Letters

Genome size and genomic GC content evolution in the miniature genome-sized family Lentibulariaceae

Introduction

Since the first measurements of genome size in the early 1950s (Swift, 1950), researchers have tried to estimate the maximum capacity of plants for genome growth and the minimum DNA content essential for proper cell function. Plants with smaller genome size soon became important subjects of study as it was possible to completely sequence their genome without the need for processing a huge amount of uninformative, repetitive DNA (Flagel & Blackman, 2012) which covers the bulk of their genomes (Bennetzen *et al.*, 2005; Ambrožová *et al.*, 2011). Unsurprisingly, the first nearly-complete genome sequence published was *Arabidopsis thaliana* (Arabidopsis Genome Initiative, 2000) as it was then considered to be the plant with the smallest genome (Bennett & Leitch, 2005). Analysis of the *Arabidopsis* genome (1C ≈ 157 Mbp; Bennett *et al.*, 2003) and the virtual removal of repetitive DNA and duplicated genes lead to the theoretical estimate of the minimum size of gene complement needed for plant functioning as 1C ≈ 50 Mbp (Bennett & Leitch, 2005).

Such small genomes were soon discovered by Greilhuber *et al.* (2006) in the carnivorous family Lentibulariaceae (Lamiales). They documented the genome size of two samples of *Genlisea aurea* as low as 1C = 63.4 Mbp (originally, one sample of *G. aurea* was misidentified as *G. margaretae*). In addition to this, relatively small genomes with 1C < 1000 Mbp were found to prevail in all three monophyletic lineages of the family, that is, the genera *Genlisea*, *Pinguicula* and *Utricularia*. Until recently, however, genome size is known only for *c.* 8% of the Lentibulariaceae species, which contains 29 *Genlisea*, *c.* 233 *Utricularia* and *c.* 101 *Pinguicula* species. This provides the challenge to search for other species with miniature genomes and possible genomic models.

Detailed sequence analyses of *G. aurea* and *Utricularia gibba* which have been published in the last months (Ibarra-Laclette *et al.*, 2013; Leushkin *et al.*, 2013) clearly confirm the expected minimalistic genome composition of these species and show that this is reached with the removal of duplicated or otherwise redundant genes (e.g. genes relating to roots in rootless *U. gibba*) and virtually all noncoding repetitive DNA (transposable elements). This finding suggests a limited role of repetitive DNA in the regulation of complex eukaryotic genomes. However, this tells nothing about the reasons and driving forces behind this extreme DNA shrinkage, which is important for understanding why variations in plant genome size and genome architecture exist. Clearly, answering this question will require future, targeted comparisons between species selected with regard to the evolutionary history of miniaturization events and the specific hypotheses addressed.

In order to extend the contemporary pool of suitable model species and to improve current knowledge on the history of miniaturization events in Lentibulariaceae, an extensive survey and phylogeny-based analysis of genome size evolution in 119 (*c.* 35%) of Lentibulariaceae species is presented. Genomic DNA base composition (GC content) is also reported for all taxa to add further to the knowledge of the process of genome miniaturization.

Materials and Methods

Samples for the measurements were mainly from the authors' private and institutional collections with a few species provided by other Czech carnivorous plant collections (Supporting Information Tables S1, S2). In most samples, original species identification was verified based on their flower morphologies. The genome size (referred to as the 1C value in this paper) and GC content were measured with flow cytometry on two CyFlow flow cytometers (Partec GmbH, Münster, Germany) using the base unspecific, intercalating fluorochrome propidium iodide (PI) and the AT-selective DAPI (4',6-diamidino-2-phenylindole). The details of the procedure and the concentrations of reagents followed Šmarda *et al.* (2008). The fully-sequenced *Oryza sativa* subsp. *japonica* 'Nipponbare' (1C = 388.8 Mbp, GC = 43.6%; International Rice Genome Sequencing Project, 2005) was the internal reference standard and four other internal standards, whose genome size and GC content were derived from comparison with this *Oryza* cultivar, were used (Methods S1). Every sample was measured at least three times (on different days) and replicated measurements were averaged (Table S3).

In addition to the measured genomic characters, information on chromosome number, life-form, altitudinal and latitudinal distribution, and distributions on particular continents was compiled from the literature or based on personal experience (Table S2, Methods S1).

For the purpose of phylogeny-based analyses, we constructed a Bayesian, ultrametric phylogenetic tree for the measured species (Figs 1, S1). The tree is based on the concatenated alignment of available sequence data from one nuclear (ITS) and three plastid regions (*rps*16, *mat*K, *trn*L-F) searched in the NCBI GenBank database (Benson *et al.*, 2013; Table S1). The details on the tree construction are found in Methods S1.

The relationships between genome size, GC content and other trait variables were tested using the phylogenetic generalized least-squares (*pgls*) in the *caper* package (function *pgls*; Orme *et al.*,

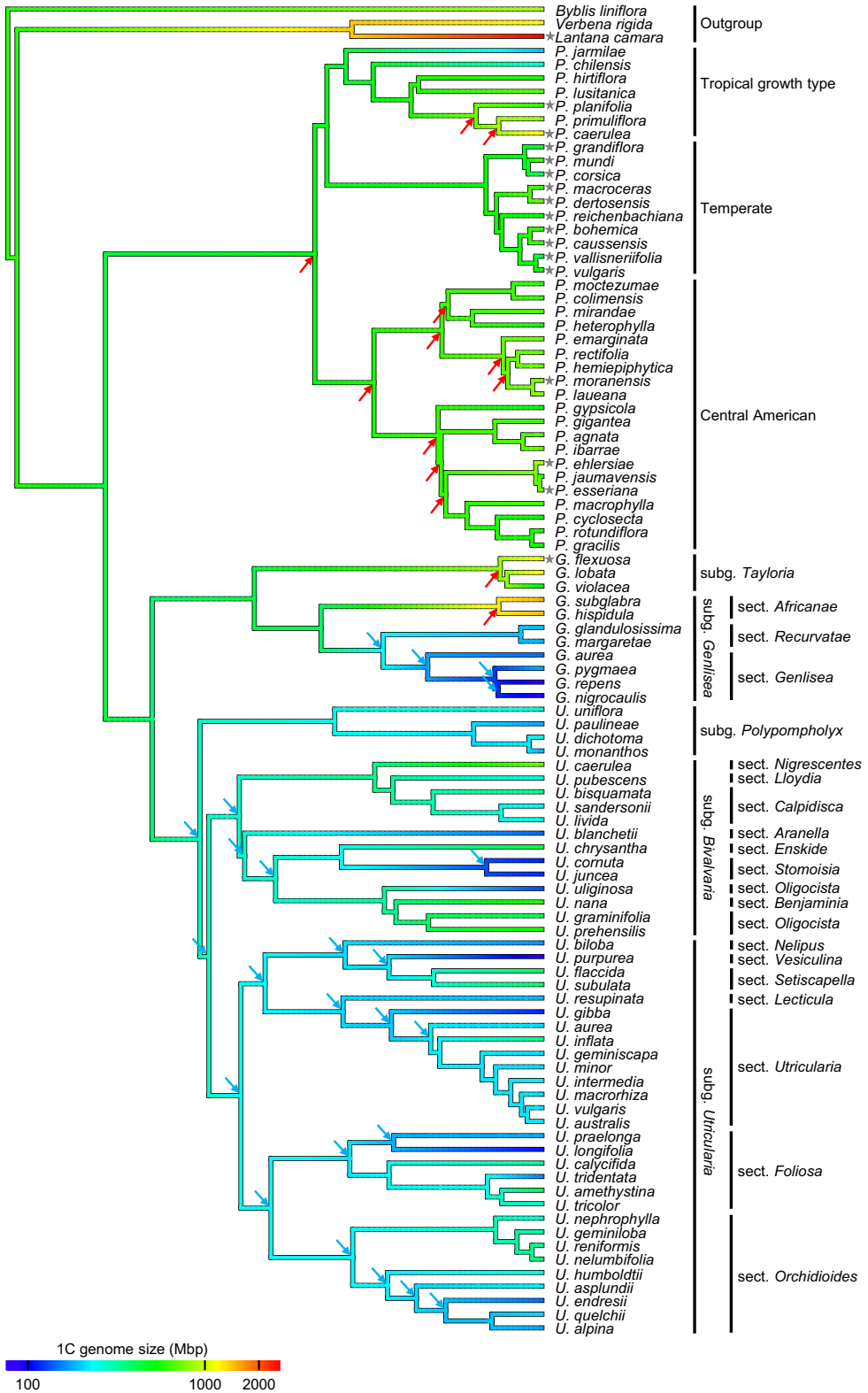

Fig. 1 Ancestral state reconstruction of genome size in Lentibulariaceae. Significant decreases and increases of genome size ($P < 0.05$) are marked, respectively, with blue and red arrows. Genome sizes referring to samples with probable recent polyploid origin are marked with grey asterisks.

2012) of R (R Core Team, 2013). Ancestral genome sizes were reconstructed using maximum likelihood (using function *ace* from R package *ape* v. 3.0-10; Paradis *et al.*, 2004) and visualized on the tree with *contMap* function of R package *phytools* v. 0.2-80 (Revell, 2012). Significant increases or decreases in genome size (Fig. 1) or GC content (Fig. S2) were detected by comparing the actual

New Phytologist

ancestral node values vs the random node values obtained with the same procedure, calculated with randomly reshuffled tip values. The randomization was repeated 999 times. All the statistics were done with $\log_{10}$ transformed data on genome sizes and logit transformed values (with natural logarithm) of the GC contents.

Results and Discussion

Summary and reliability of the data

The Lentibulariaceae species clearly have smaller genomes when compared with the related families of the Lamiales (Fig. 2). Approximately 95% of the 119 measured taxa have a 1C-value smaller than 1000 Mbp and 19 have a genome size smaller than that of *Arabidopsis* (Table 1). Our results mostly agree with those of Greilhuber *et al.* (2006), although some minor differences may appear due to the slightly different genome sizes assumed for the genome size standards (cf. Methods S1). The species with the smallest known genome size in the Lentibulariaceae (and all angiosperms) still remains *G. aurea* (63.4 Mbp; Greilhuber *et al.*, 2006). Our measurement of the genome size of this species (1C = 131 Mbp), however, is almost exactly double that reported by Greilhuber *et al.* (2006) and corresponds to a different ploidy level ('tetraploid') within this morphologically and karyologically variable species (Rivadavia, 2002; Albert *et al.*, 2010). Similarly, in *Pinguicula ehlersiae*, the two-fold difference in the measured genome size (1C = 978 Mbp in our study vs 1C = 487 Mbp by Greilhuber *et al.*, 2006) also corresponds with the existence of two

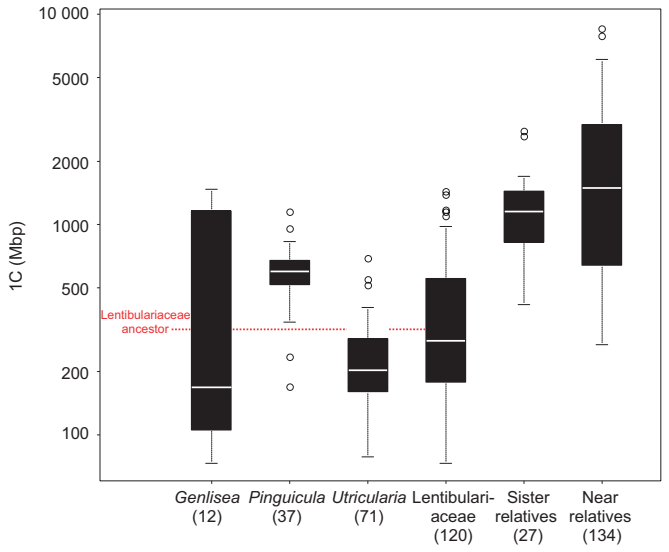

Fig. 2 Comparison of the measured genome sizes of Lentibulariaceae genera with genome size data from other Lamiales families in the Plant DNA C-value Database (Bennett & Leitch, 2005). Boxplots show the median (thick horizontal line), interquartile range (boxes), nonoutlier range (whiskers) and outliers (circles). The red horizontal line indicates the predicted genome size of the common Lentibulariaceae ancestor. Sister relatives: Acanthaceae, Bignoniaceae, Martyniaceae, Pedaliaceae, Verbenaceae; near relatives: Lamiaceae, Orobanchaceae, Paulowniaceae, Phrymaceae. Numbers of species displayed per group are given in brackets. The Lentibulariaceae family has a significantly smaller genome size than both its sister relatives and near relatives (two-sample Wilcoxon test; both comparisons $P < 0.05$).

Table 1 Results of genome size and genomic DNA base composition (GC content) measurements together with published data on chromosome number

Species	1C (Mbp)	GC (%)	2*n*
Genlisea			
aurea	131	38.9	(52^{G})
flexuosa	1121	44.3	–
glandulossisima[A]	169	34.1	–
hispidula	1417	41.5	–
lobata	1200	44.0	16^{G}
margaretae[A]	168	34.0	–
nigrocaulis clone1	80	38.9	–
nigrocaulis clone2	73	–	–
pygmaea	161	40.7	–
repens	77	38.8	–
subglabra	1471	41.7	–
violacea	460	43.7	–
Pinguicula			
agnata	651	41.1	22^{H}
bohemica	590	39.8	64^{H},(32^{H})
caerulea	1178	40.8	32^{H}
chilensis	241	39.4	16^{H}
colimensis	600	42.5	22^{H}
corsica	344	39.9	16^{H}
hirtiflora	529	40.7	28^{H}
cyclosecta	500	40.0	22^{H}
dertosensis[A]	708	38.9	64^{H}
ehlersiae	978	40.4	44^{H},(22^{H})
emarginata	717	40.9	22^{H}
esseriana	760	40.5	32^{H}
gigantea	598	40.8	22^{H}
gracilis	518	40.9	22^{H}
grandiflora	424	39.1	32^{H}
gypsicola	501	40.3	22^{H}
hemiepiphytica	702	41.8	22^{H}
heterophylla	522	39.7	22^{H}
ibarrae	676	41.2	22^{H}
jarmilae	173	42.4	–
jaumavensis	495	40.4	22^{H}
laueana	789	41.6	22^{H}
longifolia ssp. *caussensis*[A]	623	39.2	32^{H}
lusitanica	665	43.2	12^{H}
macroceras[A]	591	39.9	64^{H}
macrophylla	627	41.1	22^{H}
mirandae	663	41.2	–
moctezumae	572	41.6	22^{H}
moranensis	713	41.8	22^{H},(44^{H})
mundi	616	39.9	64^{H}
planifolia	583	43.1	32^{H}
primuliflora	830	39.8	22^{H}
rectifolia	676	41.5	22^{H}
reichenbachiana[A]	469	38.7	32^{H}
rotundiflora	547	40.8	22^{H}
vallisneriifolia	344	39.4	32^{H}
vulgaris	583	38.8	64^{H}
Utricularia			
alpina	159	39.9	18^{E}
amethystina[A]	382	40.1	–
asplundii	202	41.1	–
aurea	193	38.3	42^{E},80^{D}
aureomaculata[A]	104	35.5	–
australis	200	40.0	36^{E},38^{E},40^{E},44^{E}
bifida	245	42.4	–

Table 1 (Continued)

Species	1C (Mbp)	GC (%)	2*n*
biloba	150	39.1	–
bisquamata	308	44.5	–
blanchetii	129	40.1	–
bremii	299	40.1	36[F]
caerulea	706	43.2	36[E],40[E]
calycifida	287	43.9	–
chrysantha	404	40.3	–
cornuta	102	39.8	18[E]
dichotoma	246	41.4	28[E]
dimorphanta	187	38.6	44[F]
endresii	133	38.4	–
flaccida	349	42.1	–
floridana	100	39.9	–
fulva	120	38.4	–
geminiloba	287	38.4	–
geminiscapa[A]	191	39.1	–
gibba	103	39.9	28[E]
graminifolia[A]	377	40.8	–
hirta	152	41.3	–
humboldtii	228	41.6	–
hydrocarpa	107	36.8	–
inflata	313	40.1	–
intermedia	203	39.2	44[E]
involvens[A]	287	41.2	–
juncea	106	39.4	18[E]
laxa	381	45.1	–
livida	239	42.0	36[E]
longeciliata	234	43.3	–
longifolia	97	41.1	–
macrorhiza	193	39.4	40[E],42[E],44[E]
menziesii	274	41.4	–
microcalyx	197	42.9	–
minor	190	38.8	36[E],40[E],44[E]
minutissima	203	42.1	–
monanthos	165	40.9	–
nana[A]	561	40.5	–
nelumbifolia	349	39.7	–
nephrophylla	247	37.0	–
ochroleuca	203	39.2	40[E],44[E],46[E],48[E]
paulineae	159	39.6	–
praelonga[A]	162	42.4	–
prehensilis	526	42.8	–
pubescens	232	42.8	–
purpurea	79	34.4	–
quelchii	191	40.7	–
radiata	163	38.4	–
reflexa	270	38.8	–
reniformis	292	38.0	–
resupinata	169	39.0	36[E],44[C]
rostrata	191	41.6	–
sandersonii	204	41.4	–
stellaris	192	39.5	40[B],42[E]
striata	117	41.1	–
stygia	315	40.6	–
subulata	340	41.2	30[E]
tenuicaulis	183	38.5	40[D]
tricolor	262	41.4	28[E]
tridentata[A]	142	39.3	–
uliginosa	116	39.6	–
uniflora	245	40.8	56[E]
volubilis	211	40.6	–
vulgaris	199	39.3	36[E],40[E],44[E]

Table 1 (Continued)

Species	1C (Mbp)	GC (%)	2*n*
warburgii	324	44.3	–
welwitschii	298	42.0	–

[A]Species where flowering individuals were not available for identification. Chromosome numbers were taken from [B]Sarkar *et al.* (1980), [C]Löve & Löve (1982), [D]Tanaka & Uchiyama (1988), [E]Taylor (1989), [F]Rahman *et al.* (2001), [G]Greilhuber *et al.* (2006), [H]Casper & Stimper (2009). Chromosome counts that probably do not refer to the measured plants are in brackets.

ploidy levels (2*n* = 22, 44; Casper & Stimper, 2009). Some other disagreements reported here, such as in *Genlisea violacea*, are perhaps due to the unrecognized taxonomic diversity, noting that the *G. violacea* complex has only recently been divided into five separate species (Fleischmann *et al.*, 2011). Unrecognized karyological variability (aneuploidy) known in several Lentibulariaceae species (cf. Table 1) may cause further differences.

Our GC content estimate of *U. gibba* (39.9%) agrees well with that reported for the complete genome sequence (GC = 40.0%; Ibarra-Laclette *et al.*, 2013). However, some difference is found between our GC content estimate of *G. aurea* (38.9%) and that reported from the partial genomic sequence (40.0%) by Leushkin *et al.* (2013). This difference might arise from gaps in the genomic data and/or may correspond to a different ploidy between races of *G. aurea*, with our sample possibly being tetraploid.

Genome size evolution

The genome size of the common ancestor of the family is estimated to be 414 Mbp (95% confidence interval: 284–603 Mbp), which is less than that of any of the close Lentibulariaceae relatives (Fig. 2). In spite of this relatively small ancestral genome size, further miniaturizations can be recognized in the evolution of the family. The exceptional tendency for genome miniaturization is most remarkable in *Utricularia* (Fig. 1), where ultra-small genomes

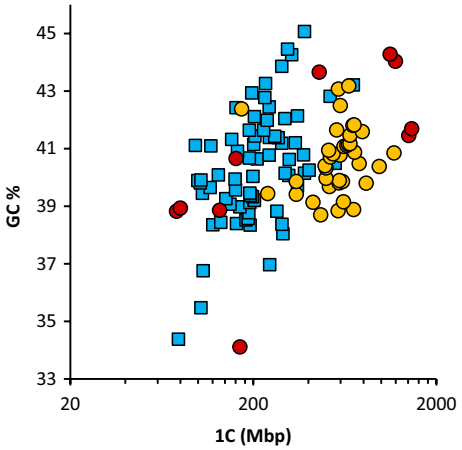

Fig. 3 Comparison of genome sizes with genomic DNA base composition (GC content) in particular Lentibulariaceae genomes. GC content is positively correlated with genome size in *Utricularia* (blue squares) and *Genlisea* (red circles) but not in *Pinguicula* (yellow circles) (*pgls* $\alpha = 0.05$).

(1C < 100 Mbp) have evolved independently in three clades: *U.* sect. *Foliosa* – (*U. longifolia*), *U.* sect. *Vesiculina* – (*U. purpurea*) and *U.* sect. *Utricularia* (*U. floridana*; not shown in the phylogenetic tree because of absence of sequence data). Beyond *Utricularia*, other prominent miniaturization is found in *Genlisea.* Here, significant genome miniaturization accompanies the evolution of *G.* sect. *Genlisea* and *G.* sect. *Recurvatae* (Fig. 1). These sections typically contain species with very small genomes (all 1C < 170 Mbp; the smallest one in our dataset represented by *G. nigrocaulis* clone 2, 1C = 73 Mbp). This contrasts with other *Genlisea* clades possessing larger genomes, with *G. subglabra* (1C = 1471 Mbp) having the largest genome in the whole family (Fig. 1).

In contrast to *Utricularia* and *Genlisea*, genome size evolution in *Pinguicula* is less dramatic, showing a consistent tendency for genome expansion. The only miniaturizations appear in *P. jarmilae* and *P. chilensis* (Fig. 1). The quiet genome size evolution of *Pinguicula* allows some of the genome size differences to be ascribed to recent polyploidy, e.g. between the closely related *P. jaumavensis* ($2n = 2x = 22$, 1C = 495 Mbp) and *P. ehlersiae* ($2n = 4x = 44$, 1C = 978 Mbp). In *Utricularia* and *Genlisea* the chromosome counts do not correlate with the observed genome sizes in any predictable way. This suggests that recent polyploidy has only a limited effect on the extreme size dynamics of Lentibulariaceae genomes. Consequently, this variation is most likely to be caused by differences in the content of noncoding repetitive DNA, as was indeed documented by the recent detailed genomic data (Ibarra-Laclette *et al.*, 2013; Leushkin *et al.*, 2013). Variation in repetitive DNA is the general reason for large-scale variation in plant genome sizes (Bennetzen *et al.*, 2005; Grover & Wendel, 2010). In *Genlisea* and *Utricularia*, however, the turnover of noncoding DNA is unusually high, with large genome size differences generated relatively quickly, even among closely related species. This provides a unique opportunity for effective study of the principles and the reasons of genome size variation in plants.

While the outcome of genome miniaturization in Lentibulariaceae is recognized, the reasons for and driving forces behind this drastic genome miniaturization remain unclear. The obvious interest in Lentibulariaceae lies in carnivory, which is an adaptation to nutrient-poor environments. As expected by Leitch & Leitch (2008), the plants with larger genomes could be disadvantaged in such places, possibly because of phosphorus and/or nitrogen limitation. Members of the Lentibulariaceae usually grow under harsh conditions of nutrient-poor soils or waters. Here, the evolutionary pressure on genome size could be very strong, thus keeping the genome sizes of Lentibulariaceae species very low. However, species with miniaturized genomes did not show any common morphological and ecological features, and genome size showed no relationship with life-form or any ecological variables tested (*pgls*, $P > 0.05$). This indicates that nutrient availability or environmental selection play perhaps only a minor role in driving the extreme genome miniaturizations. Nevertheless, nutrient limitation and associated carnivory may have been the actual reason for the initial genome size reduction in the Lentibulariaceae ancestor as well as the factor preventing excessive genome growth. This hypothesis needs further testing by comparing the genome sizes of carnivorous taxa with their noncarnivorous relatives.

Albert *et al.* (2010) and Ibarra-Laclette *et al.* (2011a,b) presented a unique mechanism of energy production which leads to the formation of reactive oxygen species. These can damage DNA molecules, possibly causing loss of the damaged DNA region. *Utricularia* and *Genlisea* might therefore be in an active process of genome downsizing without an external selection pressure. Both *Utricularia* and *Genlisea* (but not *Pinguicula*) are also known for extremely high substitution rates (Jobson & Albert, 2002; Müller *et al.*, 2004; Ibarra-Laclette *et al.*, 2011a,b), which could correspond with the influence of these reactive oxygen species. Such processes might indeed serve as a mechanistic explanation of the extremely high mutation rates and variable genome sizes observed in both genera. However, even with the data available on the complete sequence of *U. gibba*, the role of increased mutation rate in driving genome shrinkage in Lentibulariaceae genomes could not be verified (Ibarra-Laclette *et al.*, 2013).

GC content

This survey of the genomic GC contents in Lentibulariaceae has shown that both genome quantity and quality have a surprising pattern of variation within the group. The unusually wide variation of genomic GC contents appearing even within a genus (10.7% difference in *Utricularia* and 10.2% in *Genlisea*) is particularly interesting. This variation covers a substantial part of the entire known genomic GC content variation in vascular plants (ranging from 33% to 50%; Šmarda & Bureš, 2012) and represents the highest difference so far determined within a plant family or genus. The notably low GC contents are found in *G.* sect. *Recurvatae* (*G. margaretae*, *G. glandulossisima* with GC = 34.0% and 34.1%, respectively) and in *U. purpurea* (GC = 34.4%; Tables 1, S3, Fig. S2). The increased GC content is typical of *G.* sect. *Tayloria* (all GC > 43.7%) and occurs also in several clades of *Utricularia* with the most GC rich Lentibulariaceae genomes found in *U. laxa* (GC = 45.1%; Tables 1, S3).

GC content correlates well with genome size in both GC variable genera (Fig. 3), *Utricularia* (*pgls*, $\lambda = 1$, $P < 0.001$) and *Genlisea* (*pgls*, $\lambda = 1$, $P = 0.019$; excluding the outlying *G.* sect. *Recurvatae*). In *Pinguicula*, the phylogenetic trend between GC content and genome size is absent (*pgls*, $\lambda = 1$, $P = 0.497$; Fig. 3), perhaps due to the fact that *Pinguicula* genomes are mostly shaped by polyploidy (whole genome duplication) which has no direct effect on the overall genomic GC content. The correlation between GC content and genome size in *Genlisea* and *Utricularia* indicates that the extreme GC content variation of their genomes primarily relates to the high genome size dynamics and to the processes of genome miniaturization and genome growth. Assuming that coding DNA would form only a minor part of the removed or amplified DNA (because of the direct effect of gene loss or duplication on plant fitness), the most intuitive explanation for this trend would be the preferential removal or amplification of GC-rich, noncoding DNA (Šmarda & Bureš,

2012; Veselý *et al.*, 2012). However, the exact proof of this, with detailed sequence data, still poses a challenge.

Given that coding DNA is regularly the most GC-rich component of plant genomes and noncoding DNA is usually GC-poor when compared with genes (cf. Šmarda & Bureš, 2012), one would expect high GC-richness in the miniature Lentibulariaceae genomes. This work has, however, revealed several species whose very small genomes were surprisingly GC-poor (*Genlisea margaretae*, *G. glandulossisima* and *Utricularia purpurea* with 34.0%, 34.1% and 34.4%, respectively). These approach the minimum genomic GC content yet known in some Cyperaceae and Juncaceae species (Šmarda & Bureš, 2012; Šmarda *et al.*, 2012; Lipnerová *et al.*, 2013; P. Šmarda *et al.*, unpublished). These whole genome GC contents are even lower than the GC content of the noncoding genome fraction of *U. gibba* (GC = 35.9%; Ibarra-Laclette *et al.*, 2013), indicating a very different genome structure of the GC-poor species compared with the other miniature-sized genomes of Lentibulariaceae. Such a low GC content could be reached with the frequent presence of AT-rich, noncoding DNA, which is less probable due to the minimal genome size of all three species and the expected high content of coding DNA. Therefore, the depletion of GC bases must also include the coding DNA and perhaps affects the structure of genes. This suggests the existence of an additional mechanism shaping the miniature Lentibulariaceae genomes, together with the removal and amplification of noncoding DNA. Sequencing of any of the GC-poor miniature genomes of Lentibulariaceae and their comparison with the available genomic sequences for GC-rich *G. aurea* and *U. gibba* (Ibarra-Laclette *et al.*, 2013; Leushkin *et al.*, 2013) now seems to be a promising way of detecting this mechanism, which might substantially improve our understanding of the reasons behind the evolution of the GC-poor genome architectures also found in other small-genomed plants.

Acknowledgements

This study was supported by the Czech Science Foundation (projects P505/11/0551, P506/11/0890, P504/11/0783, 13-29362S) and the Academy of Sciences of the Czech Republic (to L.A., project RVO 67985939). The authors thank Brian G. McMillan for language correction, Miroslav Macák, Jaroslav Neubauer, Michal Rubeš and David Švarc for providing fresh plant material of several species, and Andreas Fleischmann for help with identification of three *Genlisea* samples.

Adam Veleba[1]*, Petr Bureš[1], Lubomír Adamec[2], Petr Šmarda[1], Ivana Lipnerová[1] and Lucie Horová[1]

[1]Department of Botany and Zoology, Masaryk University, Kotlářská 2, Brno, CZ-61137, Czech Republic;
[2]Institute of Botany, Dukelská 135, Třeboň, CZ-37982, Czech Republic
(*Author for correspondence: tel +420 728 071 865; email 184653@mail.muni.cz)

References

Albert VA, Jobson RW, Michael TP, Taylor DJ. 2010. The carnivorous bladderwort (*Utricularia*, Lentibulariaceae): a system inflates. *Journal of Experimental Botany* **61**: 5–9.

Ambrožová K, Mandáková T, Bureš P, Neumann P, Leitch IJ, Koblížková A, Macas J, Lysák MA. 2011. Diverse retrotransposon families and an AT–rich satellite DNA revealed in giant genomes of *Fritillaria* lilies. *Annals of Botany* **107**: 255–268.

Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**: 796–815.

Bennett MD, Leitch IJ. 2005. Nuclear DNA amounts in angiosperms: progress, problems and prospects. *Annals of Botany* **95**: 45–90.

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

Bennetzen JL, Ma J, Devos KM. 2005. Mechanisms of recent genome size variation in flowering plants. *Annals of Botany* **95**: 127–132.

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2013. GenBank. *Nucleic Acids Research* **41**: D36–D42.

Casper SJ, Stimper R. 2009. Chromosome numbers in *Pinguicula* (Lentibulariaceae): survey, atlas, and taxonomic conclusions. *Plant Systematics and Evolution* **277**: 21–60.

Flagel LE, Blackman BK. 2012. The first ten years of plant genome sequencing and prospects for the next decade. In: Wendel F, Greilhuber J, Doležel J, Leitch IJ, eds. *Plant genome diversity, vol. 1*. Vienna, Austria: Springer, 1–15.

Fleischmann A, Rivadavia F, Gonella PM, Heubl G. 2011. A revision of *Genlisea* subgenus *Tayloria* (Lentibulariaceae). *Phytotaxa* **33**: 1–40.

Greilhuber J, Borsch T, Müller K, Worberg A, Porembski S, Barthlott W. 2006. Smallest angiosperm genomes found in Lentibulariaceae, with chromosomes of bacterial size. *Plant Biology* **8**: 770–777.

Grover CE, Wendel JF. 2010. Recent insights into mechanisms of genome size change in plants. *Journal of Botany* **2010**: 382732.

Ibarra-Laclette E, Albert VA, Herrera-Estrella A, Herrera-Estrella L. 2011a. Is GC bias in the nuclear genome of the carnivorous plant *Utricularia* driven by ROS–based mutation and biased gene conversion? *Plant Signaling & Behavior* **6**: 1631–1634.

Ibarra-Laclette E, Albert VA, Pérez-Torres CA, Zamundio-Hernández F, Ortega-Estrada MJ, Herrera-Estrella A, Herrera-Estrella L. 2011b. Transcriptomics and molecular evolutionary rate analysis of the bladderwort (*Utricularia*), a carnivorous plant with a minimal genome. *BMC Plant Biology* **11**: 101.

Ibarra-Laclette E, Lyons E, Hernández-Guzmán G, Pérez-Torres CA, Carretero-Paulet L, Chang T-H, Lan T, Welch AJ, Juárez MJA, Simpson J *et al.* 2013. Architecture and evolution of a minute plant genome. *Nature* **498**: 94–98.

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

Jobson RW, Albert VA. 2002. Molecular rates parallel diversification contrasts between carnivorous plant sister lineages. *Cladistics* **18**: 127–136.

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

Leushkin EV, Sutormin RA, Nabieva ER, Penin AA, Kondrashov AS, Logacheva MD. 2013. The miniature genome of a carnivorous plant *Genlisea aurea* contains a low number of genes and short non-coding sequences. *BMC Genomics.* **14**: 476.

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

Löve A, Löve D. 1982. In: Löve A, ed. *IOPB chromosome number reports LXXVII. Taxon* **31**: 766–768.

Müller K, Borsch T, Legendre L, Porembski S, Theisen I, Barthlott W. 2004. Evolution of carnivory in Lamiales and Lentibulariaceae. *Plant Biology* **6**: 477–490.

Orme D, Freckleton R, Thomas G, Petzold T, Fritz S, Isaac N, Pears W. 2012. *Caper: Comparative Analyses of Phylogenetics and Evolution in R. Version 0.5.* [WWW document] URL http://cran.r-project.org/web/packages/caper/caper.pdf [accessed 25 April 2013].

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

R Core Team. 2013. *R: a language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. URL http://www.R-project.org/.

Rahman MO, Adamec L, Kondo K. 2001. Chromosome numbers of *Utricularia bremii* and *Utricularia dimorphantha* (Lentibulariaceae). *Chromosome Sciences* **5**: 105–108.

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

Rivadavia F. 2002. *Genlisea aurea* St. Hil. *Carnivorous Plant Newsletter* **31**: 54–56.

Sarkar AK, Datta N, Chatterjee U. 1980. In: Löve A, ed. *Chromosome number reports LXVII. Taxon* **29**: 360–361.

Šmarda P, Bureš P. 2012. The variation of base composition in plant genomes. In: Wendel JF, Greilhuber J, Doležel J, Leitch IJ, eds. *Plant genome diversity, vol. 1.* Vienna, Austria: Springer, 209–235.

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

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

Swift H. 1950. The constancy of desoxyribose nucleic acid in plant nuclei. *Proceedings of the National Academy of Sciences, USA* **36**: 643–654.

Tanaka R, Uchiyama H. 1988. Chromosomes of four species of *Utricularia* in Japan. *Journal of Japanese Botany* **63**: 219–223.

Taylor P. 1989. *The Genus* Utricularia. London, UK: The Royal Botanical Gardens, Kew.

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

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Detailed phylogenetic tree for the measured taxa.

Fig. S2 Ancestral state reconstruction of genomic GC content in Lentibulariaceae.

Table S1 List of species locations, details on subgeneric classification, and NCBI accesion numbers of used sequences

Table S2 Environmental data of species

Table S3 Detailed results of flow cytometry measurements

Methods S1 Details of the flow cytometry measurements, ecological traits and methods of phylogenetic tree construction.

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

Key words: carnivorous plants, flow cytometry, GC content, genome miniaturization, genome size evolution, genomic DNA base composition, genomic models, Lentibulariaceae.