	Short-time preservation (need for immediated FCM analysis)	Screening of ploidy variation across multiple species and at various spatial scales
	Suitability for genome size estimation in absolute units	Challenging transport (need for a sufficient space, phytosanitary certificate)	Genome size estimation
	Easy sample preparation		Detection of small differences in the amount of nuclear DNA
Dry seeds	Reliable FCM measurements (when cultivated plants are used)	Need for a proper timing of collection	Screening of ploidy variation across multiple species and at various spatial scales
	Suitability for the entire range of FCM applications (when cultivated plants are used)	Possible germination problems	Genome size estimation
	Easy sample transport	Possible shift in fluorescence intensity whendry seeds are analysed directly	Detection of small differences in the amount of nuclear DNA
	Convenient and long-term storage	Need for a cultivation facility	
		Unknown male parent of the individual analysed (e.g. putative hybrid origin of the seed)	
Dehydrated somatic tissue (silica- or air-dried)	Easy sample transport	Limited use for absolute genome size estimation (possible shift in fluorescence intensity)	Screening of ploidy variation across multiple species and at various spatial scales
	Convenient sample preparation	Lower resolution than in fresh tissue (challenging detection of small differences in the amount of nuclear DNA)	First insights into genome size variation

Frozen somatic tissue	Convenient sample preparation	Dependence on a basic lab facility for sample preservation (freezer)	Screening of ploidy variation in model species
		Challenging sample transport	First insights into genome size variation
		Limited use for absolute genome size estimation (possible shift in fluorescence intensity)	
		Lower resolution than in fresh tissue	
Chemically fixed somatic tissue (ethanol- or formaldehyde-based fixations)		Dependence on a basic lab facility for sample preservation	Screening of ploidy variation in model species
		Dependence on the availability of internal reference standard(s)	
		Challenging selection of a suitable reference standard for unknown sample species	
		Likely shifts in fluorescence intensity (changes in chromatin compaction)	
Glycerol-preserved nuclei	Reliable FCM measurements (fluorescence stability)	Dependence on a basic lab facility for sample preservation (freezer)	Screening of ploidy variation in model species
	Suitability for genome size estimation in absolute units	Challenging sample transport	Genome size estimation
		Dependence on the availability of internal reference standard(s)	Detection of small differences in the amount of nuclear DNA
		Challenging selection of a suitable reference standard for unknown sample species	

Finally, glycerol-preserved nuclei also offer opportunity to cytotype plant species with soft and rapidly decaying leaves, in which other modes of storage usually provide less satisfactory results or completely fail.

The major limitations of the proposed protocol stem from higher demands on sample preparation, i.e. the need for basic laboratory facilities, including a freezer, and the necessity for

suitable reference standard(s), a selection of which should meet several criteria (see Suda and Leitch 2010). In addition, the reliability of the protocol after longer-term storage (several months) needs to be assessed.

A detailed comparison of current methodologies of sample preservation for plant FCM, including their pros and cons, is provided in Table 3. Considering these facts, we assume that glycerol-treated nuclei will provide a vital alternative to other preservation methods, especially when stringent criteria on FCM analysis are required (e.g. in genome size studies) and/or if a detailed investigation of a single or a few plant species is intended. The storage of fresh tissue in cold will likely be the method of choice for short-term field trips while large-scale ploidy studies in remote areas will probably still be dominated by silica-dried samples.

Acknowledgements

Jana Krejčíková and Pavel Trávníček helped us with the flow cytometric analyses and Jan Lepš, Kateřina Štajerová and Kenneth Molem kindly assisted with the collection and determination of the tropical species. The project was supported by projects 206/08/H049 and P506/10/ 0704 from the Czech Science Foundation. JT was supported by project 138/2010/P. Additional support was provided by the Academy of Science of the Czech Republic within the Institutional Research Programme AV0Z60050516 and the Ministry of Education, Youth and Sports of the Czech Republic (projects MSM0021620828 and MSM6007665801).

References

Bainard J. D., Husband B. C., Baldwin S. J., Fazekas A. J., Gregory T. R., Newmaster S. G. & Kron P. (2011): The effects of rapid desiccation on estimates of plant genome size. – Chromosome Res 19: 825–842.

Bates D. & Maechler M. (2009): Package 'lme4'– linear mixed effects models using S4 classes. – R Foundation for Statistical Computing, Vienna.

Bendiksby M., Tribsch A., Borgen L., Trávníček P. & Brysting A. K. (2011): Allopolyploid origins of the *Galeopsis* tetraploids — revisiting Müntzing's classical textbook example using molecular tools. – New Phytol 191: 1150–1167.

Chiatante D., Brusa P., Levi M., Sgorbati S. & Sparvoli E. (1990): A simple protocol to purify fresh nuclei from milligram amounts of meristematic pea root tissue for biochemical and flow cytometry applications. – Physiol Plantarum 78: 501–506.

Cires E., Cuesta C., Peredo E. L., Revilla M. A. & Prieto J. A. F. (2009): Genome size variation and morphological differentiation within *Ranunculus parnassifolius* group (Ranunculaceae) from calcareous screes in the northwest of Spain. – Plant Syst Evol 281: 193–208.

Dart S., Kron P. & Mable B. K. (2004): Characterizing polyploidy in *Arabidopsis lyrata* using chromosome counts and flow cytometry. Can J Bot 82: 185–197.

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

Doležel J., Greilhuber J. & Suda J. (2007a): Flow cytometry with plant cells: analysis of genes, chromosomes and genomes. – Wiley-VCH, Weinheim.

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

Dušková E., Kolář F., Sklenář P., Rauchová J., Kubešová M., Fér T., Suda J. & Marhold K. (2010): Genome size correlates with growth form, habitat and phylogeny in the Andean genus *Lasiocephalus* (Asteraceae). – Preslia 82: 127–148.

Eidesen P. B., Alsos I. G., Popp M., Stensrud Ø., Suda J. & Brochmann C. (2007): Nuclear vs. plastid data: complex Pleistocene history of a circumpolar key species. – Mol Ecol 16: 3902–3925.

Greilhuber J., Temsch E. M. & Loureiro J. (2007): Nuclear DNA content measurement. In: Doležel J., Greilhuber J., Suda J. (eds) Flow cytometry with plant cells: analysis of genes, chromosomes and genomes. – Wiley-VCH Verlag, Weinheim, pp 67–101.

Halverson K., Heard S. B., Nason J. D. & Stireman J. O. (2008): Origins, distribution, and local co-occurrence of polyploid cytotypes in *Solidago altissima* (Asteraceae). – Am J Bot 95: 50–58.

Hopping M. (1993): Preparation and preservation of nuclei from plant tissues for quantitative DNA analysis by flow cytometry. – N Z J Bot 31: 391–401.

Hülber K., Sonnleitner M., Flatscher R., Berger A., Dobrovsky R., Niessner S., Nigl T., Schneeweiss G. M., Kubešová M., Rauchová J., Suda J. & Schönswetter P. (2009): Ecological segregation drives fine-scale cytotype distribution of *Senecio carniolicus* in the Eastern Alps. – Preslia 81: 309–319.

Košnar J. & Kolář F. (2009): A taxonomic study of selected European taxa of the *Tortula muralis* (Pottiaceae, Musci) complex: variation in morphology and ploidy level. – Preslia 81: 399–421.

Kron P., Suda J. & Husband B. C. (2007): Applications of flow cytometry to evolutionary and population biology. – Rev Ecol Evol Syst 38: 847–876.

Loureiro J., Rodriguez E., Doležel J. & Santos C. (2006): Flow cytometric and microscopic analysis of the effect of tannic acid on plant nuclei and estimation of DNA content. – Ann Bot 98: 515–527.

Loureiro J., Trávníček P., Rauchová J., Urfus T., Vít P., Štech M., Castro S. & Suda J. (2010): The use of flow cytometry in the biosystematics, ecology and population biology of homoploid plants. – Preslia 82: 3–21.

Nsabimana A. & Van Staden J. (2006): Ploidy investigation of bananas (*Musa* spp.) from the National Banana Germplasm Collection at Rubona-Rwanda by flow cytometry. – S Afr J Bot 72: 302–305.

Otto F. (1990): DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. In: Crissman H. A., Darzynkiewicz Z. (eds) Methods in cell biology 33. – Academic, San Diego, pp 105–110.

Popp M., Gizaw A., Nemomissa S., Suda J. & Brochmann C. (2008): Colonization and diversification in the African 'sky islands' by Eurasian *Lychnis* L. (Caryophyllaceae). – J Biogeogr 35: 1016–1029.

R Development Core Team (2009): R: A language and environment for statistical computing. – R Foundation for Statistical Computing, Vienna.

Schönswetter P., Lachmayer M., Lettner C., Prehsler D., Rechnitzer S., Reich D. S., Sonnleitner M., Wagner I., Hülber I., Schneeweiss G. M., Trávníček P. & Suda J. (2007a): Sympatric diploid and hexaploid cytotypes of *Senecio carniolicus* (Asteraceae) in the Eastern Alps are separated along an altitudinal gradient. – J Plant Res 120: 721–725.

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

Shapiro H. (2003): Practical flow cytometry, 4th edn. – Wiley-Liss, New York

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

Šmarda P. (2006): DNA ploidy levels and intraspecific DNA content variability in Romanian fescues (*Festuca*, Poaceae) measured in fresh and herbarium material. – Folia Geobot 41: 417–432.

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.

StatSoft, Inc. (2008): STATISTICA (data analysis software system), version 8.0.

Suda J. & Leitch I. J. (2010): The quest for suitable reference standards in genome size research. – Cytom Part A 77A: 717–720.

Suda J. & Trávníček P. (2006a): Estimation of relative nuclear DNA content in dehydrated plant tissues by flow cytometry. – Curr Protoc Cytom 38:7.30.1–7.30.14.

Suda J. & Trávníček P. (2006b): Reliable DNA ploidy determination in dehydrated tissues of vascular plants by DAPI flow cytometry – new prospects for plant research. – Cytom Part A 69A: 273–280.

Suda J., Kyncl T. & Jarolímová V. (2005): Nuclear DNA amounts in Macaronesian angiosperms: forty percent of Canarian endemic flora completed. – Plant Syst Evol 252: 215–238.

Suda J., Kron P., Husband B. C. & Trávníček P. (2007a): Flow cytometry and ploidy: applications in plant systematics, ecology and evolutionary biology. In: Doležel J., Greilhuber J., Suda J. (eds) Flow cytometry with plant cells: analysis of genes, chromosomes and genomes. – Wiley-VCH, Weinheim, pp 103–130.

Suda J., Weiss-Schneeweiss H., Tribsch A., Schneeweiss G., Trávníček P. & Schönswetter P. (2007b): Complex distribution patterns of di-, tetra-, and hexaploid cytotypes in the European high mountain plant *Senecio carniolicus* (Asteraceae). – Am J Bot 94: 1391–1401.

Volkova P. A., Trávníček P. & Brochmann C. (2010): Evolutionary dynamics across discontinuous freshwater systems: rapid expansions and repeated allopolyploid origins in the Palearctic white water-lilies (*Nymphaea*). – Taxon 59: 483–494.

KEY RESULTS AND CONCLUSION

GS dynamics and variation in study systems

In **Case study I** we confirmed the existence of extraordinary intraspecific genome size variation in *Taraxacum stenocephalum* (Asteraceae), which at the homoploid level ranged up to 1.262-fold in 775 F1 seedlings. Such big differences in 2C DNA content between individuals are not associated with aneuploidy or the presence of B chromosomes. The only detected karyological differences were found in the presence of satellite chromosomes, so we assume that repetitive DNA contributes to DNA content variation in the study species. DNA content values in F1 individuals were directly proportionally determined by parental genome size and no shift in the GS of F1 offspring was detected as compared to parental GS.

Case studies II and **III** demonstrate the importance of genome size for large-scale cytological screenings in ploidy-variable groups. We studied differences between di- and tetraploid cytotypes in two different species – *Galium valdepilosum* (Rubiaceae) and *Arabidopsis arenosa* (Brassicaceae). In *G. valdepilosum* we detected only one contact zone and observed no mixed-ploidy populations. Although extensive sampling (874 individuals), we found no triploids, which indicates that the contact zone is stable. The within-cytotype variation is very low. In *Arabidopsis arenosa* we sampled more than 2,900 individuals and recorded the parapatric distribution of diploids and tetraploids with two contact zones. At the landscape level, diploid and tetraploid populations are interspersed. At the population level, however, they are cytologically homogeneous. We found only a few triploid individuals, always in purely diploid populations. We therefore assume that the triploids were most likely formed by the fusion of reduced and unreduced gametes. Therefore,in *A. arenosa* we also regard the contact zone as stable. In both **Case studies (II, III),** genome size estimation brought some new unexpected findings. In *G. valdepilosum*, genome size indicates the relatedness of West-Bohemian populations to *G. valdepilosum* and not to *G. sudeticum*, as had been previously thought. For *A. arenosa* we found a new isolated diploid lineage on the Baltic coast. We were the first to publish the occurrence of triploids in *A. arenosa.* The scarcity of triploids among the nearly 3,000 sampled individuals shows how rarely are triploids generated. Mixed-ploidy populations were found in *A. arenosa* only with an even bigger sampling effort (F. Kolář et al. – as yet not published). This demonstrates the relevance of detailed studies with extensive sampling, which can reveal the influence of certain very rare evolutionary events which, however, may have important consequences for species evolution.

In **Case study IV** we studied 184 individuals of 20 species of *Lasiocephalus* (Asteraceae). The estimated range of DNA content (1.64-fold variation) indicates ploidy variation. Based on the 1.5-fold difference in genome size accompanied by high morphological similarity, we assume that triploid individuals occur in some of the species. Because this supposition was not verified karyologically, we refer to presumed triploids as DNA triploids. The smallest values of DNA content in *Lasiocephalus* revealed hybridization with the closely related species *Culcitium nivale*.

Association of genome size with biological traits – from phenotype to ecology

In **Case study I** we found a significant association between intraspecific genome size and weight of fruit, time to germination, time to flowering, length of leaves and aboveground biomass. This could be explained by the size of nuclei and the duration of cell division, which together influence the growth of individuals. In congruence with other studies, we also found a non-linear relationship between genome size and particular traits. In *Taraxacum stenocephalum* there is a unimodal distribution of leaf length, which reached its greatest values in individuals with mid-sized genomes.

Case study II illustrates the different ecological preferences of diploid and tetraploid *Galium valdepilosum*. The tetraploid cytotype was found over a larger distribution area and in more habitat types than diploid cytotypes. Although serpentine substrates represent an extreme and stressful environment, both cytotypes are able to grow on this substrate to a similar extent.

By contrast, in **Case study III** on *Arabidopsis arenosa* we did not detect any differences in ecological preferences; both cytotypes occur in habitats of various types. Differences at the microecological scale in contact zones will be the topic of a future study.

In **Case study IV** we found correlations of genome size with habitat characteristics, growth forms and elevation in diploid species. However, all these traits were correlated with phylogeny. At the generic level, phylogeny seems to be the determining factor of genome size variation. At the lower taxonomic levels, ecological conditions are important and genome size can be altered through selection.

Case study V represents the first more extensive quantitative study of the association between genome size and invasiveness carried out on a set of congeners. The study is based on 93 alien species from 32 families, naturalized in the Czech Republic. We confirmed that invasiveness is associated with small genomes. We examined alien species with various status of invasiveness. Naturalized species have smaller genomes than their non-invasive congeners. When we compared naturalized aliens with invasive aliens, we did not find any difference in genome size. This means that DNA content is important in the process of naturalization and it is not crucial exclusively for invasiveness. Based on our results we suppose that species with smaller genomes are potentially more successful and they could become naturalized or invasive over time. Another paper confirming our findings about the association between genome size and invasiveness at the interspecific level followed (Pyšek et al. 2018 – *Phragmites australis*).

In **Case study V** we extracted information about congeners and confamiliers from the Angiosperm DNA C-value database (Bennett & Leitch 2012), which provides useful data allowing to make various comparisons across the plant system and helping to reveal general trends in GS evolution. Although the estimation of genome size requires great care (numerous of old estimates have been refuted as incorrect because of methodological problems), the database is useful for revealing ranges of genome size variation and associations across the angiosperms. As stated in Knight et al. (2005), joining data on genome size with databases of functional traits in combination with GIS data promises new findings about direct effects of genome size on the lives of plants. However, phylogenetic corrections should be applied to intergeneric or interspecific comparisons of genome size because phylogeny can be the main

source of genome size differences at this level, as shown in **Case study IV**. At the generic level, the phylogeny of *Lasiocephalus* is significantly correlated with genome size and with habitat requirements, elevation and growth form. In **Case study V** we used congener pairs, so there was no need to make any phylogenetic corrections.

Overcoming limits of FCM genome size estimation

Observance of a best-practice protocol is essential in genome size studies. Estimation of genome size entails a number of necessary procedures to ensure data reliability (see Doležel & Bartoš 2005, Walker et al. 2006, Doležel et al. 2007, Greilhuber 2008, Šmarda & Bureš 2010). Despite adherence to strict rules, certain issues may complicate the estimation of genome size. One limiting requirement is the need for fresh material. This hinders the determination of genome size in plants which are collected during long field trips or in remote areas if the transport of samples to a cytometer takes a long time. In **Case studies II**, **III** and **IV** we successfully used a combination of silica-dried samples and fresh samples. This enabled us to perform extensive screening throughout the distribution area of the study species (**Case studies II, III**) and to bring samples from remote regions (**Case study IV**).

The use of silica-dried samples brought significant, progress, but it is applicable only for certain types of analyses. With most species it can be used only for flow cytometry using AT-selective DAPI staining. This makes it impossible to estimate absolute genome size, and only relative genome size can be ascertained by this method. Fortunately, in all groups under study there was a strong correlation between the results of DAPI and PI analyses, the results were comparable for the purpose of ploidy level estimation. In addition, the quality of silica-dried samples is often lower compared to fresh samples and the possible shift in the fluorescence of nuclei makes it impossible to combine the two types of data. Generally, silica-dried samples are sufficient for ploidy determination (**Case study II, III, IV**) and for extensive rapid ploidy screening because such analyses are run until the acquisition of only 3,000 particles (in contrast to analyses for the purpose of estimating absolute genome size, which are run until the acquisition of 5,000 particles and repeated on three different days). Moreover, if the material gives peaks of high quality, it is also possible to detect double peaks, which constitute the best proof of tiny differences in genome size. The big advantage is that samples can be analysed after several weeks or months after being collected and dried. For example, in **Case study IV** we successfully analysed silica-dried samples up to 18 months old. This makes it easier to process samples from remote areas or when it is helpful to postpone the analysis of samples until after the busy vegetation season.

In **Case study VI** we tried to overcome the need for fresh samples by designing a protocol for sample preservation which would be applicable also for absolute genome size estimation. We modified the protocol of Hopping (1993) and created a new protocol for the fixation of nuclei. This we tested for reliability under laboratory conditions in a time-scale experiment using six species covering a wide range of plant genome sizes. We tried to apply the modified original protocol using ice-cold glycerol (Hopping 1993), which requires electricity for the operation of a freezer. As this is unavailable in in remote regions, so we also

tested the protocol with room-temperature (RT) glycerol. We compared the quality of GS analyses of fresh samples (controls), tissue stored in plastic bags in a refrigerator, nuclei preserved in RT glycerol, nuclei in ice-cold glycerol and silica-dried samples. After one week and after fifteen days, we analysed the samples and compared the quality of different treatments. Besides control samples, the best results were reached with nuclei preserved in ice-cold glycerol, followed by the plastic bags treatment and silica-dried samples; the RT glycerol treatment produced the worst results. Nuclei suspensions fixed in ice-cold glycerol were comparable with control samples for the purposes of estimating both absolute and relative genome size. We successfully applied this protocol during research trips to Papua-New Guinea on a set of 21 species from 12 families. The samples remained usable for several weeks even though they were not kept cold during the 24 hours it took to travel to the laboratory. This or similar protocols represent a promising way to expand our knowledge about genome size in undeveloped and unexplored remote areas. It could help expand the volume of data in the Plant C-value database (Bennett & Leitch 2012), most of which are from the Northern Hemisphere. Additional data from the Southern Hemisphere and other unexplored regions could bring unexpected findings or at least consolidate our ideas about the range and distribution of genome size variation in plants.

Conclusion and future prospects

This thesis demonstrates that genome size can be used as an important marker of direct dependence of a certain trait on genome size, mainly at the cellular level, which is manifested in the form of phenotypic correlations, as well as of ecological differentiation of taxa. Genome size can be a useful marker for delimiting species and studying important evolutionary processes such as hybridization or polyploidization. Genome size can also help to explain the distribution of species or understand their phylogeny and evolutionary history. All the included case studies brought new genome size values, expanding our knowledge about plant genome size, and someinteresting associations between genome size and biological traits of plants.

Intraspecific genome size variation at the homoploid level provides a suitable subject for detailed genome size studies. The advantage of this approach is that GS variation within one species is not determined by phylogeny and that it compares genomes that have been shaped by the same evolutionary processes and at the same time. However, differences in DNA content of larger magnitude are rare in wild species at the homoploid level and there is lack of studies working with homoploid species with substantial genome size variation over several generations. As a future prospect we should look at GS variation in successive generations to see under which conditions is genome size adaptive. This could be helpful for revealing the actual effect of genome size compared to another factor during the growth of individuals.

Whole-genome duplication is a crucial process in plant speciation. Genome size estimation in angiosperms often reveals polyploidy. It can also indicate the age of polyploids or point to potential progenitors of newly formed allopolyploids. Mixed-ploidy species

constitute interesting study material for exploring the effects effect of varying nuclear DNA content. However, the whole issue of polyploidy is so complex that it is difficult or even impossible to separate the effect of genome duplication from changes in the functioning of genes, which are also connected with polyploidy. For that purpose, newly formed autopolyploids could provide suitable material.

Many previous studies describing associations of genome size with various biological traits are based on interspecies comparisons without the application of phylogenetic corrections. I suppose that a re-evaluation of these data using phylogenetic corrections might bring quite a few surprises because the role of genome size is not equal in different groups.

In general, there are some direct associations between genome size with certain cellular characteristics (nuclei volume, duration of cell division) based on physical limits, but further effects and consequences of genome size are likely to differ from species to species, as they also depend on other factors and processes (such as selection) influencing traits associated with genome size.

Lastly, the causes and consequences of genome size variation can be multifactorial, so it is necessary to take a multi-method approach to unravelling the true nature of genome size variation. Sometimes we view genome size as a primary source of variability in plants when in fact it can mirror past or ongoing evolutionary processes. The authors of some papers argue that previous works have been done only on a limited number of species or within a single genus or that different species with different evolutionary histories have been compared. In my opinion, this is the exact question we ought to ask. However, we should formulate it precisely and, accordingly, choose a suitable model system that will allow us to reach an answer. In any case, genome size significantly influences the properties and behaviour of plants and is a highly useful marker in plant studies.

REFERENCES

Bainard J. D., Husband B. C., Baldwin S. J., Fazekas A. J., Gregory T. R., Newmaster S. G. & Kron P. (2011): The effects of rapid desiccation on estimates of plant genome size. – Chromosom. Res. 19: 825-842.

Barker M. S., Arrigo N., Baniaga A. E., Li Z. & Levin D. A. (2016): On the relative abundance of autopolyploids and allopolyploids. – New Phytol. 210: 391–398.

Beaulieu J. M., Leitch I. J. & Knight C. A. (2007): Genome size evolution in relation to leaf strategy and metabolic rates revisited. – Ann. Bot. 99: 495–505.

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

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

Beest M. te, Roux J. J. Le, Richardson D. M., Brysting A. K., Suda J., Kubesova M. & Pysek P. (2012): The more the better? The role of polyploidy in facilitating plant invasions. – Ann. Bot. 109: 19-45.

Bennett M. D. (1972): Nuclear DNA content and minimum generation time in herbaceous plants. – Proc. R. Soc. London Ser. B. 181: 109–135.

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

Bennett MD & Leitch IJ. 2012: Angiosperm DNA C-values database (release 8.0, Dec. 2012) http://www.kew.org/cvalues/

Bennett MD, Leitch IJ, Hanson L. 1998. DNA amounts in two samples of angiosperm weeds. Annals of Botany 82: 121–134.

Bennetzen J. L. (2000): Transposable element contributions to plant genome evolution. – Plant Mol Biol. 42: 251-269.

Bennetzen J. L. (2005): Transposable elements, gene creation and genome rearrangement in flowering plants. – Curr. Opin. Genet. Dev. 15: 621–627.

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

Biémont C. (2008): Genome size evolution: within-species variation in genome size. – Heredity (Edinb). 101: 297–298.

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

Chiatante D., Brusa P., Levi M., Sgorbati S. & Sparvoli E. (1990): A simple protocol to purify fresh nuclei from milligram amounts of meristematic pea root tissue for biochemical and flow cytometry applications. – Physiol. Plant. 78: 501–506.

Christenhusz M. J. M. & Byng J. W. (2016): The number of known plants species in the world and its annual increase. – Phytotaxa 261: 201–217.

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

Chumová Z., Krejčíková J., Mandáková T., Suda J. & Trávníček P. (2015): Evolutionary and taxonomic implications of variation in nuclear genome size: lesson from the grass genus Anthoxanthum (Poaceae). – PLoS One 10: 1–17.

Cires E., Cuesta C., Peredo E. L., Revilla M. A. & Prieto J. A. F. (2009): Genome size variation and morphological differentiation within Ranunculus parnassifolius group (Ranunculaceae) from calcareous screes in the northwest of Spain. – Plant Syst Evol 281:193–208.

Clarkson J. J., Lim K. Y., Kovarik A., Chase M. W., Knapp S. & Leitch A. R. (2005): Long-term genome diploidization in allopolyploid Nicotiana section Repandae (Solanaceae). – New Phytol. 168: 241–252.

Comai L. (2005): The advantages and disadvantages of being polyploid. – Nat. Rev. Genet.

Dart S., Kron P. & Mable B. K. (2004): Characterizing polyploidy in Arabidopsis lyrata using chromosome counts and flow cytometry. – Can J Bot 82:185–197.

Devos K. M., Brown J. K. M. & Bennetzen J. L. (2002): Genome size reduction through illegitimate recombination counteracts genome expansion in Arabidopsis. – Genome Res. 12: 1075–1079.

Díez C. M., Gaut B. S., Meca E., Scheinvar E., Montes-Hernandez S., Eguiarte L. E. & Tenailllon M. I. (2013): Genome size variation in wild and cultivated maize along altitudinal gradients. – New Phytol. 199: 264–276.

Dodsworth S., Leitch A. R. & Leitch I. J. (2015): Genome size diversity in angiosperms and its influence on gene space. – Curr. Opin. Genet. Dev. 35: 73-78.

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

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. (1991): Flow cytometric analysis of nuclear DNA content in higher plants. – Phytochem. Anal. 2: 143-154.

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

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

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. – Ann. Bot. 82: 17-26.

Flagel L. E. & Wendel J. E. (2009): Gene duplication and evolutionaary novelty in plants. – New Phytol. 183: 557-564.

Fleischmann A., Michael T. P., Rivadavia F., Sousa A., Wang W., Temsch E. M., Greilhuber J., Müller K. F. & Heubl G. (2014): Evolution of genome size and chromosome number in the

carnivorous plant genus Genlisea (Lentibulariaceae), with a new estimate of the minimum genome size in angiosperms. – Ann. Bot. 114: 1651–1663.

Francis D., Davies M. S. & Barlow P. W. (2008): A strong nucleotypic effect on the cell cycle regardless of ploidy level. – Ann. Bot. 101: 747–757.

Gregory T. R. (2001): Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. – Biol. Rev. 76: 65-101.

Gregory T. R. (2005): Genome size evolution in animals. In T. R. Gregory. The Evolution of the Genome. San Diego: Elsevier. pp. 3–87.

Greilhuber J. & Temsch E. M. (2001): Feulgen densitometry: some observations relevant to best practice in quantitative nuclear DNA content determination. – Acta Bot. Croat. 60: 285–298.

Greilhuber J. (2005): Intraspecific variation in genome size in angiosperms: Identifying its existence. – Ann. Bot. 95: 91–98.

Greilhuber J. (2008): Cytochemistry and C-values: The less-well-known world of nuclear DNA amounts. – Ann. Bot. 101: 791–804.

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. – Ann. Bot. 95: 255–260.

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.

Grotkopp E, Rejmánek M, Rost TL. 2002. Toward a causal explanation of plant invasiveness: seedling growth and life-history strategies of 29 pine (Pinus) species. The American Naturalist 159: 396–419.

Grotkopp E., Rejmánek M., Sanderson M. J. & Rost T. L. (2004): Evolution of genome size in pines (Pinus) and its life history correlates: supertree analysis. – Evolution (N. Y). 58: 1705–1729.

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 Phytol. 210: 1195–1206.

Halverson K., Heard S. B., Nason J. D. & Stireman J. O. (2008): Origins, distribution, and local co-occurrence of polyploid cytotypes in Solidago altissima (Asteraceae). – Am J Bot 95:50–58.

Hegarty M., Coate J., Sherman-Broyles S., Abbott R., Hiscock S., Doyle J. (2013). Lessons from natural and artificial polyploids in higher plants. *Cytogenet. Genome Res.* 140 204–225 10.1159/000353361

Herben T., Suda J., Klimešová J., Mihulka S., Říha P. & Šímová I. (2012): Ecological effects of cell-level processes: genome size, functional traits and regional abundance of herbaceous plant species. – Ann. Bot. 110: 1357–1367.

Hijmans R. J., Gavrilenko T., Stephenson S., Bamberg J., Salas A. & Spooner D. M. (2007): Geographical and environmental range expansion through polyploidy in wild potatoes (Solanum section Petota). – Glob. Ecol. Biogeogr. 16: 485–495.

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., Espuny L. D. T. & Simmons E. (2010): Stomatal vs. genome size in angiosperms: the somatic tail wagging the genomic dog? – Ann. Bot. 105: 573–584.

Hopping M. E. (1993): Preparation and preservation of nuclei from plant tissues for quantitative DNA analysis by flow cytometry. – New Zeal. J. Bot. 31: 391–401.

Hülgenhof E., Weidhase R. A., Schlegel R. & Tewes A. (1988): Flow cytometric determination of DNA content in isolated nuclei of cereals. – Genome 30:565-569, https://doi.org/10.1139/g88-095.

Husband B. C. (2004): The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. – Biol. J. Linn. Soc. 82: 537–546.

Husband BC, Baldwin SJ, Suda J. The incidence of polyploidy in natural plant populations: major patterns and evolutionary processes (2013): In: Leitch IJ, Greilhuber J, Doležel J, Wendel JF, editors. Plant Genome Diversity 2: Physical Structure, Behaviour and Evolution of Plant Genomes. Springer; Wien: pp. 255–276.

Jarret R. L., Ozias-Akins P., Phatak S., Nadimpalli R., Duncan R. & Hiliard S. (1995): DNA contents in *Paspalum* spp. determined by flow cytometry. – Genetic Resources and Crop Evolution 42(3): 237–242.

Jordan G. J., Carpenter R. J., Koutoulis A., Price A. & Brodribb T. J. (2015): Environmental adaptation in stomatal size independent of the effects of genome size. – New Phytol. 205: 608–617.

Katagiri Y, Hasegawa J, Fujikura U, Hoshino R, Matsunaga S, Tsukaya H. (2016). The Coordination Of Ploidy And Cell Size Differs Between Cell Layers In Leaves. Development 143: 1120-1125.

Kelly L. J., Leitch A. R., Fay M. F., Renny-Byfield S., Pellicer J., Macas J. & Leitch I. J. (2012): Why size really matters when sequencing plant genomes. – Plant Ecol. Divers. 5: 415–425.

Knight C. A. & Ackerly D. D. (2002): Variation in nuclear DNA content across environmental gradients: a quantile regression analysis. – Ecol. Lett. 5: 66-76.

Knight C. A. & Beaulieu J. M. (2008): Genome size scaling through phenotype space. – Ann. Bot. 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. – J. Bot. 1–7.

Knight C. A., Molinari N. A. & Petrov D. A. (2005): The large genome constraint hypothesis: evolution, ecology and phenotype. – Ann. Bot. 95: 177–190.

Krahulcová A., Trávníček P., Krahulec F. & Rejmánek M. (2017): Small genomes and large seeds: chromosome numbers, genome size and seed mass in diploid Aesculus species (Sapindaceae). – Ann. Bot. 119: 957–964.

Kuester A., Conner J. K., Culley T., Baucom R. S. (2014): How weeds emerge: a taxonomic and trait-based examination using United States data. – New Phytologist 202: 1055–1068.

Lawrence M. E. (1985): *Senecio* L. (Asteraceae) in Australia: recombination systems of a polyploid series. – Australian Journal of Botany 33: 209–219.

Lavergne S., Muenke N. J. & Molofsky J. (2010): Genome size reduction can trigger rapid phenotypic evolution in invasive plants. – Ann. Bot. 105: 109–116.

Leitch and Bennett, 2004: Genome downsizing in polyploid plants. – Biol. J. Linn. Soc. 82: 651-663.

Leitch I. J., Bennett M. D. (2007): Genome size and its uses: the impact of flow cytometry. In: Doležel J, Greilhuber J, Suda J, eds. Flow cytometry with plant cells. Analysis of genes, chromosomes and genomes. Weinheim, Germany: Wiley-VCH Verlag, 153–176.

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

Leong-Škorničková J., Šída O., Jarolímová V., Sabu M., Fér T., Trávníček P. & Suda J. (2007): Chromosome numbers and genome size variation in Indian species of Curcuma (Zingiberaceae). – Ann. Bot. 100: 505–526.

Levin D. A. 1983. Polyploidy and novelty in flowering plants. American Naturalist 122: 1 – 25.

Loureiro J., Rodriguez E., Doležel J. & Santos C. (2006): Flow cytometric and microscopic analysis of the effect of tannic acid on plant nuclei and estimation of DNA content. – Ann. Bot. 98: 515–527.

Loureiro J., Trávníček P., Rauchová J., Urfus T., Vít P., Štech M., Castro S. & Suda J. (2010): The use of flow cytometry in the biosystematics, ecology and population biology of homoploid plants. – Preslia 82: 3–21.

Luo C., Chen D., Cheng X., Zhao H. & Huang C. (2017): Genome size estimations in Chrysanthemum and correlations with molecular phylogenies. – Genet. Resour. Crop Evol. 64: 1451–1463.

Lysak M. A., Koch M. A., Beaulieu J. M., Meister A. & Leitch I. J. (2009): The dynamic ups and downs of genome size evolution in Brassicaceae. – Mol. Biol. Evol. 26: 85–98.

Ma J., Devos K. M. & Bennetzen J. L. (2004): Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice. – Genome Res. 14: 860–869.

Mandáková T., Joly S., Krzywinski M., Mummenhoff K. & Lysak M. A. (2010): Fast diploidization in close mesopolyploid relatives of Arabidopsis. – Plant Cell 22: 2277–2290.

Matzk F., Meister A. & Schubert I. (2000): An efficient screen for reproductive pathways using mature seeds of monocots and dicots. – Plant J. 21: 97–108.

McMaster G. S. (1997): Phenology, development, and growth of the wheat (Triticum aestivum L.) shoot apex: a review. – Adv. Agron. 59: 63-118.

McMaster G. S. (2005): Phytomers, phyllochrons, phenology and temperate cereal development. – J. Agric. Sci. 143: 137–150.

Meagher T. R. & Vassiliadis C. (2005): Phenotypic impacts of repetitive DNA in flowering plants. – New Phytol. 168: 71–80.

Morgan H. D. & Westoby M. (2005): The relationship between nuclear DNA content and leaf strategy in seed plants. – Ann. Bot. 96: 1321–1330.

Mukherjee, S. & Sharma, A. K. (1990). Mitotic cycle duration and its relationship with nuclear DNA content, chromosome number and size and rate of growth in species of Acacia growing under stress. Ind. J. Exp. Biol., 28, 508-10.

Nemoto K., Morita S. Baba T (1995): Shoot and root development in rice related to the phyllochron. – Crop Sci. 35: 24-29.

Neumann P., Koblížková A., Navrátilová A., Macas J. (2006): Significant expansion of Vicia pannonica genome size mediated by amplification of a single type of giant retroelement. – Genetics 173: 1047-105.

Noirot M., Barre P., Duperray C., Hamon S. & Kochko A. De. (2005): Investigation on the causes of stoichiometric error in genome size estimation using heat experiments: consequences on data interpretation. – Ann. Bot. 95: 111–118.

Noirot M., Barre P., Louarn J., Duperray C. & Hamon S. (2000): Nucleus-cytosol interactions - A source of stoichiometric error in flow cytometric estimation of nuclear DNA content in plants. – Ann. Bot. 86: 309–316.

Nsabimana A. & Van Staden J. (2006): Ploidy investigation of bananas (Musa spp.) from the National Banana Germplasm Collection at Rubona-Rwanda by flow cytometry. – S Afr J Bot 72:302–305.

Obermayer R. & Greilhuber J. (2005): Does genome size in Dasypyrum villosum vary with fruit colour? – Heredity (Edinb). 95: 91–95.

Ohri D. & Pistrick K. (2001): Phenology and genome size variation in Allium L. - a tight correlation? – Plant Biol. 3: 654–660.

Oliver M. J., Petrov D., Ackerly D., Falkowski P. & Schofield O. M. (2007): The mode and tempo of genome size evolution in eukaryotes. – Genome Res. 17: 594–601.

Otto S. P. & Whitton J. (2000): Polyploid incidence and evolution. – Annu. Rev. Genet. 34: 401-437.

Pandit M. K.,White S. M. & Pocock M. J. O. (2014): The contrasting effects of genome size, chromosome number and ploidy level on plant invasiveness: a global analysis. – New Phytologist 203: 697–703.

Parisod C., Holderegger R. & Brochmann C. (2010): Evolutionary consequences of autopolyploidy. – New Phytol. 186: 5–17.

Paule J., Paule L. & Gömöry D. (2018): Small genome size variation across the range of European beech (Fagus sylvatica). – Plant Syst. Evol. 304: 577–582.

Paule J., Wagner N. D., Weising K. & Zizka G. (2017): Ecological range shift in the polyploid members of the South American genus Fosterella (Bromeliaceae). – Ann. Bot. 120: 233–243.

Peer Y. Van De, Mizrachi E. & Marchal K. (2017): The evolutionary significance of polyploidy. – Nat. Rev. Genet. 18: 411–424.

Pellicer J., Fay M. F. & Leitch I. J. (2010): The largest eukaryotic genome of them all? – Bot. J. Linn. Soc. 164: 10–15.

Pellicer J., Hidalgo O., Dodsworth S. & Leitch I. J. (2018): Genome size diversity and its impact on the evolution of land plants. – Genes (Basel). 9: 10.3390/genes9020088.

Petrov D. A., Sangster T. A., Johnston J. S., Hartl D. L. & Shaw K. L. (2000): Evidence for DNA loss as a determinant of genome size. – Science 287: 1060–1062.

Prančl J., Kaplan Z., Trávníček P. & Jarolímová V. (2014): Genome size as a key to evolutionary complex aquatic plants: polyploidy and hybridization in Callitriche (Plantaginaceae). – PLoS One 9: 10.1371/journal.pone.0105997.

Price H. J., Hodnett G. & Johnston J. S. (2000): Sunflower (Helianthus annuus) leaves contain compounds that reduce nuclear propidium iodide fluorescence. – Ann. Bot. 86: 929–934.

Pyšek P., Skálová H., Čuda J., Guo W.-Y., Suda J., Doležal J., Kauzál O., Lambertini C., Lučanová M., Mandáková T., Moravcová L., Pyšková K., Brix H. & Meyerson L. A. (2018): Small genome separates native and invasive populations in an ecologically important cosmopolitan grass. – Ecology 99: 79-90.

Ramsey J. (2011): Polyploidy and ecological adaptation in wild yarrow. – Proc. Natl. Acad. Sci. 108: 7096–7101.

Rebernig C. A., Weiss-Schneeweiss H., Blöch C., Turner B., Stuessy T. F., Obermayer R., Villaseñor J. L. & Schneeweiss G. M. (2012): The evolutionary history of the white-rayed species of Melampodium (Asteraceae) involved multiple cycles of hybridization and polyploidization. – Am. J. Bot. 99: 1043–1057.

Rejmánek M, Richardson DM, Higgins SI, Pitcairn MJ, Grotkopp E. (2005): Ecology of invasive plants: state of the art. In: Mooney HA, Mack RM, McNeely JA, Neville L, Schei P, Waage J, eds. Invasive alien species: searching for solutions. Washington, DC, USA: Island Press, 104–161.

Rejmánek M. 1996. A theory of seed plant invasiveness: the first sketch. Biological Conservation 78: 171–181.

Rejmánek M. 2000. Invasive plants: approaches and predictions. Austral Ecology 25: 497–506.

SanMiguel P., Tikhonov A., Jin Y. K., Motchoulskaia N., Zakharov D., Melake-Berhan A., Springer P. S., Edwards K. J., Lee M., Avramova Z. & Bennetzen J. L. (1996): Nested retrotransposons in the intergenic regions of the maize genome. – Science 274: 765–768.

Shapiro H. (2004): Practical flow cytometry. – Wiley-Liss, New York.

Shirasu K. (2000): A Contiguous 66-kb Barley DNA Sequence Provides Evidence for Reversible Genome Expansion. – Genome Res. 10: 908–915.

Schmuths H., Meister A., Horres R. & Bachmann K. (2004): Genome size variation among accessions of Arabidopsis thaliana. – Ann. Bot. 93: 317–321.

Schubert I. & Vu G. T. H. (2016): Genome stability and evolution: attempting a holistic view. – Trends Plant Sci. 21: 749–757.

Slovák M., Vít P., Urfus T. & Suda J. (2009): Complex pattern of genome size variation in a polymorphic member of the Asteraceae. – J. Biogeogr. 36: 372–384.

Snodgrass S. J., Jareczek J. & Wendel J. F. (2017): An examination of nucleotypic effects in diploid and polyploid cotton. – AoB Plants 9: 10.1093/aobpla/plw082.

Soltis D. E., Albert V. A., Leebens-Mack J., Bell C. D., Paterson A. H., Zheng C., Sankoff D., dePamphilis C. W., Wall P. K. & Soltis P. S. (2009): Polyploidy and angiosperm diversification. – Am. J. Bot. 96: 336–348.

Soltis D. E., Soltis P. S., Bennett M. D. & Leitch I. J. (2003): Evolution of genome size in the angiosperms. – Am. J. Bot. 90: 1596–1603.

Soltis D. E., Visger C. J. & Soltis P. S. (2014): The polyploidy revolution then and now: Stebbins revisited. – Am. J. Bot. 101: 1057–1078.

Soltis D. E., Buggs R. J. A., Barbazuk W. B., Chamala S., Chester M., Gallegher J. P., Schnable P. S. & Soltis P. S. (2012): The early stages of polyploidy: rapid and repeated evolution in Tragopogon, in Soltis PS, Soltis DE (eds): Polyploidy and Genome Evolution, pp 271–292, Springer, Heidelberg.

Suda J. & Trávníček P. (2006a): Estimation of relative nuclear DNA content in dehydrated plant tissues by flow cytometry. – Curr Protoc Cytom 38:7.30.1–7.30.14.

Suda J. & Trávníček P. (2006a): Reliable DNA ploidy determination in dehydrated tissues of vascular plants by DAPI flow cytometry - new prospects for plant research. – Cytom. Part A 69: 273–280.

Suda J., Meyerson L. A., Leitch I. J. & Pyšek P. (2015): The hidden side of plant invasions: the role of genome size. – New Phytol. 205: 994-1007.

Suda J., Weiss-Schneeweiss H., Tribsch A., Schneeweiss G. M., Trávníček P. & Schönswetter P. (2007): Complex distribution patterns of di-, tetra-, and hexaploid cytotypes in the European high mountain plant Senecio carniolicus (Asteraceae). – Am. J. Bot. 94: 1391–1401.

Sugiyama S., Yamaguchi K. & Yamada T. (2002): Intraspecific phenotypic variation associated with nuclear DNA content in Lolium perenne L. – Euphytica 128: 145–151.

Šímová I. & Herben T. (2012): Geometrical constraints in the scaling relationships between genome size, cell size and cell cycle length in herbaceous plants. – Proc. R. Soc. B Biol. Sci. 279: 867–875.

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

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

Šmarda P., Bureš P. & Horová L. (2007): Random distribution pattern and non-adaptivity of genome size in a highly variable population of Festuca pallens. – Ann. Bot. 100: 141–150.

Šmarda P., Bureš P., Horová L. & Rotreklová O. (2008): Intrapopulation genome size dynamics in Festuca pallens. – Ann. Bot. 102: 599–607.

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

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 Phytol. 200: 911–921.

Šmarda P., Horová L., Bureš P., Hralová I. & Marková M. (2010): Stabilizing selection on genome size in a population of Festuca pallens under conditions of intensive intraspecific competition. – New Phytol. 187: 1195–1204.

Temsch E. M., Greilhuber J., Hammett K. R. W. & Murray B. G. (2008): Genome size in Dahlia Cav. (Asteraceae-Coreopsideae). – Plant Syst. Evol. 276: 157–166.

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

Tsukaya H. (2013): Does ploidy level directly control cell size? Counterevidence from Arabidopsis genetics. – PLoS One 8: 1–7.

Veselý P., Bureš P., Šmarda P. & Pavlíček T. (2012): Genome size and DNA base composition of geophytes: the mirror of phenology and ecology? – Ann. Bot. 109: 65–75.

Vicient C., Suoniemi A., Anamthawat-Jónsson K., Tanskanen J., Beharav A., Nevo E. & Schulman A. (1999): Retrotransposon BARE-1 and Its Role in Genome Evolution in the Genus Hordeum. – Plant Cell 11: 1769–1784.

Vidic T., Greilhuber J., Vilhar B. & Dermastia M. (2009): Selective significance of genome size in a plant community with heavy metal pollution. – Ecol. Appl. 19: 1515–1521.

Vilhar B., Greilhuber J., Koce J. D., Temsch E. M. & Dermastia M. (2001): Plant genome size measurement with DNA image cytometry. – Ann. Bot. 87: 719–728.

Vinogradov A. E. (2003): Selfish DNA is maladaptive: evidence from the plant Red List. – Trends Genet. 19: 609–614.

Vinogradov A. E. (2004): Genome size and extinction risk in vertebrates. – Proceedings of the Royal Society of London, Series B 271:1701–1705.

Vítová J., Vít P. & Suda J. (2015): Rare occurrence of reciprocal hybridization in a sympatric population of the Czech stenoendemic Dianthus arenarius subsp. bohemicus and widespread D. carthusianorum. – Preslia 87: 329–345.

Vitte C. & Bennetzen J. L. (2006): Analysis of retrotransposon structural diversity uncovers properties and propensities in angiosperm genome evolution. – Proc. Natl. Acad. Sci. 103: 17638–17643.

Voglmayr H. (2000): Nuclear DNA amounts in mosses (Musci). – Ann. Bot. 85: 531–546.

Vu G. T. H., Schmutzer T., Bull F., Cao H. X., Fuchs J., Tran T. D., Jovtchev G., Pistrick K., Stein N., Pecinka A., Neumann P., Novak P., Macas J., Dear P. H., Blattner F. R., Scholz U. & Schubert I. (2015): Comparative genome analysis reveals divergent genome size evolution in a carnivorous plant genus. – Plant Genome 8: 10.3835/plantgenome2015.04.0021.

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

Walker D. J., Monino I. & Correal E. (2006): Genome size in Bituminaria bituminosa (L.) C. H. Stirton (Fabaceae) populations: separation of " true " differences from environmental effects on DNA determination. – Environ. Exp. Bot. 55: 258-265.

Weiss-Schneeweiss H., Emadzade K., Jang T.-S. & Schneeweiss G. M. (2013): Evolutionary consequences, constraints and potential of polyploidy in plants. – Cytogenet. Genome Res. 140: 137-150.

Weiss-Schneeweiss H., Greilhuber J. & Schneeweiss G. M. (2006): Genome size evolution in holoparasitic Orobanche (Orobanchaceae) and related genera. – Am. J. Bot. 93: 148–156.

Wendel J. F., Jackson S. A., Meyers B. C. & Wing R. A. (2016): Evolution of plant genome architecture. – Genome Biol. 17: 10.1186/s13059-016-0908-1.

Whitney K. D., Baack E. J., Hamrick J. L., Godt M. J. W., Barringer B. C., Bennett M. D., Eckert C. G., Goodwillie C., Kalisz S., Leitch I. J. & Ross-Ibarra J. (2010): A role for nonadaptive processes in plant genome size evolution? – Evolution (N. Y). 64: 2097–2109.

Wicker T., Yahiaoui N., Guyot R., Schlagenhauf E., Liu Z.-D., Dubcovsky J. & Keller B. (2003): Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and Am genomes of wheat. – Plant Cell 15: 1186–1197.

PROFESSIONAL CURRICULUM VITAE

Magdalena Lučanová (née Kubešová)

Date and place of birth
4. 11. 1981 České Budějovice, Czech Republic

Employement and university education:
od 2017: researcher at the Department of Botany, Faculty of Science at the Charles University in Prague (part time job at the projects of prof. RNDr. Karol Marhold, CSc. and RNDr. Filip Kolář, Ph.D.)

from 2017: researcher at the Department of Botany, Faculty of Science at the University of South Bohemia in České Budějovice (part time job at the projects of RNDr. Petr Koutecký, Ph.D.)

from 2014: maternity leave (only 15% job at the project of prof. Karol Marhold)

from 2012: researcher at the Department of Botany, Faculty of Science at the Charles University in Prague (Laboratory of flow cytometry - part time job at the projects of prof. Jan Suda and prof. Karol Marhold)

2007 – 2018: PhD. studies at the Department of Botany, Faculty of Science at the Charles University in Prague and at the Botanical institute, Academy of Sciences in Průhonice (part time job in Laboratory of flow cytometry)

2004 – 2007: MSc. studies of Botany (Systematics of higher plants) at the Biological Faculty at the University of South Bohemia in České Budějovice

2001 – 2004: Bc. studies of Biology at the Biological Faculty at the University of South Bohemia in České Budějovice

Professional activities:
supervision of batchelor and master theses
participation in teaching of Karyological and palynological practical class – teaching of practical fluorescent and classic karyology in botany
occasionally teaching of botany at the field practical classes and teaching practical class of morphology and system of higher planst
participation in many botanical exkursions and workshops
participation in floristic research of South Bohemia
2013 - Prize of Josef Hlávka for the Best Students and Graduates
2003 – 2011 - member of Working group for Biological olympiad preparation for students of higher school
participation in preparation of biological competition Zelená stezka – Zlatý list for students of primary school
leading of field botanical exkursions for secondary school teachers
leading of field botanical exkursions for students of pedagogy
2002, 2004, 2008 – habitats mapping for Nature Conservation Agency of the Czech Republic – project Natura 2000
2001-2002 – breeding of aphidofagous insect – Institute of Entomology, Biology centre CAS
from 2001 – member of Southbohemian branch of Czech Botanical Society

from 2000 – member of Czech Botanical Society
1998 – 2001 - participation at the Ecological Olympiad:
member of winning team in national round
1997 – 2001: participation at the Biological Olympiad
2000: 4. place in national round
2001: 5. place in national round + participation at the International Biological Olympiad in Brussels; Praemium Bohemiae Award (Bohuslav Jan Horáček Foundation) for secondary school students

Professional specialization
Systematics of higher plants (phylogeography, taxonomy), genome size in plants, polyploidy, cytogenetic methods (flow cytometry, karyology, FISH), reproductive modes, invasive plants, floristics

Experiences
Flow cytometry, karyology, grounding in fluorescent cytogenetics (FISH), vegetative scanning, experimental study of reprosystems

Grant projects

2017: Parallel adaptation to alpine environments in wild *Arabidopsis*.- GAČR 17-20357Y, RNDr. Filip Kolář, Ph.D. (researcher)

2017: Misunderstood patterns of cryptic variation in aquatic plants caused by hybridization and polyploidization events. GAČR 41-17-06825S, RNDr. Petr Koutecký (researcher)

2016: Pathways and consequences of polyploidy and chromosomal variation in plant evolution – three cases from the family Brassicaceae. GAČR 16-10809S, prof. Karol Marhold (researcher)

2012: Evolutionary patterns in polyploid complexes: congruent or distinct stories in three examples from the Brassicaceae family? GAČR P506/12/0668, prof. Karol Marhold (researcher)

2010: Immigrated vs. in situ created diversity in postglacial areas: A story of the polyploid complex *Galium pumilum* agg. GAČR P506/10/0704, prof. Jan Suda (researcher)

2005: Structure of populations of *Agrostis stolonifera* in the Czech Republic from the karyological point of view. Mattoni Awards for Studies of Biodiversity and Conservation Biology 2005. Magdalena Kubešová (main investigator)

2003: Biosystematic study of *Agrostis stolonifera* group. FRVŠ G4 b 1850/ 2003, Milan Štech, Magdalena Lučanová (co-investigator)

Synanthropic Flora of Villages in the Blanský les Protected Landscape Area. Mattoni Awards for Studies of Biodiversity and Conservation Biology, Jakub Těšitel, Magdalena Kubešová, Petr Koutecký (co-investigator)

SCI publications

Pyšek P., Skálová H., Čuda J., Guo W.-Y., Suda J., Doležal J., Kauzál O., Lambertini C., **Lučanová M.**, Mandáková T., Moravcová L., Pyšková K., Brix H., Meyerson L. A. (2018): Small genome separates native and invasive populations in an ecologically important cosmopolitan grass. – Ecology 99(1): 79-90.

Šrámková-Fuxová G., Záveská E., Kolář F., **Lučanová M.**, Španiel S., Marhold K. (2017): Range-wide genetic structure of *Arabidopsis halleri* (Brassicaceae): glacial persistence in multiple refugia and origin of the Northern Hemisphere disjucntion. – Botanical Journal of the Linnean Society 185(3): 321-342.

Kolář F, **Lučanová M**, Záveská E, Fuxová G, Mandáková T, Španiel S, Senko D, Svitok M, Kolník M, Gudžinskas Z, Marhold K (2016): Ecological segregation does not drive the intricate parapatric distribution of diploid and tetraploid cytotypes of the *Arabidopsis arenosa* group (Brassicaceae). – Biological Journal of the Linnean Society 119(3):673-688.

Meyerson L. A., Cronin J. T., Bhattarai G. P., Brix H., Lambertini C., **Lučanová M.**, Rinehart S., Suda J., Pyšek P. (2016): Do ploidy level and nuclear genome size and latitude of origin modify the expression of *Phragmites australis* traits and interactions with herbivores? – Biological Invasions 18(9): 2531-2549.

Kolář F.*, Fuxová G.*, Záveská E.*, Nagano A. J., Hyklová L., **Lučanová M.**, Kudoh H., Marhold K. (2016): Northern glacial refugia and altitudinal niche divergence shape genome-wide differentiation in the emerging plant model *Arabidopsis arenosa*. – Molecular Ecology 25(16): 3929-3949, *equal contribution.

Pyšek P., Manceur A., Alba C., McGregor K., Pergl J., Štajerová K., Chytrý M., Danihelka J., Kartesz J., Klimešová J., **Lučanová M.**, Moravcová L., Nishino M., Sádlo J., Suda J., Tichý L. & Kühn I. (2015): Naturalization of central European plants in North America: species traits, habitats, propagule pressure, residence time. – Ecology 96: 762–774 (doi: 10.1890/14-1005.1).

Hohmann N., Schmickl R., Chiang T.Y., **Lučanová M.**, Kolář F., Marhold K., Koch M.A. (2014): Taming the wild: resolving the gene pools of non-model *Arabidopsis* lineages. – BMC Evolutionary Biology 14: 224.

Těšitel J., Těšitelová T., Bernardová A., Drdová-Janková E., **Lučanová M.**, Klimešová J. (2014): Demographic population structure and fungal associations of plants colonizing High Arctic glacier forelands, Petuniabukta, Svalbard. - Polar Research 33:20797.

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., **Lučanová M.**, Fuxová G., Záveská E., Španiel S., Marhold K. (2013): Cytological variation of *Arabidopsis arenosa* in its Carpathian diversity centre. – Acta Biologica Cracoviensia Series Botanica 55:51.

Fuxová G., Záveská E., Kolář F., **Lučanová M.**, Záveská E., Španiel S., Marhold K. (2013)**:** Phylogeography and taxonomy of *Arabidopsis halleri* in the Carpathians. – Acta Biologica Cracoviensia Series Botanica 55:44.

Krejčíková J., Sudová R., **Lučanová M.**, Trávníček P., Urfus T., Vít P., Weiss-Schneeweiss H., Kolano B., Oberlander K., Dreyer L. L. & Suda J. (2013): High ploidy diversity and distinct

patterns of cytotype distribution in a widespread species of *Oxalis* in the Greater Cape Floristic Region. – Annals of Botany 111: 641–649.

Kolář F., **Lučanová M.**, Vít P., Urfus T., Chrtek J., Fér T., Ehrendorfer F. & Suda J. (2013): Diversity and endemism in deglaciated areas: Ploidy, relative genome size and niche differentiation in the *Galium pusillum* complex (Rubiaceae) in Northern and Central Europe. – Annals of Botany 111: 1095–1108.

Trávníček P., Jersáková J., Kubátová B., Krejčíková J., Bateman R. M., **Lučanová M.**, Krajníková E., Těšitelová T., Štípková Z., Amardeilh J.-P., Brzosko E., Jermakowicz E., Cabanne O., Durka W., Efimov P., Hedrén M., Hermosilla C. E., Kreutz K., Kull T., Tali K., Marchand O., Rey M., Schiestl F. P., Čurn V. & Suda J. (2012): Minority cytotypes in European populations of the *Gymnadenia conopsea* complex (Orchidaceae) greatly increase intraspecific and intrapopulation diversity. – Annals of Botany 110: 977–986.

Kolář F., **Lučanová M.**, Těšitel J., Loureiro J. & Suda J. (2012): Glycerol-treated nuclear suspensions-an efficient preservation method for flow cytometric analysis of plant samples. – Chromosome research 20(2): 303-315.

te Beest M., Le Roux J. J., Richardson D. M., Brysting A. K., Suda J., **Kubešová M.** & Pyšek P. (2012): The more the better? The role of polyploidy in facilitating plant invasions. – Annals of Botany 109: 19-45.

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

Dušková E., Kolář F., Sklenář P., Rauchová J., **Kubešová M.**, Fér T., Suda J. & Marhold K. (2010): Genome size correlates with growth form, habitat and phylogeny in the Andean genus *Lasiocephalus* (Asteraceae). – Preslia 82: 127-148

Hülber K., Sonnleitner M., Flatscher R., Berger A., Dobrovsky R., Niessner S., Nigl T., Schneeweiss G. M., **Kubešová M.**, Rauchová J., Suda J. & Schönswetter P. (2009): Ecological segregation drives fine-scale cytotype distribution of *Senecio carniolicus* in the Eastern Alps. – Preslia 81(3): 309–319

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

Suda J., Loureiro J., Trávníček P., Rauchová J., Vít P., Urfus T., **Kubešová M.**, Dreyer L. L., Oberlander K. C., Wester P. & Roets F. (2009): Flow cytometry and its applications in plant population biology, ecology and biosystematics: New prospects for the Cape flora *(conference abstract)*. – South African Journal of Botany 75(2): 389

Other publications

Lepší P., Lepší M., Boublík K., Štech M., Hans V. [eds.] (2013): Červená kniha květeny jižní části Čech. – Jihočeské muzeum v Českých Budějoivicích, 501 pp.

Kolář F., Matějů J., **Lučanová M.**, Chlumská Z., Černá K., Prach J., Baláž V., Falteisek L. (2012): Ochrana přírody z pohledu biologa. – Dokořán, Praha. 216 pp. + 16 stran barevná příloha, ISBN 978-80-7363-414-8

Baláž V., Falteisek L., Chlumská Z., Kolář F., **Kubešová M.**, Matějů J., Prach J. & Rezková K. (2010): Ochrana přírody z pohledu biologa. Biologická olympiáda 2010 – 2011, 45. ročník, přípravný text pro kategorie A, B. – Česká zemědělská univerzita v Praze. Praha. 191 pp., ISBN 978-80-213-2085-7

Karlík P., Ešnerová J., Urfus T., Kuneš I., Baláš M., Vítámvás J., Koňasová T., **Kubešová M.**, Fér T. & Vít P. (2010): Problematika určování druhů bříz *Betula* L. ve světle průtokové cytometrie. In: Bříza – strom roku 2010. Sborník z konference (Kostelec nad Černými lesy 23. září 2010). – Česká zemědělská univerzita v Praze, Praha. 117 pp., ISBN 978-80-213-2098-7

Štajerová K., Kolář F.**,** **Kubešová M.**, Sekerka L., Molem K. & Lepš J. (2008): Role nepůvodních druhů během sukcese na opuštěných políčkách Papuy Nové Guineje [The role of alien plants in succession in a shifting agriculture of Papua New Guinea] – Zpr. Čes. Botan. Společ. 43, Materiály 23: 63-72

Kolář F., **Kubešová M.**, Těšitel J. & Koutecký P. (2007): Květena vesnic CHKO Blanský les [Synanthropic flora of villages in the Blanský les Protected Landscape Area] – Zpr. Čes. Botan. Společ., Praha, 42(1): 89-105

Kubešová M. (2007): Phenotypic plasticity and cytotypes of *Agrostis stolonifera* in the Czech Republic. – Ms. Thesis, in Czech – 46 p., Faculty of Biological Sciences, The University of South Bohemia, České Budějovice, Czech Republic

Falteisek L., Fikáček M., Hodač L., Chmátal L., Jedelský P., Koutecká E., Koutecký P., Krtková J., **Kubešová M.**, Lišková J., Matějů J., Smýkal V., Šípek P. (2006): Organismy a abiotické faktory prostředí. Přípravný text Biologické olympiády kategorie A, B. – Národní institut dětí a mládeže MŠMT. Praha

Lučan R. K., Radil J. & **Kubešová M.** (2006): Prostorová aktivita mateřské kolonie netopýra vodního (*Myotis daubentonii*) v různých fázích reprodukčního cyklu: telemetrické přiblížení. In: BRYJA J. & ZUKAL J. (eds.): *Zoologické dny*. *Brno 2006*. *Sborník abstraktů z konference 9.–10. února* [*2006*]. Ústav biologie obratlovců AV ČR, Brno, 268 pp.

Lučan R. K., Radil J., Jahelková H., Pithartová T., **Kubešová M.**, Půža V., Šabacká M., Frantová D., Nedoma R. & Pech P. (2005): Spatial activity of maternity colonies of Daubenton's bats in different phases of the reproductive cycle as revealed by radiotracking. In: BRYJA J. & ZUKAL J. (eds.): *Zoologické dny*. *Brno 2006*. *Sborník abstraktů z konference 9.–10. února* [*2006*]. Ústav biologie obratlovců AV ČR, Brno, 268 pp.

Čepička I., Jedelský P., **Kubešová M.**, Lišková J., Matějů J., Strádalová V. (2005): Pohyb. Přípravný text Biologické olympiády kategorie A, B. – Národní institut dětí a mládeže MŠMT. Praha

Kubešová M. (2004): Karyological study of *Agrostis stolonifera* in the Czech Republic. – Bc. Thesis, in Czech – 28 p., Faculty of Biological Sciences, The University of South Bohemia, České Budějovice, Czech Republic

Conference participations

Fuxová, G., Záveska, E., Kolář, F., **Lučanová, M.**, Španiel, S. & Marhold, K.: Phylogeography and taxonomy of *Arabidopsis halleri* in the Carpathians. In: Anonymous, ed., Biogeography of the

Carpathians: Evolution of biodiversity in a spatiotemporal context, September 26-28, 2013, Kraków, Poland. Acta Biol. Cracov., ser. Bot. 55 (Suppl. 1): 44. (*poster*)

Kolář, F., **Lučanová, M.**, Fuxová, G., Záveská, E., **Španiel, S.** & Marhold, K.: Cytological variation of *Arabidopsis arenosa* in its Carpathian diversity centre. In: Anonymous, ed., Biogeography of the Carpathians: Evolution of biodiversity in a spatiotemporal context, September 26-28, 2013, Kraków, Poland. Acta Biol. Cracov., ser. Bot. 55 (Suppl. 1): 51. (*poster*)

Kolář F., **Lučanová M.**, Píšová S., Fér T., Vít P., Urfus T., Chrtek J., Koutecký P., Suda J.: Evolutionary history of *Galium pumilum* agg. polyploid complex in deglaciated Europe. (*poster*) - BioSyst. EU 2013 Global systematics!, 18. - 22. 2. 2013, Vienna, Austria

Těšitel J., Bernardová A., Janková Drdová E**., Lučanová M.**, Těšitelová T., Klimešová J.: Brave pioneers: growth and flowering of plants colonizing glacier forelands in Petuniabukta, Svalbard. BES Annual Meeting, 17. – 20. 12. 2012, University of Birmingham, United Kingdom

Meyerson L. A., Pyšek P., Suda J., Lambertini C., Brix H., Cronin J. T., **Lučanová M.**: The hidden side of plant invasions: The role of genome-wide processes in plant success. (*oral*) – Ecological Society of America Annual Meeting, 5. – 10. 8. 2012, Portland, Oregon, USA

Lučanová M., Loureiro J., Suda J.: The dynamics and evolutionary consequences of intraspecific variation in genome size. (*poster*) – International Conference on Polyploidy, Hybridization, and Biodiversity, 7. – 10. 5. 2012, Průhonice, Česká Republika

Kolář F., **Lučanová M.**, Píšová S., Fér T., Vít P., Urfus T., Chrtek J., Koutecký P., Suda J.: Evolutionary history of *Galium pumilum* agg. polyploid complex in deglaciated Europe. (*poster*) – International Conference on Polyploidy, Hybridization, and Biodiversity, 7. – 10. 5. 2012, Průhonice, Česká Republika

Krejčíková J., Trávníček P., **Lučanová M.**, Kalůsková J., Weiss-Schneeweiss H., Dreyer L. L., Oberlander K. C., Suda J.: Hidden diversity of the Cape flora: Large ploidy variation in Oxalis hirta L. (*poster*) – International Conference on Polyploidy, Hybridization, and Biodiversity, 7. – 10. 5. 2012, Průhonice, Česká Republika

Temsch E. M., Jang T., **Lučanová M.**, Parker J., Speta F., Weiss-Schneeweiss H.: Genome size variation in *Prospero* (Hyacinthaceae). (*poster*) – International Conference on Polyploidy, Hybridization, and Biodiversity, 7. – 10. 5. 2012, Průhonice, Česká Republika

Suda J., Krejčíková J., **Lučanová M.**, Sudová R., Trávníček P., Urfus T., Vít P., Dreyer L., Oberlander K., Weiss-Schneeweiss H.: Understanding the whole genome processes in a world biodiversity hotspot: flow cytometric investigations of the Cape Flora. (*oral*) – International Conference on Polyploidy, Hybridization, and Biodiversity, 7. – 10. 5. 2012, Průhonice, Česká Republika

Kubešová M., Loureiro J. & Suda J. (2011): Genome size variation in *Taraxacum stenocephalum*. (*poster*) – BioSytematics Berlin, 21. – 27. 2. 2011, Německo

Suda J., Krejčíková J., **Lučanová M.**, Sudová R., Trávníček P., Urfus T., Vít P., Dreyer L., Oberlander K., Weiss-Schneeweiss H. (2011): Understanding the whole genome processes in a world biodiversity hotspot: flow cytometric investigations of the Cape flora. ČSAC

Suda J., Kolář F., **Kubešová M.**, Rauchová J., Trávníček P., Urfus T. &Vít P. (2010): The use of flow cytometry in plant biosystematics. – oral presentation of JS at "XIth Symposium of the International Organization of Plant Biosystematists", Aurangabad, India.

Karlík P., Ešnerová J., Urfus T., Kuneš I., Baláš M., Vítámvás J., Koňasová T., **Kubešová M.**, Fér T. & Vít P. (2010): Problematika určování druhů bříz *Betula* L. ve světle průtokové cytometrie. In: Bříza – strom roku 2010. Sborník z konference (Kostelec nad Černými lesy 23. září 2010). – Česká zemědělská univerzita v Praze, Praha. 117 pp., ISBN 978-80-213-2098-7

Kubešová M., Kolář F., Loureiro J., Těšitel J., Trávníček P. & Suda J. (2009): Glycerol-based preservation technique of plant tissues for flow cytometric analyses. *(oral)* – Analytická cytometrie V – konference ČSAC, Olomouc, Česká republika

Kubešová M., Loureiro J., Trávníček P., Urfus T., Vít P. & Suda J. (2009): Patterns and dynamics of genome size variation in *Taraxacum stenocephalum* (Asteraceae). *(poster)* – International Conference on Polyploidy, Hybridization and Biodiversity, Saint Malo, France

Kolář F., Štech M., Trávníček P., Rauchová J., Urfus T., Vít P., **Kubešová M.** & Suda J. (2009): Primary and secondary contact zones of di- and tetraploid *Knautia arvensis* agg. (Dipsacaceae). *(poster)* **–** International Conference on Polyploidy, Hybridization and Biodiversity, Saint Malo, France

Urfus T., Krahulec F., Vít P., **Kubešová M.** & Trávníček P. (2009): Variation in *Pilosella officinarum* F. W. Schultz et Sch. Bip. In central Europe: ploidy levels, breeding systems and their correlation with morphology. *(poster)* **–** International Conference on Polyploidy, Hybridization and Biodiversity, Saint Malo, France

Vít P., Suda J., **Kubešová M.**, Urfus T. & Trávníček P. (2009): Hybridization of *Cerastium alsinifolium*: caryological and molecular evaluation of serpentine endemic species. *(poster)* **–** International Conference on Polyploidy, Hybridization and Biodiversity, Saint Malo, France

Suda J., Rauchová J., Trávníček P., **Kubešová M.** & Loureiro J. (2009): Applications of flow cytometry to plant invasion biology. *(přednáška)* – 10th International Conference on the Ecology and Management of Alien Plant Invasions, Stellenbosch, South Africa

Suda J., Loureiro J., Trávníček P., Rauchová J., Vít P., Urfus T., **Kubešová M.**, Dreyer L. L., Oberlander K. C., Wester P. & Roets, F. (2009): Flow cytometry and its applications in plant population biology, ecology and biosystematics: New prospects for the Cape flora. (*hlavní přednáška*) **–** 19.1.2009-22.1.2009. 35th Annual Conference of the South African Association of Botanists, Stellenbosch, JAR

Štajerová K., Kolář F., **Kubešová M.**, Sekerka L., Molem K. & Lepš J. (2007): The role of alien plants in a shifting agriculture in Papua New Guinea. *(poster)* **–** Conference of the Czech botanical society, Plant invasions in the Czech Republic: present state, research and management., Praha, Czech Republic

Těšitel J., Kolář F., **Kubešová M.** & Koutecký P. (2006): Factors Affecting Synanthropic Flora of Villages in Blanský Les, South Bohemia. *(poster)* – 12. Österreichisches Botanikertreffen, Kremsmünster, Austria. Beiträge zur Naturkunde Oberösterreichs 16: 636

Štech M., **Kubešová M.** & Horváthová V. (2004): Bemerkungen zur Gattung *Agrostis* in der Tschechischen Republik. *(poster)* – 11. Österreichisches Botanikertreffen, Wien, Austria. Kurzfassungen der Beiträge: 70

Meyerson L. A., Cronin J. T., Bhattarai G. P., Brix H., Lambertini C., **Lučanová M**., Rinehart S., Suda J., Pysek P. (2016): Do ploidy level and nuclear genome size and latitude of origin modify the expression of *Phragmites australis* traits and interactions with herbivores? – Biological Invasions 18(9): 2531-2549

Educational stay

February + March, October + November 2011 – Plant Evolutionary Cytogenetics Group (laboratory of Hanna Schneeweiss), Department of Systematic and Evolutionary Botany, University of Vienna, Austria – cytogenomics (karyology, FISH)

Other interests

travelling: Austria, Slovakia, Slovenia, Hungary, Croatia, Monte Negro, Italy, France, Switzerland, Spain (+ Canary Islands), Portugal, Belgium, Germany, Norway (+ Svalbard), Finland, Estonia, Lithuania, Latvia, Poland, Ukraine, Romania, Cyprus, Turkey, Syria, Sweden, Jordan, Russian, Nepal, Malaysia, Papua-New Guinea, Egypt, Republic of South Africa, Venezuela, Great Britain, Bahrain

birdwatching