The LINNEAN SOCIETY of London BOTANICAL Journal of the Linnean Society

Botanical Journal of the Linnean Society, 2010, **164**, 10–15. With 3 figures

The largest eukaryotic genome of them all?

JAUME PELLICER*, MICHAEL F. FAY FLS and ILIA J. LEITCH FLS

Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond Surrey TW9 3AB, UK

Received 30 June 2010; revised 12 July 2010; accepted for publication 19 July 2010

We report the largest eukaryotic genome to date in the monocot *Paris japonica* (Melanthiaceae, 1C = 152.23 pg), measured using flow cytometry. This value is 15% larger than any previous estimate and extends the range of genome sizes to *c.* 2400-fold across angiosperms and *c.* 66 000-fold across eukaryotes. © 2010 The Linnean Society of London, *Botanical Journal of the Linnean Society*, 2010, **164**, 10–15.

ADDITIONAL KEYWORDS: C-value – flow cytometry – genome size – guard cells – *Kinugasa* – Melanthiaceae – *Paris* – stomata.

INTRODUCTION

The diversity of eukaryotic genome sizes has long fascinated, but at the same time puzzled, scientists who have asked how and why such diversity evolved. These are important questions because we know that the total amount of DNA in the nucleus has both biological and ecological consequences that affect the distribution and persistence of biodiversity. There is a staggering diversity of genome sizes (the amount of DNA in the nucleus) in eukaryotes, with available data for over 10 000 species showing that they currently vary *c.* 57 000-fold (this is less than often quoted because the commonly cited extreme values for amoebae, e.g. 700 000 Mbp for *Amoeba dubia*, are excluded due to considerable uncertainty about their accuracy; Gregory, 2005b). The smallest genome so far reported is in the microsporidian *Encephalitozoon intestinalis*, which parasitizes a range of mammals, including humans (Vivares, 1999). Its genome comprises just 0.0023 pg of DNA (= 1C-value; Greilhuber *et al.*, 2005), which corresponds to 2.3 Mbp (1 pg = 978 Mbp; Doležel *et al.*, 2003). At the other end of the range, the genome of the marbled lung fish, *Protopterus aethiopicus*, with *c.* 130 000 Mbp (1C = 132.83 pg; Pedersen, 1971), is one of the largest ever reported. Given that the length of one nucleotide is estimated to be *c.* 0.34 nm, this diversity translates into lengths of only *c.* 2 mm of DNA per somatic nucleus in *E. intestinalis* and *c.* 88 m in *P. aethiopicus*, with our own genome (1C = 3 pg) measuring *c.* 2 m. Such enormous variation and lack of apparent correlation with organismal complexity has long caught the attention of biologists, including Thomas (1971), who coined the phrase 'the C-value paradox' (more recently termed 'the C-value enigma'; Gregory, 2005b).

Across angiosperms, genome size has been revealed to be especially diverse, ranging *c.* 2000-fold. The carnivorous plant *Genlisea margaretae* Hutch. (Lentibulariaceae) has the smallest angiosperm genome so far reported, with only 63.4 Mbp (1C = 0.0648 pg) of DNA (Greilhuber *et al.*, 2006; Chase *et al.*, 2009), approximately 40% of the value for the genetic model species *Arabidopsis thaliana* (L.) Heynh. (157 Mbp; Bennett *et al.*, 2003). At the upper end of the range, large genomes, such as that found in *Fritillaria assyriaca* Baker (1C = 127.4 pg; Bennett & Smith, 1976), and, more recently, in the hybrid *Trillium* × *hagae* Miyabe & Tatew. (1C = 132.50 pg; Zonneveld, 2010) have been reported. Despite the existence of such large genomes, nearly all angiosperm taxa have small genomes, which has led to hypotheses relating to the evolutionary costs of genome obesity in plants (Knight, Molinari & Petrov, 2005). Among the factors that play an important role in the diversification of genome sizes in plants, polyploidy and the accumulation of repetitive DNA

*Corresponding author. E-mail: j.pellicer@kew.org

sequences, often derived from retrotransposons, are of special interest (Bennett & Leitch, 2005; Parisod *et al*., 2009), as are their mechanisms of regulation (Lisch, 2009, Grover & Wendel, 2010). Episodes of polyploidy are frequent in plants (Soltis & Soltis, 2000; Wendel, 2000; Cui *et al*., 2006), leading to genome reorganization at functional and structural levels that occurs during and after genome duplication, some of them involving genome size changes (Leitch & Bennett, 2004).

Paris japonica (Franch. & Sav.) Franch. (also known as *Kinugasa japonica* (Franch. & Sav.) Tatew. & Sutô.) is a rhizomatous geophyte endemic to Japan, where it inhabits the subalpine regions of the mountains of northern and central Honshu. It is a member of the family Melanthiaceae as currently circumscribed (APG II, 2003; Zomlefer *et al*., 2006; APG III, 2009). *Paris* L., *Trillium* L. and related genera were previously often included in Trilliaceae (e.g. Takhtajan, 1983, 1997; Thorne, 1992), but molecular studies have shown Trilliaceae to be embedded in Melanthiaceae, leading to them being treated as tribe Parideae Bartl. of Melanthiacae (e.g. Chase *et al*., 1995, 2000; Fuse & Tamura, 2000; Rudall *et al*., 2000; Zomlefer *et al*., 2001, 2006).

As part of ongoing research into the causes and consequences of genome-size diversity in plants (particularly monocots), we have found, as reported here, the largest eukaryotic genome so far known. This emphasizes the need for an in-depth understanding about the biological processes leading to these large genomes.

MATERIAL AND METHODS

Study species

Specimens of *Paris japonica* studied (*J. Pellicer s.n*., K) were collected from the foothills of Shirouma-dake Nagano-ken, Honshu, and donated to the Royal Botanic Gardens, Kew, by the Japanese Alpine Rock Garden Society.

Chromosome preparations and ploidy determination

Root-tip meristems were obtained from plants growing in pots at the Royal Botanic Gardens, Kew. They were pretreated in a saturated solution of 1-bromonaphthalene at 20 °C for 24 h. Material was then fixed in absolute ethanol and glacial acetic acid (3:1) and stored at 4 °C. To make chromosome preparations, roots were washed in double distilled water for 20 min, hydrolysed in 1 M hydrochloric acid (HCl) for 5 min at 60 °C, stained with Schiff's reagent for 30 min and squashed on slides in a drop of 2% aceto-orcein. Metaphase plates were photographed with a digital camera (SPOT RT; Diagnostic Inc.) mounted on a Zeiss Axioplan Imaging microscope.

Nuclear DNA content estimations

The genome size of *P. japonica* was measured using three calibration standards to obtain a mean nuclear DNA content. For each analysis, four specimens were selected and three independent samples were prepared as follows. Young leaves of *P. japonica* and the calibration standard were co-chopped in 2 mL of the 'General purpose isolation buffer' (Loureiro *et al*., 2007) with the addition of 3% PVP-40 following the one-step procedure described by Doležel, Greilhuber & Suda (2007). The suspension of nuclei was filtered through a 30-μm nylon mesh, then 30 μl of a solution of 100 μg mL^{-1} ribonuclease A (RNase A; Sigma) was added, and finally the nuclei were stained with propidium iodide (Sigma; 1 mg mL^{-1}) at a final concentration of 60 μg mL^{-1}. Samples were kept on ice for 30 min and analysed using a Partec Cyflow SL3 flow cytometer fitted with a 100-mW green solid state laser (Cobolt Samba). For each run, 5000 nuclei were analysed and three runs were made per sample.

The first analysis used *Allium cepa* L. 'Ailsa Craig' as the calibration standard [1C = 16.75 pg (Van't Hof & Sparrow, 1963); Fig. 1A], selected because it has the largest genome of the widely used and commonly accepted calibration standards (e.g. Bennett & Smith, 1976; Doležel *et al*., 2007). However, it is generally agreed that the genome size of the internal calibration standard should be as close as possible to the species of interest to avoid potential technical problems relating to, e.g., linearity. As the genome size of *A. cepa* is *c.* 10 × smaller than *P. japonica*, two additional genome-size estimates were made using calibration standards with genomes closer in size to *P. japonica*. The first experiment used *Trillium rivale* S.Watson (a member of a genus closely related to *Paris*) with a genome size estimated to be 1C = 27.11 pg using *A. cepa* as a calibration standard (Fig. 1B), and the second one used *Trillium sessile* L. (Fig. 1C) as the standard, with an even larger 1C-value of 54.08 pg (also estimated using *A. cepa*).

RESULTS

Using *A. cepa*, the genome size of *P. japonica* was estimated to be 1C = 152.13 ± 0.39 pg. Similar results were obtained with *T. rivale* (1C = 152.15 ± 0.54 pg) and *T. sessile* (1C = 152.41 ± 0.42 pg) as the internal standards. Thus, the genome size of the species comprises *c.* 150 000 Mbp (1C = 152.23 pg). Fluorescence histograms illustrating the results obtained by flow cytometry are presented in Fig. 1. The plants of *P.*

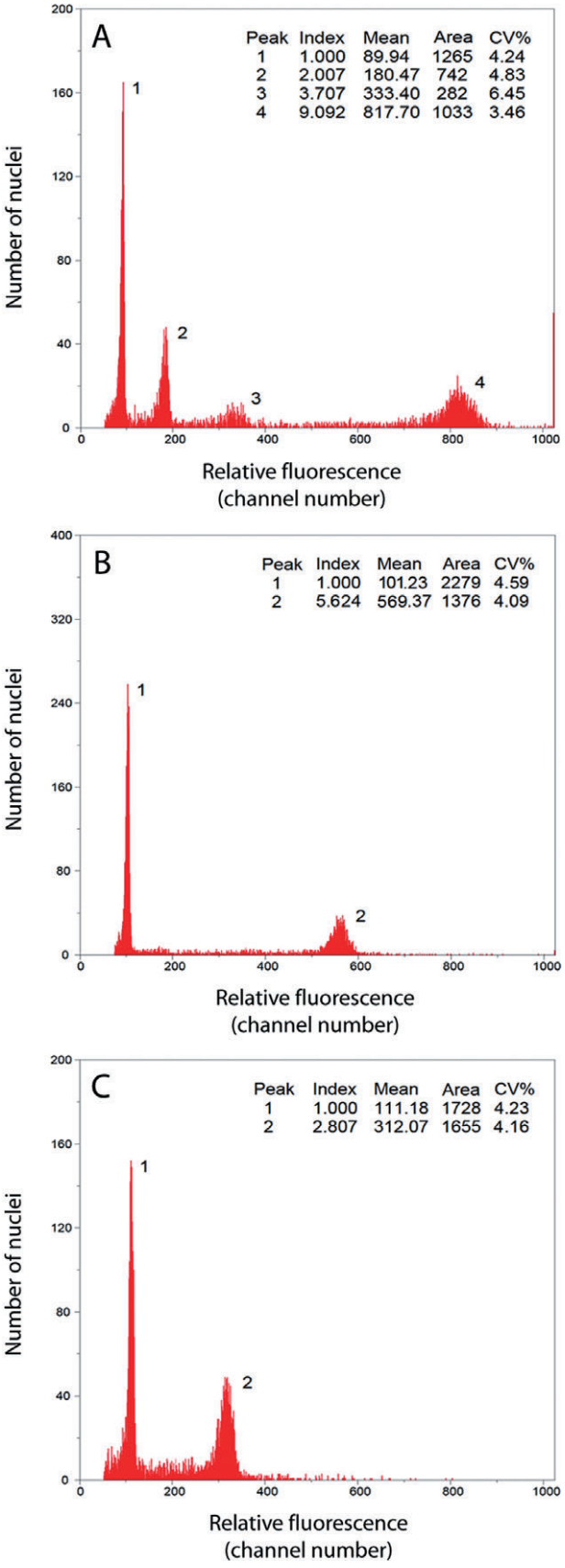

Figure 1. Representative flow histograms of relative fluorescence obtained after isolation of nuclei from *Paris japonica* with three internal calibration standards. A, *Allium cepa* (peaks 1, 2 and 3 = *A. cepa* 2C, 4C and 8C; peak 4 = *P. japonica* 2C). B, *Trillium rivale* (peak 1 = *T. rivale* 2C, peak 2 = *P. japonica* 2C). C, *Trillium sessile* (peak 1 = *T. sessile* 2C, peak 2 = *P. japonica* 2C). Statistical analysis of peaks is also given: Index = mean channel number of sample/mean channel number of standard; Mean = mean channel number; Area = number of nuclei in peak; CV% = coefficient of variation of peak.

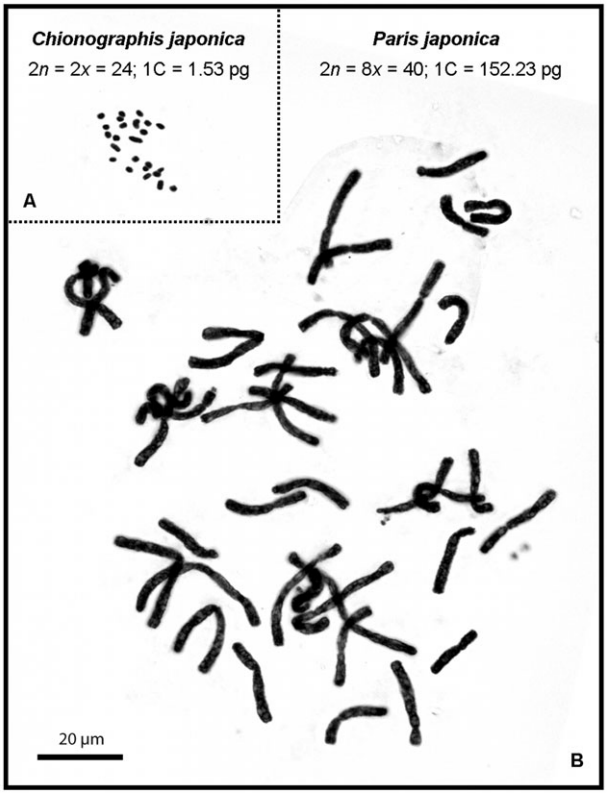

Figure 2. Comparison of the 100-fold difference in genome sizes between (A) *Chionographis japonica* (1C = 1.53 pg) and (B) *Paris japonica*, the species with the largest known eukaryotic genome. Both pictures are at the same magnification.

japonica studied were shown to be octoploid, with $2n = 8x = 40$ chromosomes (Fig. 2B).

DISCUSSION

Previous studies have shown that chromosome numbers in tribe Parideae are based on $x = 5$ (Tamura, 1995, and references therein), with polyploid series up to $8x$ in *Paris* (*P. japonica*, labelled as *Kinugasa japonica*; Haga, 1937; Hara, 1969), and our

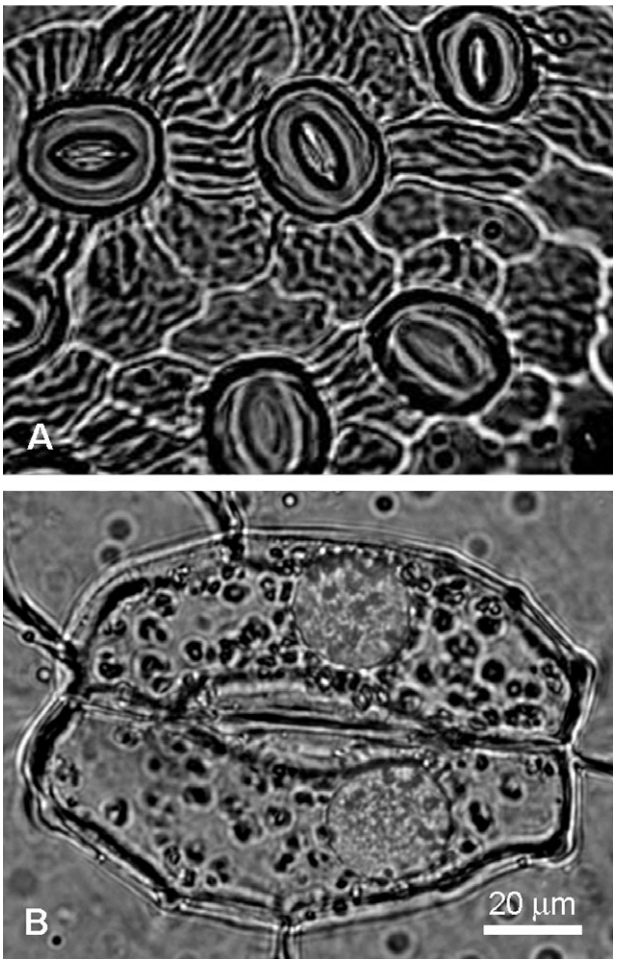

Figure 3. Comparison of abaxial leaf epidermal images showing guard cells. A, *Platanus orientalis* 1C = 1.27 pg (image from Knight & Beaulieu, 2008, and reused here with permission of the authors). B, *Paris japonica* 1C = 152.23 pg (image used with permission of P. Franks). Both pictures are at the same magnification.

results confirm these reports. A comparison of the chromosome complement of *P. japonica* with that of another member of Melanthiaceae, *Chionographis japonica* (Willd.) Maxim. (Fig. 2A), which is reported to have the smallest genome size in Melanthiaceae (1C = 1.53 pg, Leitch *et al*., 2010), illustrates the great karyological diversity within the family. There is a > 10-fold difference in the chromosome length between these species (Tamura, 1995) and *c.* 100 × more DNA in *P. japonica*.

Having such a large genome has direct phenotypic consequences (e.g. size and density of stomatal cells, Fig. 3; Beaulieu *et al*., 2008; Knight & Beaulieu, 2008) and the example of *P. japonica* supports the trend of increased guard-cell size linked to genome-size expansion (although this does not necessarily apply across narrow ranges of small genome sizes; Rupp *et al*., 2010).

The genome of *P. japonica* is 15% larger than any previous estimate for a eukaryote and its discovery extends the range of genome sizes encountered in eukaryotes to c. 66 000-fold. But have we uncovered the full extent of genome size diversity? At the lower end of the scale, the minute genome of *E. intestinalis*, which is even smaller than many bacterial genomes, is already extremely compact, with reductions at many levels, including the number and size of genes and the minimal amounts of non-coding repetitive DNA (Vivares & Metenier, 2000). Further extensive reductions of genome size in eukaryotes therefore seem unlikely.

At the upper end of the range, species with 'very large' genomes (i.e. 1C > 100 pg) have evolved independently in only a few eukaryotic lineages. Plant species with large chromosomes, and hence genomes, are known in several monocot families, but these are mostly diploid with only a few tetraploids and hexaploids. As *P. japonica* is the only known octoploid with large chromosomes reported in these well-studied families, it seems unlikely that larger genomes will be uncovered in plants.

For animals, there has been some debate as to whether the reported genome size estimates for *Protopterus aethiopicus* and three other species comprising the genus are over- or underestimates (Gregory, 2005a). The value of 1C = 81.60 pg for the only tetraploid *P. dolloi* ($2n = 64$) is now considered an underestimate and more recent studies suggest its genome is closer to 1C = 125 pg; finding animals with genomes exceeding 150 pg seems unlikely.

In the Special Issue of *Science* celebrating its 125th anniversary, one of the 100 big questions listed under 'What don't we know' (Kennedy & Norman, 2005) was 'Why are some genomes really big and others quite compact?'. This is an important question, as many studies have documented how there are clear biochemical costs and biological consequences associated with increasing DNA amounts (Gregory, 2001; Leitch & Bennett, 2007), such as reduced brain complexity in some animals (Roth, Blanke & Wake, 1994; Andrews & Gregory, 2009) and increased risk of extinction in plants (Vinogradov, 2003). Although the recent surge in molecular data has already contributed much to our understanding of the DNA sequences that comprise genomes of different sizes and the mechanisms that bring about these changes (Grover & Wendel, 2010), such studies have focused almost exclusively on species with small to medium-sized genomes.

We are still profoundly ignorant about why some genomes, including that of *P. japonica*, are so big and how they operate and function. In searching for the

answers, it seems likely that next generation and third generation sequencing technologies will deliver a wealth of genomic and epigenomic data from which insights will be gained. It is imperative that such approaches are extended to truly obese genomes if we are to get a holistic view of genome size diversity across eukaryotes.

ACKNOWLEDGEMENTS

We wish to thank Richard Kernick for growing the material of *Paris japonica* and the Japanese Alpine Rock Garden Society for supplying it. Peter Franks and Charlie Knight kindly allowed us to use images. This research was supported by research grant NE/G020256/1 from NERC.

REFERENCES

Andrews CB, Gregory TR. 2009. Genome size is inversely correlated with relative brain size in parrots and cockatoos. *Genome* **52:** 261–267.

APG II. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants. *Botanical Journal of the Linnean Society* **141:** 399–436.

APG III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* **161:** 105–121.

Beaulieu JM, Leitch IJ, Patel S, Pendharkar A, Knight CA. 2008. Genome size is a strong predictor of cell size and stomatal density in angiosperms. *New Phytologist* **179:** 975–986.

Bennett MD, Leitch IJ. 2005. Genome size evolution in plants. In: Gregory TR, ed. *The evolution of the genome*. San Diego: Elsevier, 90–151.

Bennett MD, Leitch IJ, Price HJ, Johnston JS. 2003. Comparisons with *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.

Bennett MD, Smith JB. 1976. Nuclear DNA amounts in angiosperms. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **274:** 227–274.

Chase MW, Christenhusz MJM, Sanders D, Fay MF. 2009. Murderous plants: Victorian Gothic, Darwin and modern insights into vegetable carnivory. *Botanical Journal of the Linnean Society* **161:** 329–356.

Chase MW, Duvall MR, Hillis HG, Conran JG, Cox AV, Eguiarte LE, Hartwell J, Fay MF, Caddick LR, Cameron KM, Hoot SB. 1995. Molecular phylogenetics of Lilianae. In: Rudall PJ, Cribb PJ, Cutler DF, Humphries C, eds. *Monocotyledons: systematics and evolution*. Kew: Royal Botanic Gardens, 109–137.

Chase MW, Soltis DE, Soltis PS, Rudall PJ, Fay MF, Hahn WH, Sullivan S, Joseph J, Molvray M, Kores P, Givnish TJ, Sytsma KJ, Pires JC. 2000. Higher-level systematics of the monocotyledons: an assessment of current knowledge and a new classification. In: Wilson KL, Morrison DA, eds. *Monocots: systematics and evolution*. Melbourne: CSIRO, 3–16.

Cui L, Wall PK, Leebens-Mack JH, Lindsay BG, Soltis DE, Doyle JJ, Soltis PS, Carlson JE, Arumuganathan K, Barakat A, Albert VA, Ma H, dePamphilis CW. 2006. Widespread genome duplications throughout the history of flowering plants. *Genome Research* **16:** 738–749.

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

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

Fuse S, Tamura MN. 2000. A phylogenetic analysis of the plastid *mat*K gene with emphasis on Melanthiaceae *sensu lato*. *Plant Biology* **2:** 415–427.

Gregory TR. 2001. Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. *Biological Reviews* **76:** 65–101.

Gregory TR. 2005a. Genome size evolution in animals. In: Gregory TR, ed. *The evolution of the genome*. San Diego: Elsevier, 4–71.

Gregory TR. 2005b. Synergy between sequence and size in large-scale genomics. *Nature Reviews Genetics* **6:** 699–708.

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.

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

Grover C, Wendel JF. 2010. Recent insights into mechanisms of genome size change in plants. *Journal of Botany* **2010:** Article ID 382732 (8 pages).

Haga T. 1937. Chromosome complement of *Kinugasa japonica* with special reference to its origin and hebavior. *Cytologia* **8:** 137–141.

Hara H. 1969. Variations in *Paris polyphylla* Smith, with reference to other Asiatic species. *Journal of the Faculty of Sciences, University of Tokyo, Section III* **10:** 141–180.

Kennedy D, Norman C. 2005. What don't we know? *Science* **309:** 75.

Knight CA, Beaulieu JM. 2008. Genome size scaling through phenotype space. *Annals of Botany* **101:** 759–766.

Knight CA, Molinari NA, Petrov DA. 2005. The large genome constraint hypothesis: evolution, ecology and phenotype. *Annals of Botany* **95:** 177–190.

Leitch IJ, Beaulieu JM, Chase MW, Leitch AR, Fay MF. 2010. Genome size dynamics and evolution in monocots. *Journal of Botany* **2010:** Article ID 862516 (18 pages).

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

Leitch IJ, Bennett MD. 2007. Genome size and its uses: the

impact of flow cytometry. In: Doležel J, Greilhuber J, Suda J, eds. *Flow cytometry with plant cells*. Weinheim: Wiley-VCH, 153–174.

Lisch D. 2009. Epigenetic regulation of transposable elements in plants. *Annual Review of Plant Biology* **60:** 43–66.

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

Parisod C, Alix K, Just J, Petit M, Sarilar V, Mhiri C, Ainouche M, Chalhoub B, Grandbastien MA. 2009. Impact of transposable elements on the organization and function of allopolyploid genomes. *New Phytologist* **186:** 37–45.

Pedersen RA. 1971. DNA content, ribosomal gene multiplicity, and cell size in fish. *Journal of Experimental Zoology* **177:** 65–79.

Roth G, Blanke J, Wake DB. 1994. Cell-size predicts morphological complexity in the brains of frogs and salamanders. *Proceedings of the National Academy of Sciences of the United States of America* **91:** 4796–4800.

Rudall PJ, Stobart KL, Hong W-P, Conran JG, Furness CA, Kite GC, Chase MW. 2000. Consider the lilies: systematics of Liliales. In: Wilson KL, Morrison DA, eds. *Monocots: systematics and evolution*. Melbourne, Australia: CSIRO, 347–359.

Rupp B, Samuel R, Russell A, Temsch EM, Chase MW, Leitch IJ. 2010. Genome size in *Polystachya* (Orchidaceae) and its relationships to epidermal characters. *Botanical Journal of the Linnean Society* **163:** 223–233.

Soltis PS, Soltis DE. 2000. The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences of the United States of America* **97:** 7051–7057.

Takhtajan A. 1983. A revision of *Daiswa* (*Trilliaceae*). *Brittonia* **35:** 255–270.

Takhtajan A. 1997. *Diversity and classification of flowering plants*. New York: Columbia University Press.

Tamura MN. 1995. A karyological review of the orders Asparagales and Liliales (Monocotyledonae). *Feddes Repertorium* **106:** 83–111.

Thomas CA. 1971. The genetic organization of chromosomes. *Annual Review of Genetics* **5:** 237–256.

Thorne RF. 1992. Classification and geography of the flowering plants. *The Botanical Review* **58:** 225–348.

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

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

Vivares CP. 1999. On the genome of *Microsporidia*. *Journal of Eukaryotic Microbiology* **46** (Suppl.): 16A.

Vivares CP, Metenier G. 2000. Towards the minimal eukaryotic parasitic genome. *Current Opinion in Microbiology* **3:** 463–467.

Wendel JF. 2000. Genome evolution in polyploids. *Plant Molecular Biology* **42:** 225–249.

Zomlefer WB, Judd WS, Whitten WM, Williams NH. 2006. A synopsis of Melanthiaceae (Liliales) with focus on character evolution in tribe Melanthieae. *Aliso* **22:** 566–578.

Zomlefer WB, Williams NH, Whitten WM, Judd WS. 2001. Generic circumscription and relationships in the tribe Melanthieae (Liliales, Melanthiaceae), with emphasis on *Zigadenus*: evidence from ITS and *trnL-F* sequence data. *American Journal of Botany* **88:** 1657–1669.

Zonneveld BJM. 2010. New record holders for maximum genome size in eudicots and monocots. *Journal of Botany* **2010:** Article ID 527357 (4 pages).