Received: 12 May 2021 | Revised: 10 August 2021 | Accepted: 26 August 2021

DOI: 10.1002/cyto.a.24499

REVIEW ARTICLE

ISAC CYTOMETRY Journal of Quantitative Cell Science PART A

Application-based guidelines for best practices in plant flow cytometry

Elwira Sliwinska[1] | João Loureiro[2] | Ilia J. Leitch[3] | Petr Šmarda[4] | Jillian Bainard[5] | Petr Bureš[4] | Zuzana Chumová[6,7] | Lucie Horová[4] | Petr Koutecký[8] | Magdalena Lučanová[6,8] | Pavel Trávníček[6] | David W. Galbraith[9,10]

[1]Laboratory of Molecular Biology and Cytometry, Department of Agricultural Biotechnology, UTP University of Science and Technology, Bydgoszcz, Poland

[2]Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Coimbra, Portugal

[3]Kew Science Directorate, Royal Botanic Gardens, Kew, Richmond, Surrey, UK

[4]Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic

[5]Swift Current Research and Development Centre, Agriculture and Agri-Food Canada, Swift Current, Saskatchewan, Canada

[6]Czech Academy of Sciences, Institute of Botany, Průhonice, Czech Republic

[7]Department of Botany, Faculty of Science, Charles University, Prague, Czech Republic

[8]Department of Botany, Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

[9]School of Plant Sciences, BIO5 Institute, Arizona Cancer Center, Department of Biomedical Engineering, University of Arizona, Tucson, Arizona, USA

[10]Henan University, School of Life Sciences, State Key Laboratory of Crop Stress Adaptation and Improvement, State Key Laboratory of Cotton Biology, Key Laboratory of Plant Stress Biology, Kaifeng, China

Correspondence
Elwira Sliwinska, Laboratory of Molecular Biology and Cytometry, Department of Agricultural Biotechnology, UTP University of Science and Technology, Kaliskiego Ave. 7, 85-796 Bydgoszcz, Poland.
Email: elwira@utp.edu.pl

Funding information
111 Project, Grant/Award Number: D16014; Akademie Věd České Republiky, Grant/Award Number: Research Development Project RVO 67985939; Centro de Ecologia Funcional, Grant/Award Number: FCT/UIDB/04004/2020; Chinese National Key Research and Development Program, Grant/Award Number: 2OI6YFDOl0l006; Grantová Agentura České Republiky, Grant/Award Number: GA19-13231S, GA19-18545S GA20-15989S GA17-18080S; National Natural Science Foundation of China, Grant/Award Number: 31770300; Program for Innovative Research Teams in Science and Technology at a University of Henan Province, Grant/Award Number: 18IRTSTHNO23; Programa Operacional Regional do Centro 2014-2020, Grant/Award Number: CENTRO-01-0145-FEDER-000020

Abstract

Flow cytometry (FCM) is currently the most widely-used method to establish nuclear DNA content in plants. Since simple, 1-3-parameter, flow cytometers, which are sufficient for most plant applications, are commercially available at a reasonable price, the number of laboratories equipped with these instruments, and consequently new FCM users, has greatly increased over the last decade. This paper meets an urgent need for comprehensive recommendations for best practices in FCM for different plant science applications. We discuss advantages and limitations of establishing plant ploidy, genome size, DNA base composition, cell cycle activity, and level of endoreduplication. Applications of such measurements in plant systematics, ecology, molecular biology research, reproduction biology, tissue cultures, plant breeding, and seed sciences are described. Advice is included on how to obtain accurate and reliable results, as well as how to manage troubleshooting that may occur during sample preparation, cytometric measurements, and data handling. Each section is followed by best practice recommendations; tips as to what specific information should be provided in FCM papers are also provided.

KEYWORDS
cell cycle, DNA base composition, DNA content, endoreduplication, flow cytometric seed screening, genome size, in vitro cultures, intraspecific variation, ploidy

1 | INTRODUCTION

Most of the applications of flow cytometry (FCM) in plant science are based on estimating the amount of DNA in cell nuclei. Depending on the research question, the studies focus on the measurements of genome size per se, or on analyzing genome-size derived parameters (such as ploidy level) or comparing the genome sizes of different cell populations (e.g., for estimating levels of endopolyploidy and stages of the cell cycle, or for flow cytometric seed screening). While the basic principles of the measurements and instrumentation are similar, individual applications differ in their focal parameters (e.g., genome size estimation with high precision, estimating the number of nuclei in individual cell populations, or fast screening of many individuals with lower precision). Here, we discuss the typical FCM applications for plants, and outline the methodological considerations and best practice recommendations specific for each application. In Sections 2 to 5, we describe the most widely used applications in detail (i.e., estimation of ploidy, genome size, and DNA base composition). In Sections 5 to 9, we outline how the methods can be applied for more specialized analyses. These are focused on either (i) how the methods can be used to quantify cells with different ploidy levels within a sample to assess levels of endopolyploidy or for cell cycle analysis, or (ii) specific methods needed to analyze certain types of plant material (i.e., seeds, in vitro-derived material). For these more specialized analyses, we focus mainly on explaining the underlying biological rationale of the approaches. This is because the technical details are broadly similar to those outlined in Sections 2 to 5, thus only specific requirements are highlighted where appropriate.

The majority of the text refers to the analysis of plants with a dominant sporophytic phase of the life cycle (i.e., angiosperms, gymnosperms, ferns and fern allies, and lycophytes) as they comprise the vast majority of land plants. However, the recommendations are also generally valid for analyzing material from the gametophytic life cycle phases (e.g., bryophytes, autonomous gametophytes of lycopods and ferns, and pollen), and there is no obstacle in using sporophytic genome size standards to analyze gametophytic samples. Researchers working with plants in the gametophytic phase of the life cycle just need to be aware that it is necessary to adjust some of the calculations accordingly (e.g., when calculating the genome size of the analyzed material and especially in estimating its monoploid genome size, 1Cx).

2 | PLOIDY DETERMINATION

One of the most popular applications of FCM in plant sciences over the last few decades has been the estimation of ploidy level, that is, detection of copy number of sets of homologous chromosomes. Ploidy determination is central to numerous plant breeding applications, including the breeding of polyploid species/cultivars, (doubled) haploid production, somatic hybridization by protoplast fusion, and detection of somaclonal variation in tissue cultures.[1] In addition, establishment of the ploidy levels of the embryo and endosperm in seeds provides information on the reproductive pathway of the plant (see Section 8). In biosystematics research, ploidy assessment is used for cytotype screening (categorizing plants of the same species that differ in chromosome number/ploidy),[2–4] and the detection of intercytotype hybrid zones,[5,6] intercytotype gene flow,[7,8] and minority cytotypes.[9,10] Due to the presence of breeding barriers, individual ploidy levels within a polyploid complex are frequently classified as autonomous taxa, and ploidy level information provides an independent taxonomic characteristic that allows searching for differences in morphology, distribution, or ecology.[11,12] Phylogenetic relationships and pathways of polyploid origin can be inferred based on ploidy.[13] Knowledge of plant ploidy levels is of ecological interest as well, as it may play a role in plant invasiveness,[14–16] competitive ability,[17,18] and responses to stress.[19,20] These are the most common examples, but there exists a virtually unlimited application space within which the ploidy level plays a significant role.

FCM has almost completely replaced traditional methods for establishing ploidy, whether direct (chromosome counting, microdensitometry) or indirect (the diameter of pollen grains or spores, the length of stomatal guard cells, or number of chloroplasts in stomata). Nevertheless, since FCM is an indirect method for ploidy estimation, it has to be supported by chromosome counting (see below and reference 21 for example). Advantages of FCM over other methods of ploidy estimation include (i) no need for special preparation of plant material (most tissues of the growing plant can be used),[22] (ii) rapid sample preparation, allowing daily processing of dozens or even hundreds of samples, (iii) an ability to measure DNA content/ploidy in mitotically inactive cells (as compared to chromosome counting, which requires mitotic cells), (iv) nondestructive sampling (a very small sample, about 5–10 mg fresh weight, is sufficient to provide thousands of nuclei for ploidy determination), which enables investigation of rare and endangered species with minimal damage, and allows the same individual to be used for further analyses, (v) the possibility of analysis of bulked samples of pooled individuals at one time, thereby providing large-scale screening, and (vi) very modest operating costs. It must be realized, however, that it is only possible to estimate ploidy levels from FCM measurements within a certain taxonomical rank (mostly within closely related species). A chromosome count is still recommended to calibrate ploidy established by FCM, although for some special cases, such as in agmatoploid or symploid species (when chromosome number changes are due to chromosome fragmentation or fusion), a genome size measured using FCM is more relevant for estimating the actual ploidy level than chromosome counts.[23] When a ploidy estimate is not verified by chromosome counts, it should be referred to as 'DNA-ploidy'.[24]

The nuclear DNA content of G_0/G_1 nuclei reflects the ploidy of a plant. Depending on how much prior information is available about the species under study, there are various approaches to determine ploidy of an unknown individual. If an individual of known ploidy is available (e.g., as determined from chromosome counts), the DNA content of another individual plant of the same species of unknown ploidy can be directly compared to the DNA content of this plant

using FCM.[25–27] This comparison is done either by using internal standardization or by using external standardization.[28,29] Both methods of standardization are outlined below. For ploidy estimation, any fluorochrome binding specifically and stoichiometrically to DNA, base-specific or base-nonspecific, can be applied (for fluorochromes see Section 5).

2.1 | Standardization and determination of ploidy levels

For almost all plant samples, internal standardization is strongly recommended. As an internal standard, a species unrelated to the study material that has a genome size close to, but not overlapping any of the studied cytotypes should be used (e.g., one of the recommended standard species).[28] If an unrelated species is used as a standard, first the DNA index (ratio) should be calculated based on the measured mean fluorescence intensities (*MFI*) (for example histogram see Figure 1 in Section 3):

$$DNA\,index = \frac{sample\,G_0/G_1\,MFI}{standard\,G_0/G_1\,MFI}.$$

This ratio is an estimate of the relative genome size of the sample, where the relative fluorescence of the standard is given in arbitrary units. It is possible to use this equation with any fluorochrome, but the same fluorochrome should be used throughout the experiment.

The ploidy level of the studied sample is then inferred by comparison of its DNA index to the DNA index of an individual of the same species with known ploidy level (e.g., established by chromosome counting) according to the equation:

$$sample\,ploidy\,level = \left(\frac{sample\,DNA\,index}{DNA\,index\,of\,sample\,with\,known\,ploidy}\right) \times known\,ploidy\,level.$$

Both the calculation of the peak index and of the sample ploidy level are independent of whether the tissues analyzed are gametophytic or sporophytic, as long as the same tissue type is used for the individual used for calibration and for the unknown sample.

It is also possible to use an individual of the studied species with known ploidy level as an internal standard, but this approach has one major drawback: if samples are taken from plants having the same ploidy level as the standard, only one nuclear peak will be observed, and it is impossible to determine whether sample nuclei are even present.

External standardization can also be used for ploidy estimation; however, it should be limited to screening of plant material of a known genetic origin such as breeding material, material that has been produced in vitro (e.g., haploids or induced polyploids), or for population ploidy screening. If external standardization is used, a plant of the same species with known ploidy serves as a standard. It should be run independently from the samples at the beginning of the ploidy analyses to calibrate the flow cytometer (set appropriate gain or voltage) and then every 10–20 samples to correct the settings, or more often if fluctuation of the sample peak position is observed (e.g., due to warming of the UV lamp). Problematic issues arising from poor instrument stability have declined considerably with the availability of more recent generations of flow cytometers.[26]

If a standard of the same species is used as an external reference, the ploidy of the sample is calculated according to the equation:

$$sample\,ploidy\,level = \left(\frac{sample\,G_0/G_1\,MFI}{standard\,G_0/G_1\,MFI}\right) \times standard\,ploidy\,level.$$

2.2 | Bulk samples

To facilitate rapid detection of ploidy across hundreds of samples, while still enabling detection of individual differences, bulk samples that include tissue from 2–10 individuals can be processed in a single FCM run.[8,30,31] Such an approach is possible only if: (i) endopolyploidy and/or a pronounced G_2 phase are absent in the species under study, (ii) single outliers from the majority of the fluorescence signal(s) are clearly identified, and (iii) the plant material provides no problems in terms of background fluorescence, and the nuclei provide precise and accurate fluorescence signals. The number of individuals that can be accommodated in a bulked sample depends on the ability to detect the ploidy level of all individuals (high quality FCM measurements are essential) and on the variability in the population under investigation. Use of bulked samples from preserved (e.g., silica-dried) materials is not recommended, because tissue preservation may be accompanied by a lower sample peak quality, shifted peak positions, and greater amounts of background signals.[32] The use of the same species as an internal standard is also not recommended for bulk analyses, as it is not possible to decide whether some of the nuclei in the standard position belongs to the sample of interest, even though other peaks may be present.

2.3 | Flow cytometer settings

The flow cytometer settings should be adjusted to allow detection of peaks that have a fluorescence intensity two or more times higher than the highest known ploidy of the species under study and also two times lower than the lowest known ploidy, unless the material has been reliably confirmed to be diploid by karyological methods and haploids are not expected. This step may require repeated adjustments to the cytometer's gain (voltage) settings but will ensure that unexpected ploidy levels are detected. If endopolyploid nuclei are present, extra care must be taken to ensure that the lowest ploidy level is detected, otherwise the ploidy assignment (at the level of the individual plant/taxa) will be incorrect.

To calculate the mean peak position with appropriate precision, the measurement of fluorescence of ≥1000 nuclei for a G_0/G_1 peak appears sufficient when external standardization is used.[33] If a histogram

contains more than two G_0/G_1 peaks (due to internal standardization, see above) or additional peaks due to endoreduplication (see Section 7), the total number of nuclei analyzed should be increased so that the fluorescence of ≥1000 nuclei are measured for each G_0/G_1 peak. Similarly, when using bulked samples, more nuclei should be analyzed, typically around 3000–5000, to ensure that the presence of a single individual with a different ploidy level is not missed in the analysis. Higher nuclei counts are also needed when searching for aneuploidy (gain or loss of one or more chromosomes at a given ploidy level).

2.4 | Assignment of ploidy, potential pitfalls, and how to avoid them

Classifying ploidy by FCM is a relatively easy task. However, to avoid potential errors, the following steps should be followed:

2.4.1 | Identifying a cluster of results sufficiently similar to be regarded as one cytotype

Distinct clusters of results corresponding to different ploidy levels are identified based on the ratio of fluorescence intensities of the chosen standard and samples under study. Taxa without intracytotype variation in genome size (see Section 4) should provide clear clusters (peaks) without overlap, and the mean relative fluorescence should correspond to the expected ratios of the euploidy levels. If some outliers appear, aneuploidy and other sources of genome size variation should be considered.

2.4.2 | Associating each peak/cluster with the ploidy level

To determine the ploidy level, there are several options depending on the amount of prior knowledge available for the species of interest. The easiest situation exists when at least one of the measured individuals has a preassigned ploidy level based on chromosome counting. Once the relative fluorescence of the sample of known ploidy is determined (using a standard), the fluorescence profile of a sample of unknown ploidy can then be compared to this value. To associate other samples with different ploidy levels, a simple rule of linear dependency applies – a peak with twice the relative fluorescence indicates an individual with double the known ploidy level (but see comments on genome downsizing and genome size variability below).

If the nature of the study does not allow the independent assignment of ploidy levels, there are several ways to handle this. After careful review of the literature available for the species under study, preferably sampled from the same geographic region, the peak having the lowest relative fluorescence can be assigned the lowest ploidy level known for the species. Further assignments of ploidy levels must always fit the mean ratios of fluorescence intensities observed between peaks, however, this is also subject to interpretation. For example, if the ratios between the mean fluorescence of three peaks are 1: 1.5: 2, these peaks could either be assigned to ploidy levels of 2*x*, 3*x*, and 4*x*, or 4*x*, 6*x*, and 8*x*. Their final assignment depends on the convention and expectation for the species under study and its close relatives. If no data on the species are available, then the lowest fluorescence intensity can be considered the diploid level. When ploidy levels are assigned in this way (i.e., indirectly), they should be termed DNA-ploidy, as mentioned above, to unequivocally express the way ploidy assignment was done, and to distinguish them from cytogenetically inferred ploidy levels.

It is also important to note that the process of genome downsizing following polyploid formation may result in DNA contents that are smaller than expected (i.e., monoploid genome size, 1Cx-value, tends to decrease with increasing ploidy). In other words, DNA amounts at higher ploidies are often less than full multiples of the 1Cx-value of diploids and the effect can become more pronounced with increasing ploidy levels (i.e., the 1Cx-value decrease may be nonlinear).[34] Similarly, when studying a polyploid complex of several species, each may have a specific 1Cx-value and the differences may be huge.[13,35] In such cases, estimation of ploidy from genome size and/or the ratio to a given species is not straightforward. These issues require careful interpretation of the data before assigning DNA-ploidy.[21,24,36] A chromosome count may be required to unequivocally assign a ploidy level to a plant sample, especially for a newly detected cytotype, but this ploidy assignment may not be possible in symploid and agmatoploid species (e.g., Cyperaceae, Juncaceae) where chromosome numbers can vary among and within species of the sample ploidy due to chromosomal fusion and fragmentation.

2.4.3 | Identifying outliers as potential aneuploids

Another tricky issue in ploidy determination is coping with outliers from clusters and their assignment to potential aneuploids. Firstly, various technical artifacts must be excluded (see above). The existence of outliers is likely due to intracytotype genome size variation (see Section 4), supernumerary B-chromosomes, or aneuploidy. Some of these options can be ruled out based on the type of detected variability: continuous variability likely points to intracytotype genome size variation, or the presence of B-chromosomes,[37] while discrete variability with distinct steps suggests aneuploidy.[8] However, one should be aware of the limitations resulting from the technical resolution of FCM estimation. Detection and assignment of aneuploidy can be completely different based on the size and number of chromosomes in the plant under study. For example, assuming similar chromosome sizes, if the euploid has $2n = 10$ chromosomes, the presence of one additional chromosome of similar size to the other autosomes will result in an approximately 10% difference in DNA content, whereas if $2n > 50$, one additional chromosome will result in a less than 2% difference. Thus, the justification to consider cluster outliers as aneuploids requires corroboration via chromosome counting. Likewise, in aneuploid-prone material (e.g., when studying hybridization between plants of different ploidy), researchers should be aware that

they may easily overlook aneuploids (scoring them as euploids). In such cases, internal standardization and genome size-quality analyses are required (see Section 4). If aneuploids are present in the analyzed population, the use of bulked samples becomes limited, as a minor peak of one aneuploid individual may be hidden within the major peak of an euploid, depending on the genome size difference, the quality of a histogram, and the number of individuals in the bulked samples.

2.5 | Best practices

- For ploidy estimation, the employment of an individual cytotype/line/cultivar/clone of which the ploidy has been verified by chromosome counting is recommended as a control. If ploidy is not verified by chromosome counting, it should be called DNA-ploidy.
- For nuclei isolation, adopt recommendations provided by Loureiro et al.[38]
- Internal standardization is recommended; however, external standardization can be used for breeding/in vitro-produced materials and for population ploidy screening.
- Beware of the presence of aneuploids, especially when bulk samples are analyzed.
- When assigning the ploidy level to different cytotypes, consider possible variability in the monoploid genome size (e.g., due to genome downsizing).

3 | GENOME SIZE ESTIMATION

The amount of DNA in the nucleus of a cell is typically referred to as the genome size, and is measured either in terms of picograms (pg, i.e., 1×10^{-12} g) or base pairs (with 1 pg = 978 Mbp).[39] Nevertheless, given that the amount of DNA varies across the cell cycle (i.e., G_2 nuclei have twice the DNA amount as G_1 nuclei; see Section 6), and following meiosis or endoreduplication (see Section 7), considerable confusion can arise when discussing genome sizes.[40] Here the widely-adopted terminology proposed by Greilhuber et al.[41] and Greilhuber and Doležel[42] is employed. Thus, the holoploid genome size (or C-value) refers to the total amount of DNA in the nucleus regardless of generative ploidy. In diplophasic organisms (most of the vascular land plants), the 1C-value corresponds to the amount of DNA in the unreplicated gametic nucleus (e.g., sperm or egg cell of angiosperms), while the 2C-value is the amount of DNA in a somatic cell at the G_0/G_1 stage of the cell cycle (see Section 6). In contrast, the monoploid genome size (or Cx-value) corresponds to the DNA amount in the unreplicated basic (x) chromosome set. It can be calculated by dividing the 2C-value by the ploidy level (e.g., in a tetraploid where $2n = 4x$, the 1Cx-value = 2C-value ÷ 4). For diploids (i.e., $2n = 2x$) the 1C-value (holoploid) and 1Cx (monoploid) genome sizes are the same; however, in polyploids, which contain more than two monoploid genomes in a somatic cell, the 1Cx-value is always smaller than the 1C-value. Calculating an 'average' size of the monoploid genome in polyploids, as expressed by the 1Cx-value, can be helpful for phylogeny-dependent statistical analyses of genome size evolution, although it should be realized that this measure can be a bit misleading in some polyploid species where the different genomes within the polyploid nucleus differ considerably in size.

As the use of FCM to estimate genome sizes has risen over the years, it has greatly increased the rate at which new genome size data are being generated.[43] This has enhanced our understanding not only of the huge diversity of genome sizes encountered across different plant lineages,[44] but also uncovered considerable diversity between closely related species and even within species, both in terms of differences in ploidy level (see Section 2) and the existence of inter- and intraspecific genome size variation (see Section 4), independent of polyploidy.[13,21,35,45] In some cases, genome size data can even be used to distinguish between related species having the same chromosome number.[45] Moreover, the size of the genome has been shown to correlate with a huge diversity of genomic, phenotypic and ecological traits, demonstrating the importance of documenting and understanding genome size diversity and how it can impact not only the ecology and evolutionary trajectory of a species but also the response to environmental changes.[46–49]

Knowledge of genome sizes is also essential for more practical reasons, for example, in planning whole genome sequencing where it is necessary to know how big the genome is in order to estimate time and costs.[50] In addition, for research involving tissue culture, knowledge of genome size is vital to assess somaclonal variation (see Section 9).

Given the importance of understanding the full extent of plant genome size diversity outlined above, and the acceleration in the rate at which plant genome size data have accumulated, this has led to plants being the most thoroughly explored eukaryote lineage, and the discovery that angiosperms have the largest range of genome sizes for any comparable group (ranging c. 2400-fold).[43,44] The majority of published genome size data have been collated into the publicly available Kew Plant DNA C-values Database (https://cvalues.science.kew.org/), which currently (release 7.1, April 2019) contains data for 12,273 species comprising 10,770 angiosperms, 421 gymnosperms, 334 bryophytes, and 445 algae.

The ability to use FCM to estimate genome size is based on comparing the fluorescence intensities of nuclei that have been quantitatively stained with a DNA-specific fluorochrome (usually propidium iodide [PI]) between the sample and the internal calibration standard (a plant with a known genome size, Figure 1).[28]

3.1 | Quantification of genome sizes using FCM

3.1.1 | Material for measurements

In principle, genome sizes can be estimated from any part of a plant that contains intact nuclei.[22] Most measurements of genome size are typically made using leaves, but other parts of the plant have been used, including flower stalks, petioles, petals, tree phloem, roots, or

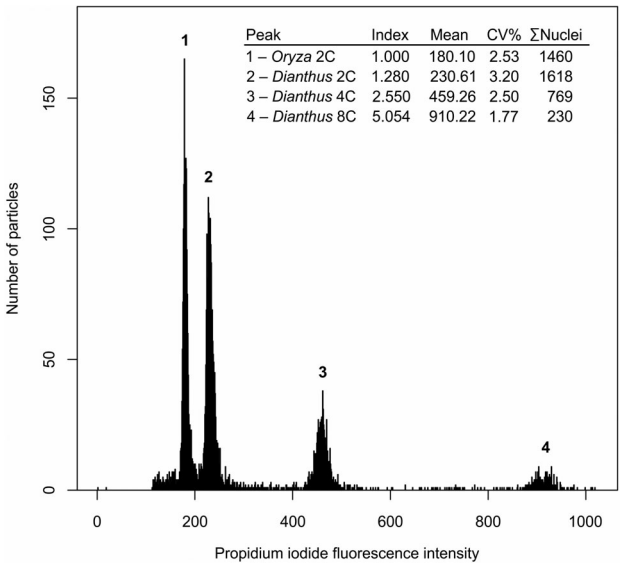

FIGURE 1 Calculating the genome size from the flow histogram. The 1D fluorescence histogram from measurements of 5000 PI-stained particles (nuclei) of *Dianthus capillifrons* (sample) and *Oryza sativa* 'Nipponbare' (internal standard). Debris particles with fluorescence below channel 120 were removed from the histogram by the software. Given the sample/standard ratio of the 2C peaks (i.e., mean of peak 2 divided by mean of peak 1) is 1.280 and the genome size of the standard is 2C = 0.795 pg (777.64 Mbp), the 2C genome size estimate of the sample from this particular measurement is 1.280 × 0.795 = 1.018 pg (995.4 Mbp)

seeds. For some species, preserved or fixed samples (e.g., herbarium vouchers, silica-gel-dried or deep-frozen samples) can provide an approximate estimation of genome size.[51] However, since the fluorescence of nuclei from these nonvital, preserved tissues is typically lower than from the corresponding fresh tissues,[32,52,53] the use of preserved or fixed samples is therefore not recommended for robust genome size measurements in plants. Also using seeds for precise genome size estimation should be avoided.[22]

3.1.2 | Selection of the internal standard

An essential requirement for measuring genome size is a genome size calibration internal standard, which enables the relative fluorescence signal of the sample nuclei to be converted into an absolute genome size measurement with units (e.g., pg, Mbp). For plants, these standards are plant nuclei of known genome size.[28] The use of calibration beads is not recommended, given their spectral characteristics are quite different from stained plant nuclei. Likewise, use of nucleated blood cells (cf. chicken or fish red blood cells) is not recommended due to potential differences in the staining properties of these versus plant nuclei.

To ensure similar staining and fluorescence conditions for nuclei from both the sample and the standard species, they should be processed together (by co-chopping tissues of both species in the same dish in a nuclei isolation buffer), and the mixed suspension of sample and standard nuclei then measured in a single run. Internal standardization is particularly important in plants because their cells can contain a wide variety of secondary metabolites that affect the quality of the DNA staining,[38] which leads to errors in genome size estimation. Co-processing the sample and standard material ensures these metabolites or other factors affect both the standard and sample nuclei to the same extent and hence allows the fluorescence ratio to remain constant, even under suboptimal staining conditions.[6,38]

To identify the appropriate internal standard, an approximate genome size of the sample of interest (if not known) can first be quickly estimated using a rapid procedure. For example, it may be estimated by measuring the sample fluorescence at a given gain (voltage) and comparing its peak position with that of common standards measured on the same machine with the same settings. Such initial measurements of just the sample of interest also provide an approximate idea about the quality of the sample peaks and whether there is a need to explore different nuclei isolation buffers or alternative tissues for measurement. However, users should be aware of the potential for secondary metabolites to interfere with staining by noting whether the position of the peak changes when running a single or combined sample. An alternative approach to estimate the approximate genome size is to prepare a combined sample with an internal standard that has a mid-range genome size, for example *Bellis perennis* (2C = 3.159 pg, see Dataset S01 in reference 54) and then run the sample, preferably on a logarithmic scale, knowing the expected position of the standard peak. Nevertheless, if only one peak appears in the FCM histogram, it is difficult to determine whether this is due to the overlap of the sample and standard peaks or due to low or absent DNA signal from the sample.

The peaks of the internal standard selected for the actual genome size measurements should not overlap with any peak of the sample, and the genome size of the standard should not differ more than ±3-times from that of the sample (i.e., the sample/standard fluorescence ratio should be 0.3–0.9 or 1.1–3.0, assuming that peaks differing 0.9- and 1.1-fold in their mean fluorescence do not overlap). Measurements of samples with very different genome sizes can suffer from various nonlinearity effects intrinsic to the cytometer,[55] and in general should be avoided particularly when using older instruments. One approach that has been suggested to overcome such problems is the use of polysomatic standards which have nuclei that have undergone endoreduplication.[56] The peaks of highly endopolyploid nuclei (usually 16C and higher; see Section 7), however, can show decreased fluorescence compared to their expected genome size, when analyzed using some instruments. In addition to instrumental nonlinearity, this might also be a consequence of differences in chromatin state and DNA stainability of highly endopolyploid nuclei, or their possible under-replication.[57] Therefore, the use of highly endoreduplicated nuclei from polysomatic standards should be carefully considered, and the measurements always controlled appropriately by reference to known DNA content values.

3.1.3 | DNA fluorochromes and staining

To measure DNA in absolute units: (i) the applied fluorochrome must bind specifically and quantitatively to DNA, (ii) its binding must be sequence-independent and thus not affected by

differences in base composition between the sample and the standard, and (iii) the fluorochrome binding to DNA must have the same stoichiometry under the test conditions to ensure the measured differences in fluorescence intensity of the stained nuclei (i.e., amount of fluorochrome bound to DNA) between the sample and the standard are directly proportional to the differences in the amounts of DNA in the nuclei.[58] Given these requirements, only intercalating fluorochromes should be used for absolute genome size estimation, whereas the use of 4′-6-diamidine-2-phenylindole (DAPI) and other base-specific fluorochromes for such studies is unacceptable. It is also not appropriate to estimate genome sizes from samples that have been stained simultaneously with multiple fluorochromes (as common in immunological and medical applications of FCM). This is because the presence of more than one fluorochrome can impact the stoichiometry of DNA staining due to competition for DNA binding sites and/or interactions between the different fluorochromes.[59] Although base-specific fluorochromes are sometimes used to express a genome size relative to the standard in arbitrary units (see Section 2), even this use is questionable unless the similarity in DNA base composition between the compared samples can be guaranteed, such as in studies of autopolyploid complexes.[60] Instead, the use of base-specific dyes should be limited to ploidy estimation and the measurement of intraspecific genome size variation (see Sections 2, 4, and 8).

For genome size estimation in absolute units, the most popular and commonly used fluorochrome is PI. This is because it has a higher DNA binding affinity and produces sharper peaks in the flow histogram, compared with, for example, its structural analogue ethidium bromide (EB).[61] The optimal concentration of PI for genome size measurements using leaf homogenates is typically 50 μg/ml. At this concentration, nuclear staining is nearly saturated, with double sized nuclei showing exactly double the fluorescence, and with the peaks having the lowest coefficient of variation (CV; Figure 2). Deviations from this recommended PI concentration and the optimal amount of material used (to be estimated empirically for a given tissue type and species)[22,38] may lead to non-stoichiometric binding and should be avoided. The PI-DNA binding can also be affected by numerous metabolites commonly present in the cytosol of plant cells and that are released during nuclei isolation.[38] Beyond cell metabolites, the concentration of the fluorochrome or nuclei and issues relating to instrumental reasons, the quality of peaks in the flow histograms may be affected also by the chopping technique, incubation time, and age and storage of the chemicals and solutions used.[38,62]

Since PI can bind to both double-stranded RNA and DNA, most PI-staining protocols for estimating genome size routinely include RNase. However, plant homogenates often contain endogenous RNase, which may be sufficient to eliminate RNA contributions to PI fluorescence. Nevertheless, the addition of exogenous RNase is important in the analysis of RNA-rich tissues such as meristems and seeds; in RNA-poor tissues (such as leaves) the use of RNase is frequently unnecessary, but should be confirmed in preliminary experiments.[38]

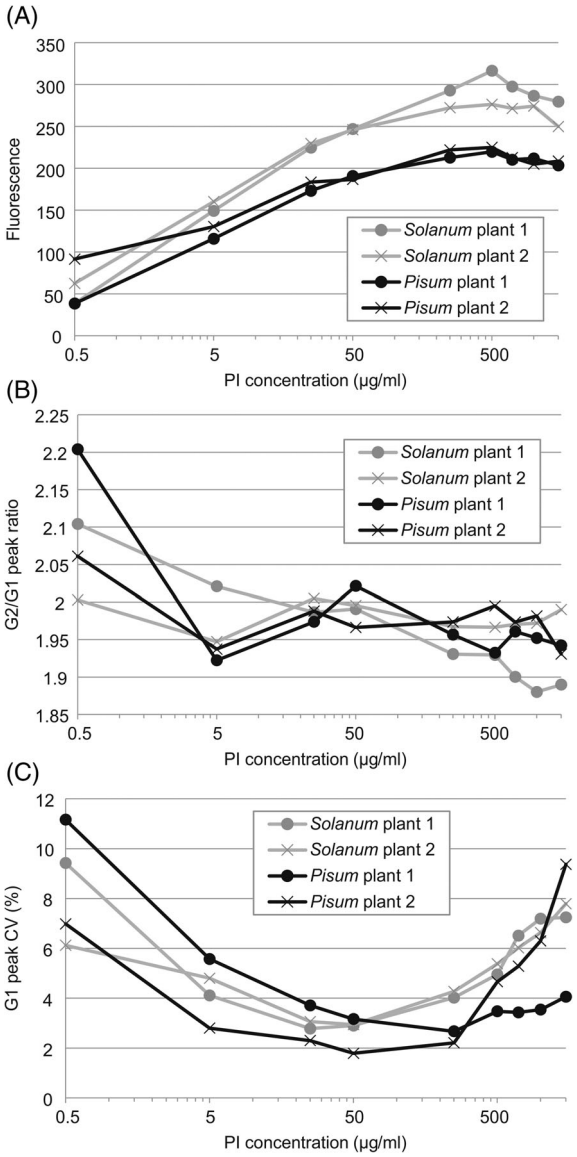

FIGURE 2 Optimizing the concentration of propidium iodide (PI) used to stain plant nuclei. The graphs show how the amount of nuclei fluorescence and the peak CV are influenced by the PI concentration used to analyze two plants of *Solanum lycopersicum* 'Stupické polní tyčkové rané' and two of *Pisum sativum* 'Ctirad', each plant measured on a different day (plant no. 1: 6.3.2019, plant no. 2: 11.3.2019). (A) Relative fluorescence of nuclei at different PI concentrations showing perfect day-to-day reproducibility with PI staining around 50 μg/ml and unstable staining in *Solanum* at high PI concentrations and *Pisum* at low PI concentrations indicated by the difference in the shape of the curves. The PI fluorescence for both plants of the same species was measured at the same gain (voltage), *Pisum* at a much lower gain than *Solanum* so its fluorescence is lower despite its genome size being larger than *Solanum*. (B) Fluorescence ratio between G_2 and G_0/G_1 nuclei at different PI concentrations showing the optimal PI concentration for stoichiometric staining (i.e. where the expected ratio should be two since the G_2 nuclei should have twice the DNA amount as the G_0/G_1 nuclei) at concentrations of PI around 50 μg/ml. Although PI is typically used in this range, it is noted that the optimal concentration may change with the amount of material chopped and the genome size of the analyzed nuclei (i.e. concentration and amount of DNA to be stained). (C) Coefficients of variation (CVs) of G_0/G_1 peaks suggesting that the best quality peaks (i.e., with lowest CVs) are at PI concentrations around 50 μg/ml

3.1.4 | Flow cytometer settings for optimizing the measurement

To optimize the capture of data and hence ensure robust genome size estimations, the following rules should be followed:

- Optimize settings on the flow cytometer to ensure that data from all appropriate and available parameters are gathered. This will depend on the options available for the particular instrument but in addition to fluorescence, if available, this should include side-scatter, forward-scatter, and time. Set the amplification parameters to the appropriate scales, if this is permitted for the cytometer. For older instruments, signal amplification can employ linear or logarithmic amplifiers. In these cases, log amplification of the fluorescent signals is not recommended for measurements of genome size but amplification of side- and forward-scatter (if applicable) can be set to logarithmic scales. For the latest generation of cytometers, signal amplification across all fluorescence and scatter channels is exclusively linear, and whether or not the information is displayed on linear or logarithmic scales is irrelevant since it does not change the linear values of the underlying data. Logarithmic scale displays are encouraged when dealing with measurements having large dynamic ranges, and for situations displaying extensive endoreduplication, such as in the developing maize endosperm.
- Set visual layout allowing the control of all parameters. If possible, it is recommended that 2D plots that is, fluorescence versus side-scatter and fluorescence versus time are displayed.
- Set the threshold for the machine to automatically stop after a desired number of nuclei have been recorded (e.g., 5000 nuclei) or a specified sample volume has been analyzed; stopping the machine manually is also acceptable. For fast screening of multiple samples, analysis of a total of 1000–3000 nuclei should be sufficient. For making the final measurements, it is recommended that the analyzed peak (i.e., G_0/G_1) contains at least 1000 nuclei, so stopping the measurement after a total of 5000 nuclei have been detected is usually enough for most measurements (Figure 1). If additional peaks of nuclei are present (e.g., when measuring genome size in endopolyploid tissues or in seeds) and/or there are high levels of debris, the total number of analyzed particles must be increased appropriately. Unequal heights of the sample and standard peaks do not compromise the measurement accuracy (except for the visual presentation) as long as the number of nuclei in the smaller target peak does not fall below 1000. The minimum number of nuclei to be detected in each peak could be verified in preliminary experiments to check for effects of nuclei number on peak position.
- Check that the instrument is set up and functioning properly at the beginning of any series of measurements. This can be done by measuring a single calibration standard, such as *Pisum sativum*, which typically produces a very high quality and stable peak. If the settings are optimal, this measurement should give a perfectly symmetrical G_0/G_1 peak with a CV below 2%. If the peak quality is worse than this, one should consider adjusting the machine (if possible) and altering the method of production and staining of nuclei. In addition, if the peak quality deteriorates as more measurements are made, first consider carefully cleaning the machine before trying to adjust the machine settings. Note that peak quality depends also on the quality and age (storage) of the chemicals and solutions that are used, and so it is important to ensure that all reagents and buffers are appropriately stored and checked for contamination before each use (e.g., check if the solutions are cloudy which might indicate bacterial or fungal contamination). If the machine variability is chronic, one should consider moving to the latest generation instrument.
- Select the appropriate standard for the sample to be measured following the procedure above, and optimize buffer composition.[38]
- When undertaking the final measurements, set the gain (voltage) so that the peak of nuclei with the lowest relative fluorescence on the flow histogram is clearly distinguished from the debris with lower fluorescence.
- Set the flow rate to that which is optimal for the number of nuclei to be analyzed. The analysis of too many nuclei and/or setting the sample flow rate too high (or also too low) typically leads to higher peak CVs due to hydrodynamic focusing issues.
- Check peak symmetry and CV. Analyses showing asymmetric peaks (a common symptom when secondary metabolites or degraded nuclei are present) or peaks having CV > 5% should be discarded. In such cases, the instrument should be cleaned and adjusted, and/or sample preparation procedure modified. Measurements should then be repeated until high quality peaks are obtained.
- Record all parameters available as the machine output. The fluorescence should be stable during the time-course of the measurement (conveniently visualized on the 2D plot of fluorescence intensity vs. time). The 2D dot-plot showing fluorescence versus side-scatter is especially helpful when the presence of large amounts of debris and high background signals obscures signals from small(−genome) nuclei or samples with endopolyploid nuclei in a standard 1D fluorescence histogram. The signal from side-scatter visualized in a 2D plot enables debris to be clearly distinguished from relevant signals from the sample or standard nuclei and hence allows the use of unbiased gating of relevant fluorescence events. When such gating is applied, analysis using the derived (debris-cleaned) 1D histogram can be performed (Figure 3).
- Each plant should be analyzed at least three times on three different days to minimize the effects of any random measurement errors (day-to-day fluctuations), and to provide a basic insight about the extent of measurement variation. Measurements of a single plant repeated using the same machine in the same lab on consecutive days can show coefficients of variation (calculated as

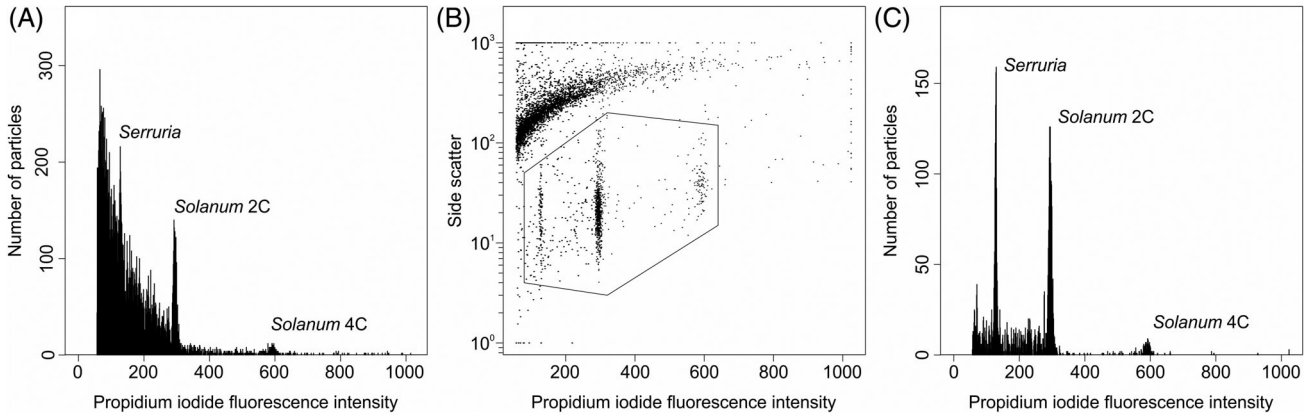

FIGURE 3 Identifying and removing debris by gating to increase the clarity of the FCM histogram. Measurement of a combined sample comprising *Serruria aemula* (sample, Proteaceae) and *Solanum pseudocapsicum* (standard) nuclei on a 2D dot-plot (B) showing high amounts of debris that obscure peak identification in the standard 1D fluorescence histogram (A). The debris particles, however, have different side-scatter profiles (forming a curve in B) and may be separated from the nuclei by appropriate gating (see box in B). This gating produces a cleaner 1D histogram (C) clearly showing the true (i.e., without contamination from the debris) peaks arising from the sample and standard nuclei, with target peaks having low CVs (*Serruria* - 2.35%, *Solanum* - 1.66%) and enough nuclei (1056 and 1416 nuclei in the G_0/G_1 peaks respectively; 12,000 particles measured in total)

100 × the standard deviation of individual measurements divided by their mean) of up to 3%. Outlying measurements should be discarded, and measurements repeated on different days until the variation in the data is below this threshold.

- Wherever possible, the genome size of a species should be measured in several plants from a population or from different localities. However, the measurement of a single individual per species is generally acceptable given the small variability, i.e., constancy of genome size, typically observed within a species (this variability is generally much smaller than the difference between measurements made of the same species in different laboratories[55]) and the relative rarity of plants showing genuine and considerable intraspecific genome size variation,[63] see also Section 4).
- Record and archive all primary data and metadata.

3.2 | Calculating genome size and presentation of results

3.2.1 | Calculation of genome size

Knowing the *MFI* (i.e., mean channel number) of the sample and the standard species from the fluorescence histogram, and the genome size of the standard ($gs_{standard}$), the genome size of a sample (gs_{sample}) is calculated according to the rule of three as (Figure 1):

$$gs_{sample} = \frac{MFI_{sample}}{MFI_{standard}} \times gs_{standard}.$$

This calculation assumes signal linearity and perfect proportionality between the amount of fluorescence and DNA content, which may not always be met (Figure 2). If this is not the case, or if uncertainty exists as to whether these assumptions are met, and a high level of accuracy of the genome size estimate is required, the sample should be prepared together (co-chopped) with two or more standards; such an approach largely overcomes such problems.[64] At its simplest, this may include the measurement of a sample with two internal standards (A and B), preferably one with a smaller genome size and the other with a larger size than that of the sample. In this case the genome size of the sample is calculated as:

$$gs_{sample} = \frac{MFI_{sample}}{(MFI_{standard_A} + MFI_{standard_B})} \times \left(gs_{standard_A} + gs_{standard_B}\right).$$

For alternative and more complex approaches used for generating very precise measurements of genome size, as needed for example when estimating the genome size of a new internal standard, see the reference 28.

3.2.2 | Reporting genome sizes as 1C, 2C, and Cx values

There is no consensus as to whether genome size should be expressed as 1C or 2C; instead, this choice tends to be field and/or group specific. It means that all genome size (nuclear DNA content) estimates that are reported must always be unambiguously denoted as either 1C or 2C, and ideally used consistently through the whole study (i.e., eliminating dual usage of 1C and 2C in one paper). When deciding whether to use 1C or 2C, one could also consider that 2C is more biological meaningful for studies focused on vascular plants as this is the typical state of somatic cell nuclei during their life cycle. 1C may be more meaningful for organisms where the dominant phase of

the life cycle is haplophasic (such as for bryophytes, i.e., mosses, hornworts and liverworts), for pollen studies, or for genomic projects.

Calculation of the monoploid genome size (1Cx; possible only if the ploidy level of the sample is known) may be the choice for studies of polyploid complexes and plant breeding material of different ploidy or for studies of genome evolution where recent polyploid events can make it difficult to select the most appropriate evolutionary model. It should be noted that the term 'genome size' for geneticists and molecular biologists conventionally refers to the haplophasic (1C) state. For example, the widely cited size of the human genome (3 billion bp) is the haplophasic DNA content as found in germ (sperm or egg) cells.

3.2.3 | Reporting genome sizes in picograms or base pairs

DNA content can be expressed either in picograms (pg) of DNA or as the number of base pairs (usually in Mbp or Gbp). The preference for either is mainly driven by historical traditions, with DNA content being described in terms of pg per 2C-nucleus for FCM studies. In contrast, data on genome sizes derived from sequencing projects are almost exclusively described as 1C-values and listed as the number of base pairs.

The amount of DNA in pg is derived from the molecular weights of the individual DNA bases. Using the assumption of a 1:1 ratio of AT to GC base pairs (i.e., GC content = 50%), and ignoring the existence of modified nucleotides, 1 pg of DNA is equivalent to 978 Mbp[39] (977.82 Mbp if a higher precision is needed). Although the GC proportion in plant genomes is nearly always lower than 50%[21,65] (see Section 5), the molecular weights of the AT and GC base pairs are almost identical (615.383 and 616.3711 g/mol, respectively) and deviations of the base composition from 1:1 has no practical effect on the final estimate of DNA content in pg. For example, for the plant with the lowest GC content estimated to date, that is, *Calypso bulbosa* var. *occidentalis* GC = 23.9%,[66] the conversion will be 1 pg = 978.23 Mbp, while for a typical nongrass angiosperm with GC ≈ 40%, 1 pg corresponds to 977.97 Mbp, nearly exactly corresponding to the rounded up conversion value 1 pg = 978 Mbp.[39]

3.2.4 | Repeating measurements and reporting measurement accuracy

Each estimate should be supplied with some measure of variation to document its methodical accuracy. This variation should be reported together with the average estimate of the repeated measurements, either calculated from the sample/standard ratios or from the genome size estimates from individual repeated measurements. The most frequently used measures of variation are standard deviation, standard error, and/or coefficient of variation. Providing the range of values observed is also useful. If the range of values are reported, one should bear in mind that this will increase with the number of replicates and cannot be interpreted as evidence of genuine intraspecific genome size variation. More robust tests are needed to demonstrate this (see Section 4).

In reporting the results, one should also report peak CV values and always present sample histograms representative of the measurements. The peak CV largely (but not necessarily) correlates with measurement accuracy unless any manipulation of the raw peak data has been carried out. For example, narrowing the signal range by manually gating the peak on a 1D flow histogram to reduce signal from debris is sometimes desirable but the resulting peak CV cannot then be used to argue for measurement accuracy. For older instruments, reporting peak CVs is meaningless for data that has been logarithmically scaled. However, the latest generation of cytometers take this into account, providing CV values for individual peaks that are calculated using linear scales even if the data are displayed using logarithmic axes.

3.2.5 | Including primary data in publications

It must be noted that the FCM estimate is only a hypothesis about genome size, dependent to a large extent on the reliability of the genome size estimation of the internal standard used. Although we consider that the genome size estimates for the currently available standards are reliable, other 'recommended' values for the same standard also exist,[28,67] and it is likely that even more exact estimates of their actual genome size will be provided in the future as more precise methods and techniques become available. Therefore, it is always advised to give the primary result of the FCM measurement (i.e., the standard/sample fluorescence ratio) in a publication (e.g., in supplementary material) or in a publicly available repository, to enable the future recalculation of genome sizes using an updated estimate for the standard. To be useful, the standard/sample fluorescence ratio should always be given with the name of the standard (including species, cultivar, or source locality or donor if a wild plant species has been used) and the genome size of the standard used for calculation of the sample genome size. Reporting the genome size estimate and the standard alone is less useful compared with the raw sample/standard ratio, since the estimate is typically rounded which decreases the numeric precision of any back-calculations that might be undertaken. Good practice in systematics research also includes preparing and storing a herbarium voucher of the analyzed plant, indicating the location of, and contact information for, the herbarium, as well as/or a photo 'voucher' of the measured plant to enable any future taxonomic revisions. All data from the FCM measurements should be made freely accessible over the internet.

3.3 | Potential pitfalls and how to avoid them

3.3.1 | Switching of sample and standard peak positions

One of the problems that may arise when measuring genome size is a switching in the position of the sample and standard peak in the FCM

histogram between different measurements. An awareness of this possibility is important, especially if the sample and standard peaks are quite close together (e.g., their fluorescence ratio is between ~0.7 and ~1.3), and always if the estimated genome size does not correspond to previously published values. One approach to ensure correct identification of the standard peak(s) in FCM histograms is to add additional standard nuclei (prepared as a separate sample of isolated nuclei) to the remaining sample + standard nuclei mixture, and then repeat the measurement with this new mixture. The peak corresponding to the standard will increase in height (or relative number of nuclei in the peak compared to the other peaks) when compared with the original histogram.

3.3.2 | Missing the G_0/G_1 peak

Another common problem is to misidentify the G_0/G_1 peak, particularly when estimating genome sizes using highly endopolyploid plant tissues (e.g., leaves of some Brassicaceae, Orchidaceae, Crassulaceae, and fleshy, succulent plants in general; see Section 7). In such tissues, the G_0/G_1 nuclei may occur infrequently, especially in older tissues, and hence the G_0/G_1 peak may be effectively absent, or may remain hidden within the debris region. A solution to this problem is to repeat measurements using the youngest and most actively proliferating tissue available (e.g., youngest leaves or root tips) which typically contain G_0/G_1 nuclei at a much higher frequency than older tissues.

For orchid species suspected of being characterized by partial endoreplication (see Section 7), genome size is most reliably measured from samples of immature ovaries where the cells remain mostly at the G_0/G_1 stage.[57,66] Some researchers have tried to overcome this problem by measuring genome size using seeds (embryos). This is certainly a good method to check for the presence of G_0/G_1 peak in a sample. However, it is still debated whether seeds are suitable for direct measurements of genome size and to what extent such measurements are universally comparable to those obtained from fresh tissues. This is because highly desiccated seed tissue differs from non-desiccated tissues both metabolically and epigenetically,[68] which, together with difference in chromatin structure,[69,70] can affect the accessibility of DNA for binding and its stoichiometry. Furthermore, the desiccated DNA is reported to adopt different conformations compared to the DNA in fresh tissue which have unclear effects on PI staining. This practice is therefore not recommended.

3.3.3 | Contamination leading to 'ghost' peaks

Sometimes 'ghost' peaks can appear in the FCM histogram of the plant, which may represent nuclei derived from parasites, symbionts, or pathogens. In such cases, the best approach is to repeat measurements with different, healthy plant tissues, where such peak(s) usually disappear. In extreme situations, axenic plants (i.e., obtained from tissue culture entirely free of all other contaminating organisms) can be sampled.

3.3.4 | Checking results with previously published genome size and ploidy data

When a new genome estimate has been generated, it is often useful to compare it with already published values for the same species or its closest relatives, for example, using data available in the Plant DNA C-values Database.[71] Nevertheless, one should be aware that some previously published estimates obtained using alternative methods, such as Feulgen densitometry, may be less reliable due to technical and instrumental difficulties.[72] In addition, older data generated using FCM may differ substantially from more recent FCM estimates and are often considered to be less reliable. In part this is because of progress in understanding the factors that influence the generation of robust genome size estimates and hence the adoption of best practice approaches for FCM (e.g., some older data were obtained using external standardization, base-specific fluorochromes, and/or unsuitable genome size standards). To check the reliability of the new estimates, one can compare them against all previously published values for a species in the Kew Plant DNA C-values Database (if available) by selecting the 'All estimates' option in the 'Show estimates' search type (the default option setting in the database is to show 'Prime estimates only').

Differences in the genome size obtained by using different internal standards can also introduce variation between genome size estimates for the same species. These differences can be reduced by recalculating genome sizes using the sample/standard ratios from the original paper (if given) and the currently accepted standard values.[28] Nevertheless, even in the recalculated data, differences up to 10% in FCM genome size estimates by different authors and labs may remain due to different instrumental and methodological factors.[55] On the positive side, comparison of archival values for DNA contents retrieved from the Kew Plant DNA C-values Database often reveal excellent linearity in comparison to fluorescence peak values for nuclei prepared from the same species.[73]

Another useful practice to uncover potential errors is to compare the estimated genome size of the measured species with its known ploidy (e.g., from the Chromosome Counts Database; http://ccdb.tau.ac.il).[74] This can be done most effectively by calculating the monoploid genome size (1Cx) for the measured species and its congeners. A doubling or other approximate multiplication of the monoploid genome size compared to the congeners usually reflects an incorrect association of the measured genome size with the known ploidy. If this is not the case and the monoploid genome size differs from previous estimates for the species, this could indicate that the sample belongs to a different species (which could easily be the case in taxonomically-difficult polyploid aggregates), or the sample peak used to calculate genome size did not correspond to the G_0/G_1 nuclei. If neither of the two previous options applies, the genome size estimate indeed suggests that the individual analyzed may be of a new ploidy level for the species. This should be verified by a chromosome count from the plant used to estimate genome size.[9,21] This fixes the genome size to a specific ploidy level and eliminates further doubts about this genome size estimate and its

ploidy. Nevertheless, there are examples when a doubling or other multiplication of the genome size compared to its congeners is not due to a difference in ploidy, but to differences in the abundance of repetitive DNA (e.g., see references 21,75,76). When counting chromosomes of the FCM-measured plant, it is good practice to screen the genome size in the tissue of the plant part used for chromosome counting (e.g., the root tip) to avoid any doubts about the association of the FCM estimate and the chromosome count. This is because there is a risk that making a chromosome count from roots of germinated seeds obtained from the mother plant used to estimate genome size could lead to errors if the ploidy level differs between the hybrid seed and the mother plant.

3.4 | Best practices

- For nuclei isolation, adopt recommendations provided by Loureiro et al.[38]
- Always use a single intercalating fluorochrome such as PI for the measurements. Results with base-specific fluorochromes (e.g., DAPI) or from different fluorochrome mixtures are unacceptable for precise genome size measurements, although these may be useful for preliminary genome size screening or for the detection of intraspecific genome size variation.
- For measuring genome size, use fresh material. The DNA in nuclei from dried (e.g., from a herbarium voucher, silica-dried sample, or seed) or fixed materials can have staining properties different from native DNA and may lead to unreliable results.
- Always use a plant internal standard, which is co-chopped and stained together with the sample (not chopped and stained separately and then mixed prior to analysis). Data obtained using external standardization are unacceptable.
- Use a plant internal standard whose genome size is similar to the sample being analyzed (ideally 0.5-2-times the size of the sample and certainly not less than 0.3- or more than 3-times).
- If very accurate estimates of genome size are needed, consider using two or more internal genome size standards (ideally with the genome size of the sample falling between those two standards, see reference 28).
- Note that currently available estimates are still hypotheses of the actual DNA content. Therefore, always report the primary data of the FCM measurements, that is, the standard/sample ratio. This will enable future recalculations and corrections to be made if a more accurate genome size estimate for an internal standard is determined.
- Repeat measurements of genome size for an individual using nuclei isolated on different days are desirable to get an idea of the variation between estimates and hence assess the potential presence of methodological issues. This variation, expressed as the standard deviation or coefficient of variation of the genome size estimate or sample/standard ratio, should always be documented together with the mean estimate of the measurements. Note that recording peak CV, although also recommended, is less informative in this respect.
- If low quality results (for example, target peaks with peak CVs > 5%, or containing less than 1000 nuclei) are obtained, optimize the sample preparation and/or instrument settings and repeat the measurements until the quality criteria are satisfied.
- Compare new estimates with available data for the same or closely-related species. Where discrepancies are apparent, make sure that the standard and sample peak positions have not been confused and/or that the G_0/G_1 peak has been correctly identified.
- In systematic studies handling nonmodel species, always prepare a herbarium voucher of the measured plants to enable taxonomic verification.
- Always indicate whether your genome size estimates are 1C or 2C, and use either of those exclusively and consistently in your reports to avoid any confusion.

4 | INTRASPECIFIC VARIATION IN GENOME SIZE AT THE HOMOPLOID LEVEL

For differences in genome size between species to evolve, such divergence is first expected to arise at the population level, resulting in the presence of individuals having different genome sizes. The underlying processes giving rise to intraspecific variation can be complex, ranging from changes in the activity of retrotransposons and other types of repetitive DNA sequences, and deletions/insertions of various DNA fragments, to variation in chromosome numbers (e.g., aneuploidy, B-chromosomes). Variation in genome size due to the occurrence of different ploidy levels or arising from taxonomic issues are not considered here to represent intraspecific genome size variation. Knowledge of intraspecific genome size differences may play a significant role in understanding the link between speciation and genome size evolution in plants.[63] Due to the difficulties of detecting very small differences in genome size between individuals and demonstrating that they are real rather than a result of technical artifacts, the study of intraspecific variation in genome size represents one of the most methodologically challenging applications of plant FCM.

4.1 | Sources of errors

False evidence of intraspecific variation can easily arise from various methodological bias and poor practice, and at least 50 studies that reported intraspecific variation in genome size have subsequently been shown to be false due to a variety of reasons.[63] Many of these reports have been based on using methods such as Feulgen densitometry, where accuracy depends on multiple critical steps and is therefore quite error-prone.[72,77,78] Nevertheless, there are also numerous erroneous reports of intraspecific genome size variation using FCM.[63]

The most common sources of errors when using FCM include:

- Lack of standardization or use only of external standardization.
- The presence of different types and amounts of metabolites, which interfere with the quantitative binding of the fluorochrome to DNA.[38,79–82] These metabolites can cause errors in genome size estimates even if an internal standard and a buffer supplemented with antioxidant is used.
- Aging of the hardware components of the cytometer, and a lack of awareness of the thermal dependence of the signal multipliers in FCM machines[83] both of which are difficult to control. This means that some cytometers can produce slightly different results (minute, but statistically significant) on different days or at different times of the day.

4.2 | Specific requirements and problem solving

The sample processing and FCM measurements used to detect intraspecific variation do not actually differ from the standard protocol for genome size estimation (see Section 3), except for the very stringent requirements for measurement accuracy. The greater the accuracy, the lower the chance of the results providing false evidence of variation, and the higher the probability of reliably detecting even minute differences in genome size. In addition, there are specific requirements for measurement design to avoid generating erroneous data that might lead to false claims of intraspecific variation.

4.2.1 | Improving accuracy

- Using the most appropriate nuclei extraction and staining procedure for the studied species,[38] or a different fluorochrome compared with conventional measurements of genome size (see below).
- Where possible, selecting only the most healthy, fresh, and secondary metabolite-free parts of the sample and standard tissue for measurements.[22,38]
- Using a plant internal standard whose genome size is close to that of the species being studied. This gives more reproducible results in repeated analyses of the same individual than when the genome size of the internal standard is more distant from the species being analyzed.
- Increasing the number of nuclei in the G_0/G_1 peak (to at least 3000).
- Using an internal standard that has been grown under the same conditions for the whole series of measurements. This helps to minimize measurement bias due to changes in metabolites, which can interfere with the stoichiometry of binding of the fluorochrome to the DNA.
- Growing the plants being investigated under the same conditions for a reasonable time for the reasons given above for the internal standard. If plants need to be transplanted for analysis beyond one growing season or to be grown in different environmental conditions, they should first develop new leaves so that these can be used for measurements.

4.2.2 | Generating indisputable evidence

Document the differences in genome size by demonstrating the occurrence of double peaks or a bifurcating peak in the FCM histogram when making measurements of co-chopped samples predicted to differ in genome size (Figure 4). This is the only direct and indisputable evidence of intraspecific variation. In general, split or bifurcated peaks can only be observed in samples where the difference in genome size is twice that of their peak CV.[84] With the most precise methods of sample preparation and nuclei staining, in practice differences as small as 1.02–1.04-fold have been detected.[85] The difference observed between the samples in the double peak(s) should be the same as the difference estimated from running each individual separately with the internal FCM standard (for examples, see references 86–88). For testing the agreement between genome size estimates from double peaks and from individual estimates, see references 89 and 90.

Because knowledge of genome size in absolute units in this kind of study is generally less important than measurement accuracy and the verification that the intraspecific variation as genuine, it may be easier to work with alternative fluorochromes or internal standards to those recommended for measurements of genome size. The most popular fluorochrome used for detecting intraspecific variation is the AT-specific DAPI, which produces narrower and much more easily distinguishable peaks in most plants using the standard sample preparation procedures (see Section 5) (e.g., see references 60,61,84,90,91), with peak CVs sometimes even below 1%.[85] It should be noted, however, that results from measurements with DAPI must not be used to convert relative fluorescence values into absolute genome size. If such data are required, base-independent binding fluorochromes (such as PI) must be employed (see Section 3).

4.2.3 | Proper measurement design

Where it is not possible to demonstrate the existence of double peaks, but variation is suspected and is considered biologically interesting (e.g., there is a correlation between the genome size difference and some biological factors), then it is essential that the measurements of the individual samples are performed following the recommendations below, to minimize errors arising from temporal/temperature bias of flow cytometers:

- Samples being compared should always be prepared using the same method for nuclei isolation and be measured with the same flow cytometer.
- Special attention should be paid to repeat measurements of the same samples on different days and to measure all samples within each measurement run in random order with respect to the hypothesis being tested. The random order should be generated before any measurements are made, ideally using a random number generator (available, for example, in MS Excel or R). Thus, if the aim is to show a correlation between genome size and, for

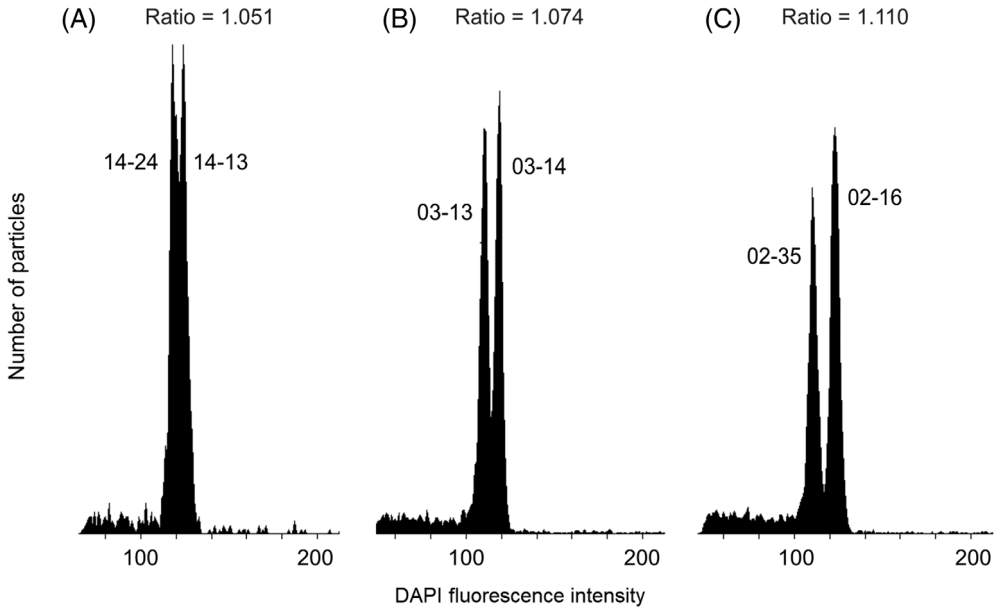

FIGURE 4 Examples of a bifurcated peak (A) and double peaks (B, C) obtained in a single flow cytometry measurement (histogram) indisputably documenting the existence of differences in genome size between individual progeny from mother plants of *Festuca pallens* (Poaceae). The individuals differing in genome size were co-chopped in one dish and the filtered nuclei mixture was stained with DAPI. Sample number (mother number-progeny number) and the ratio of the mean peak fluorescence are given. Adapted from Šmarda et al.[89]

example, altitude (or other geographical variables), samples from lower altitudes should not all be measured at a different time of day or even on different days from the higher altitude samples.

- Ideally, all samples should be measured within the same day and on the same machine. If there are too many samples for this to be feasible, then the effect of analyzing samples on different days must be properly handled by randomizing the order of analyzed samples (as noted above). The samples should then be measured consecutively in the given order irrespective of the day. It should be noted, however, that as the number of measurement days increases, so too does the introduction of instrumental bias in the data, and hence the ability to detect genuine intraspecific variation decreases.
- Results of repeated measurements of the same samples should not be averaged (as this leads to a loss of information about measurement error), but the hypotheses based on the data should be tested using the appropriate statistical methods for handling repeated measurements and associated measurement errors (e.g., using mixed-effect models and treating blocks or days of measurements as random factors).

Studies reporting small difference in the standard/sample ratios (genome size) between samples (or sample groups) that were measured on different days or using different cytometers cannot be accepted as evidence of genuine genome size differences. This is one of the very common mistakes reported in many published FCM studies. It often arises because many authors, in attempting to discuss all their genome size data, observe differences in genome size between individuals, even when the study of intraspecific variation was not the primary aim of the study, and hence the experimental design did not follow the best practice approaches outlined above. Without careful measurement design and/or confirmatory proof with double peaks, any intraspecific differences in genome size should be considered a priori as unreliable.

4.3 | Best practices

- Beware of the potential false evidence of intraspecific variation caused by such factors as different concentrations of metabolites in different tissues, plants grown under different environmental conditions, and/or temporal measurement bias due to aging and/or temperature fluctuations of flow cytometer components.
- Improve the accuracy and resolution of measurements by rigorously following the recommendations: (i) selecting the best isolation buffer giving the lowest CVs, (ii) ensuring the fluorescence signal of the standard is close to that of the sample, (iii) ensuring there is sufficient material of a single individual of the internal standard for the whole series of measurements, and (iv) working with the standard and all samples cultivated under the same controlled conditions (although this can differ for the standard and the samples in some specific cases) throughout the duration of the experiment.
- Consider using alternative fluorochromes, such as AT-specific DAPI since in these kinds of studies improving the accuracy and resolution limits of the flow cytometer are more important than the knowledge of the genome size in absolute units.
- Confirm, wherever possible, any reported intraspecific differences in genome size with a FCM histogram showing a double or bifurcating peak obtained from measurements of co-chopped samples.
- If analyzing and discussing differences that cannot be demonstrated by double peaks, ensure the following:
  - Samples must be measured in a random order with respect to the tested hypothesis to prevent any impact of temporal measurement bias of the cytometer. Small differences in genome size (not confirmed by double peaks), even if statistically significant, cannot be considered as reliable without proper (explicitly described in the paper) random measurement design.
  - Measurements should be repeated several times on different days or seasons using the same nuclei isolation procedure and

cytometer and the results of repeated measurements should be analyzed using statistical methods that can handle data arising from repeated measurements of the same sample (e.g., mixed-effect models).

5 | DNA BASE COMPOSITION MEASUREMENT

Like genome size, the base composition of a genome, that is, the proportion of adenine (A), guanine (G), cytosine (C) and thymine (T) in the DNA is a basic species characteristic that may be measured using FCM.[58,65,92–94] The DNA base composition is most frequently reported as the percentage content of GC and AT bases, reflecting the DNA base pairing. The usage of GC or AT content to express DNA base composition is a matter of choice as these measures are mutually convertible, given that AT% + GC% = 100%. GC content is more frequently used in the existing literature and is more easily searchable in full texts since the 'GC' double-letter combination is rarely encountered in the English language. Knowledge of GC content may find application in genomics as a cheap independent check for genomic sequence completeness and its possible base bias,[94,95] searching for species with peculiar genomic composition worthy of future sequencing, and tracing genomic processes related to base composition in evolution or for systematic purposes.[35,54]

Due to its broad variation, GC content is commonly used to characterize prokaryotic genomes.[96–98] It varies considerably across vascular plant species, ranging between 24% and 51%.[21,66] This can have several consequences for plants due to the different physicochemical properties of GC and AT base pairs. For example, DNA that is GC-rich has higher thermal stability, higher bendability, lower packability, and higher metabolic (synthetic) costs compared with AT-rich DNA.[54,99,100]

5.1 | Principles

The measurement of GC content using FCM is conceptually similar to the measurement of genome size, in that the sample comprises an unknown species and an internal standard (see Section 3), and that the analysis can easily be implemented. The simplicity of the method and the low requirements of material favor FCM over other analytical methods (e.g., thermal denaturation, paper or high-performance liquid chromatography, buoyant density centrifugation, measurement of UV absorbance after DNA bromination, reviewed in reference 93), and therefore nearly 90% of existing data on plant GC content has been obtained with FCM.

The principle lies in the comparison of two sample species/internal standard fluorescence ratios–one obtained from analysis of the FCM sample stained with a base-nonspecific fluorochrome (usually PI), and the other one with a base-specific fluorochrome (usually the AT-specific DAPI). If a sample species with fewer AT bases than the internal standard is stained with the AT-specific fluorochrome, the resulting fluorescence ratio will be lower compared to the ratio obtained from measurements made using a base-nonspecific fluorochrome (Figures 5 and 6). This species- and dye-specific shift in the sample species/internal standard fluorescence ratios, termed the dye factor, is the basis for calculating the AT/GC proportions.

5.2 | Measurement

The measurement must follow the best practice recommendations given for the measurement of genome size (see Section 3). However, two cytometers differing in excitation sources are typically employed for the measurement of GC content, and therefore double the amount of plant material of the sample species and the standard is typically co-chopped in double the amount of isolation buffer, compared with the amounts used for measuring genome size. Following filtration, the nuclei suspension is divided into two parts, and each is stained by a different fluorescent dye (e.g., PI and DAPI). These are analyzed separately on different cytometers, using the appropriate excitation source (e.g., green light for PI and UV light for DAPI). While it is possible to isolate nuclei for PI- and DAPI-staining separately and analyze them on different days or in different laboratories, this approach increases the measurement error and is not recommended. Simultaneous staining and analysis of nuclei with two or even more fluorescent dyes should also be avoided as the competition of the dyes for binding sites on the DNA causes unpredictable effects on the fluorescence intensities detected for each dye, rendering exact calculations of base composition unreliable.

5.3 | Base-specific fluorochromes and binding lengths

Several fluorochromes are commercially available that are specific for AT- or GC-DNA bases, such as AT-specific DAPI, 4′,6-bis (2′-imidazolinyl-4H.5H)-2-phenylindole (DIPI), and several Hoechst dyes, and GC-specific chromomycin, olivomycin,[93] and mithramycin.[101] DAPI seems to be particularly preferred, since: (i) its signal is highly reproducible and provides sharper peaks (lower CVs, even <1%) compared with other AT-specific dyes,[85,102,103] (ii) its small size limits problems encountered with staining the DNA in large nuclei with highly condensed chromatin using standard FCM protocols, and (iii) staining with DAPI does not require an RNase treatment (which is necessary when using intercalating dyes; see Section 3). It is therefore frequently used in other FCM applications and is thus readily available in many laboratories.

DAPI binds preferentially to the DNA minor groove, forming highly fluorescent complexes with at least three consecutive A or T bases.[104] When bound, the DAPI fluorescence increases more than 20-fold, with maximum excitation at 363 nm and maximum emission at 448 nm.[104] At the concentrations typically recommended for plant FCM applications (i.e., 4 μg/ml),[60] DAPI seems to behave like a standard intercalating dye, as it also binds to GC-rich DNA and double-stranded RNA, however, this binding results in no or much lower fluorescence compared to

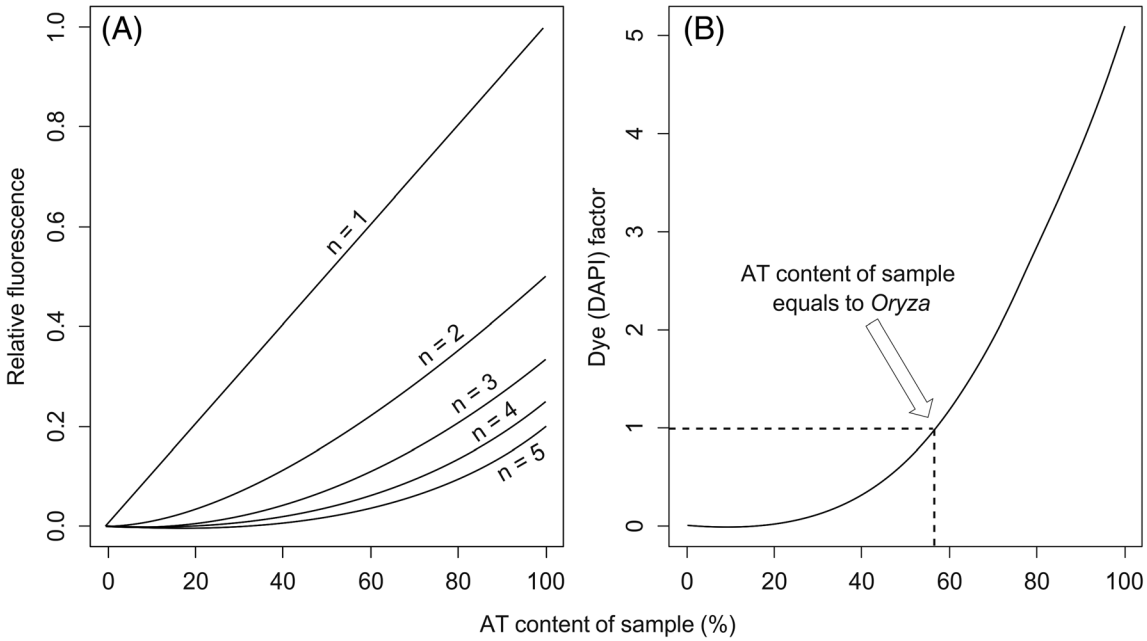

FIGURE 5 Relationship of sample fluorescence and binding length (n) of an AT-specific fluorochrome with AT content of a sample, assuming a random distribution of base pairs. (A) Relationships between the sample fluorescence (in relative units from minimum = 0 to maximum = 1) and the AT content of a sample for different binding lengths of an AT-specific fluorochrome. (B) Relationship between a sample's AT content and its dye factor obtained from a sample measurement using the AT-specific fluorochrome DAPI (binding length $n = 4$) and *Oryza sativa* 'Nipponbare' (AT = 56.4%) as the standard. A dye factor > 1 indicates that the AT content of a sample is higher than *Oryza*, while a dye factor < 1 indicates that AT content of a sample is lower than in *Oryza*

when it is bound to AT bases within the minor groove, and the emission maximum shifts when it binds to RNA.[104–106]

$$DF_{sample} = \frac{MFI_{sample}DAPI / MFI_{standard}DAPI}{MFI_{sample}PI / MFI_{standard}PI}.$$

5.4 | Standards

For practical reasons, the internal standards used for measuring the DNA base composition are the same as those used for estimating genome size (see Section 3). With possible differences in the positions of the sample and the standard peaks when measured with different dyes, identifying a standard fulfilling the criteria of Section 3 (i.e., a small difference in the positions of the sample and standard peaks, but with no overlap) is usually more difficult for this application, and sometimes requires prescreening of the peak position with both selected fluorochromes. Use of animal standards should be avoided because of differences in genome organization and distribution of bases compared to plants[107], which might affect the binding properties of the base-specific fluorochromes.

5.5 | Calculation

The simple measure of the relative base composition of a sample is provided by the dye factor (*DF*). It is calculated from the *MFI* of the nuclei of the sample and the standard stained with the base-specific dye (e.g., DAPI) and the base-nonspecific dye (e.g., PI), according to the following equation[92]:

While the application of base-nonspecific fluorochromes enables comparisons of nuclear fluorescence in a linear fashion (such as is used when calculating genome size), the base-specific fluorochromes usually require several consecutive base pairs of the same type (AT or GC) to form fluorescent complexes with DNA.[93] With DAPI, for example, the number of bound bases, that is, the binding length of the dye, varies between 2.77 and 4.22, depending on the particular genomic sequence and base pair motif, with binding length equal to four bases being the most universal for FCM measurements in plants.[92,94] Because the probability of such base clustering depends nonlinearly on the AT/GC content of the genome and the observed fluorescence is therefore nonlinearly dependent on the amount of target bases (Figure 5A), the estimation of GC or AT contents from FCM data in absolute percentages requires a more sophisticated calculation.[92,108,109] For practical reasons, all calculations assume: (i) a random distribution of bases in the genome and (ii) a constant number of contiguous base pairs needed for a fluorochrome to bind to the DNA molecule (i.e., 'binding length'). Knowing the dye factor of the sample (DF_{sample}), the AT content of the standard ($AT_{standard}$, ranging from 0 to 1), and the binding length of the AT-specific dye (b), the AT content of the sample (AT_{sample}) can be calculated from the following relationship[109]:

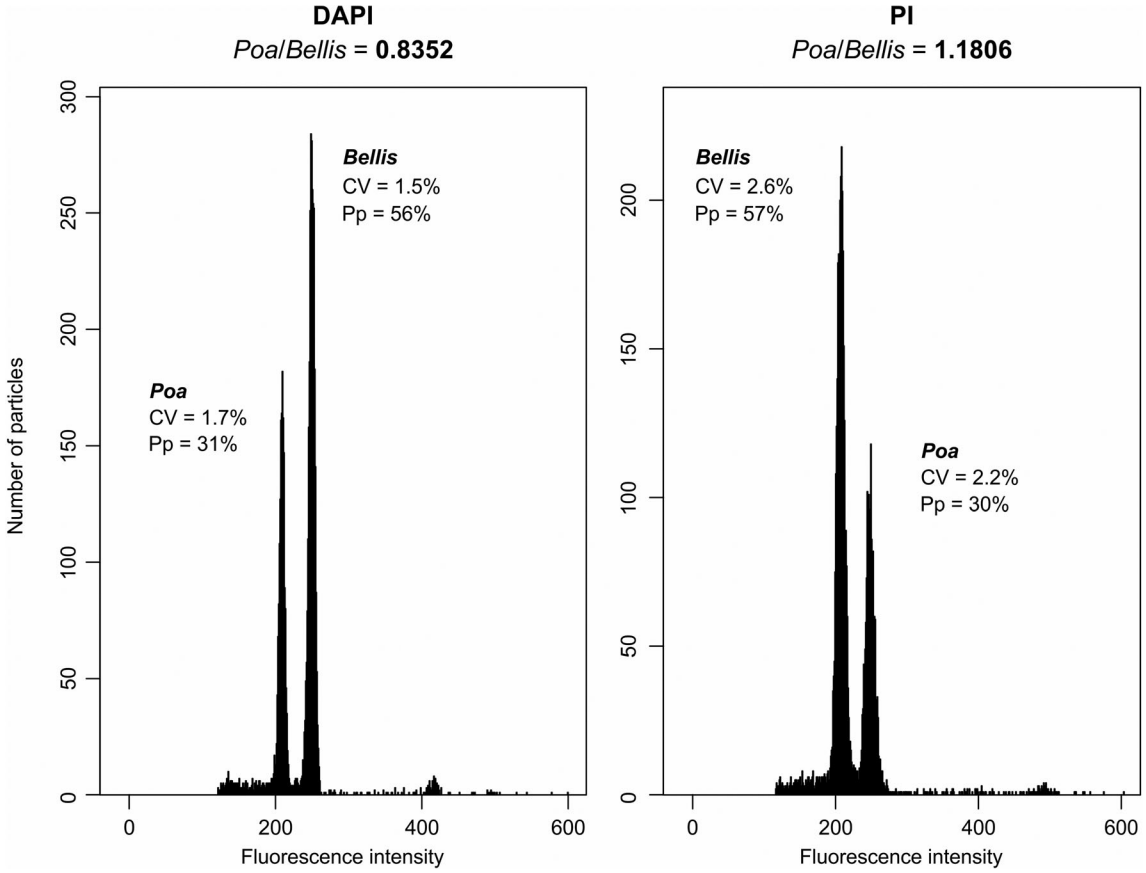

FIGURE 6 Comparison of nuclei fluorescence in FCM histograms obtained by measuring two species with contrasting GC contents (AT-rich, GC-poor *Bellis perennis* and AT-poor, GC-rich *Poa palustris*) using base-nonspecific PI and AT-specific DAPI. While *Bellis* clearly has a smaller DNA content than *Poa* based on measurements using PI, the measurement of its AT fraction using DAPI indicates it has a more AT-rich genome than *Poa*. Because nuclei were isolated together for measurements with both dyes, peaks have a similar proportion of analyzed particles, which can help with their (peak) identification, alongside the observation that only *Poa* has a small G_2 peak

$$\frac{(1-AT_{sample})AT_{sample}{}^{b}}{1-AT_{sample}{}^{b}} = DF_{sample}\frac{(1-AT_{standard})AT_{standard}{}^{b}}{1-AT_{standard}{}^{b}}.$$

Solving for AT_{sample} requires some approximations. This can be done, for example, using the *regula falsi* method, which is available in a Microsoft Excel sheet that can be downloaded from http://www.sci.muni.cz/botany/systemgr/download/Festuca/ATGCFlow.xls.[35] For calculating GC contents from measurements using GC-specific fluorochromes, the 'AT' in the above formula is replaced with 'GC'. As an example, from a comparison of nuclei fluorescence obtained from analyzing nuclei of *Poa palustris* and *B. perennis* stained either with DAPI or PI (Figure 6), the $DF_{Poa} = 0.8352/1.1806 = 0.7074$. Assuming a binding length of 4 for DAPI and an AT content of *B. perennis* of 0.6046,[54] the DNA base composition of *P. palustris* is GC = 45.8% and AT = 54.2%.

5.6 | Measurement quality and estimation of errors

Since DAPI usually provides tighter peak CV values than PI, the quality criteria should be more stringent for measurements with DAPI than for standard genome size estimations using PI (see Section 3); thus, only peaks having CVs below 3.5% should be accepted for measurements with DAPI.

Calculating the degree of variation (measurement error) in the measurements must be done using the dye factors only, because the mathematical nature of calculating the final AT/GC proportions or percentages disallows the use of some mathematical operations like averaging or subtracting. There is no strict rule, but it is recommended that the measurements are acceptable if the coefficient of variation (i.e., 100 × standard deviation/mean) of dye factors from repeated measurements of the same event is below 3%. If higher than this threshold, outliers should be omitted, and measurements repeated until the 3% threshold is passed. An alternative is to express the measurement error as the minimum-maximum range of the results. However, this measure naturally increases with the number of measurements and can be difficult to compare between estimates based on different numbers of repetitions.

5.7 | Common errors, reliability verification, and method limitations

The recommended ways to avoid errors and to verify the reliability of the measurements are similar to those described for genome size

measurements (see Section 3). However, compared to measurements of genome size, GC content measurements are more susceptible to operator error because of the potential to misidentify the sample and standard peaks. This is because the amount of nuclear fluorescence, and, consequently, the relative positions of the peaks in the histograms, can differ considerably between DAPI and PI staining (Figure 6). This difference in peak positions is particularly prominent when the sample and standard species differ considerably in GC content. Thus, caution is needed, for example, when analyzing species of Poaceae, which typically have high GC contents. In general, a problem with peak identification occurs only when the sample/standard ratio falls in the range of 0.7–1.4. Since incorrect assignments of peaks will dramatically change the resulting GC content calculation, potential errors of this type can be detected by comparing the obtained values with those published for the same species or its close relatives (e.g., see reference 21). Further verification of peak assignments can be done by adding post hoc stained nuclei of the standard to the sample/standard nuclei mixture and observing which peak changes in peak area or height in the repeated measurement (see Section 3). Since the reported GC contents of most flowering plant species range between 35% and 45%, with the exceptions of the Poaceae (GC = 43%–50%), Orchidaceae (GC = 24%–51%) and Cyperaceae (GC = 33%–42%), for GC% values outside these ranges, it is recommended that the reliability of the measurements are confirmed with different internal standards.[21]

One of the problems leading to errors in measurements of DNA base composition is the assumption of a random base distribution within the genome, a situation, which is infrequently encountered in nature.[92,94] Another problem is the uncertainty associated with dye binding length, indicated by the slightly different affinity of DAPI to AT motifs of different lengths and composition.[110,111] Despite these problems, FCM-based estimates of GC contents are in good agreement with results using other methods such as DNA melting.[94] In addition, although there have been few studies to date, there is also emerging evidence of a good agreement between FCM-established GC contents and estimates obtained from whole genome sequencing data, for example, in *Utricularia gibba* (GC = 39.9% from FCM[112] and 39.93% from sequencing[113]). A final problem is that the FCM method combines results of two separate measurements, and is therefore susceptible to the combination of errors, leading to higher variability in these results compared with estimations of genome size. Considering all these reasons, differences in FCM estimates of GC content lower than 1% among closely related species, or up to 2% among distantly related ones, must be interpreted with caution.[94]

5.8 | Best practices

- Follow best practice recommendations for genome size measurements for nuclei isolation and measurements with any fluorochrome (see Section 3).
- Ideally, prepare single samples of nuclei, and divide each into two sub-samples for measurement of the fluorescence of two different fluorochromes.
- Do not stain the same nuclei simultaneously with more than one fluorochrome.
- Beware of peak shifts in samples stained with different fluorochromes when the sample and standard species differ considerably in GC content. Take care to examine whether peak switches may have inadvertently occurred, especially when measuring unknown species which have nuclear fluorescence values that are within 0.7–1.4-fold of the standard.
- Beware of the nontrivial calculation of GC content from the FCM data and use the appropriate tools for its calculation, (e.g., available at http://www.sci.muni.cz/botany/systemgr/download/Festuca/ATGCFlow.xls).
- Carefully check all GC proportion estimates that fall outside the 35%–43% range (40%–50% in the case of Poaceae), or that deviate from estimates reported for closely-related species. Verify such estimates by re-measuring the sample with an alternative standard that differs in GC content from the previous one.
- When calculating measurement error in the data, perform these calculations with dye factors only, and not with final GC contents, since this would be mathematically incorrect.
- Keep in mind the error inherent in the GC content estimation, thus small differences (<1%) found in routine measurements within closely-related species, or larger differences (up to 2%) between unrelated taxa are unlikely to be biologically relevant and should be interpreted with caution.

6 | CELL CYCLE ANALYSES

The mitotic cell cycle is a process characteristic of proliferating cells, required to increase the cell count during growth of eukaryotic organisms.[114] It comprises four phases: Gap 1 (G_1), DNA synthesis (S), Gap 2 (G_2), and mitosis (M) (Figure 7). Cells typically spend the greatest proportion of the cell cycle in the G_1 phase, wherein they are metabolically active and undergoing growth through cellular expansion. In the G_1 phase, diplophasic nuclei possess a 2C DNA content. During the following S phase, the amount of nuclear DNA gradually increases due to DNA synthesis, to reach a doubled 4C DNA content on entry into G_2, the phase within which cells prepare to undergo mitotic division. The G_2 phase involves a period of additional cellular growth and protein synthesis, but in the absence of increases in nuclear DNA content. The G_2 phase is followed by mitosis, during which the replicated nuclear DNA is divided equally between the two daughter cells, so they each contain a 2C DNA content. After proliferation is completed, cells exit the cell cycle, usually in the G_1 phase, and enter a quiescent stage termed G_0. Haplophasic nuclei have a 1C DNA content in G_1 and G_0 and a 2C DNA content in G_2. Based on the different DNA contents in the nuclei of cells within specific phases of the cell cycle, FCM analysis can provide precise measurements of the distribution of nuclei between the G_1, S, and G_2 phases (Figure 7A). However, because of the lack of a nuclear envelope during mitosis, cells in M phase can only be detected in single cell/protoplast suspensions (i.e., they cannot be detected by nuclear FCM). Routinely used

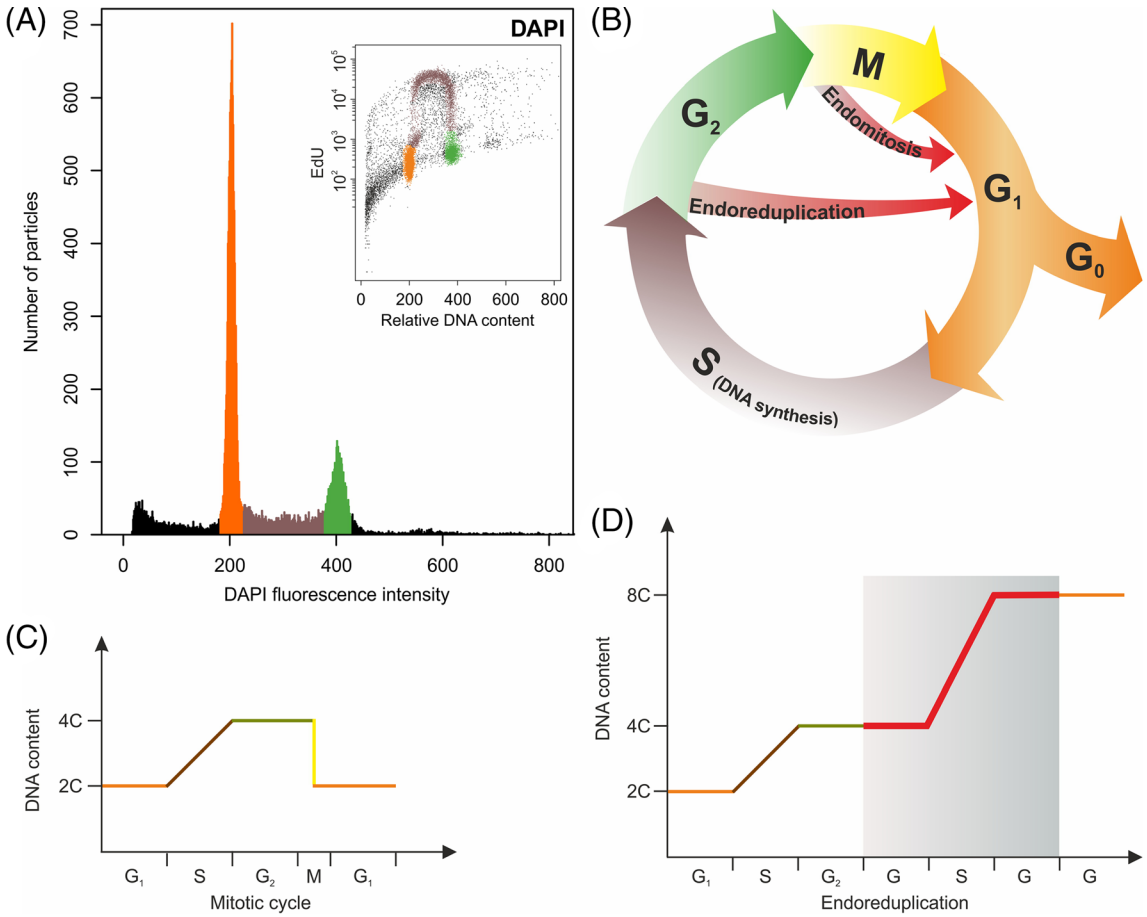

FIGURE 7 Changes in nuclear DNA content during the mitotic and endoreplication cycles. (A) Cell cycle analysis in root cells of wheat (*Triticum aestivum*) using FCM. The root nuclei were pulse-labeled with halogenated 5-ethynl-2′-deoxyuridine (EdU). In the inset, the X-axis represents relative DNA content measured as the intensity of DAPI fluorescence (linear scale). The Y-axis shows the extent of EdU incorporation as measured by Alexa Fluor 488 fluorescence intensity (log scale). The analysis was kindly provided by P. Cápal (UEB, Olomouc). (B) Simplified model of the cell cycle and endoreplication. The difference between the duration of two types of endoreplication, that is endomitosis and endoreduplication, is shown. C-DNA content in different phases of the cell cycle (C) and endoreduplication (D). The first endocycle is shaded gray (D). The corresponding cell cycle phases in (A), (C), and (D) are marked using the same color and follow the colors shown in (B) [Color figure can be viewed at wileyonlinelibrary.com]

software (e.g., Flomax, ModFit, FCS Express) enables the proportions of the different cell cycle phases to be quantified using univariate FCM analysis. Peaks can also be marked manually, and the proportions calculated using any spreadsheet program (e.g., Excel).

However, not all phases can be distinguished using this approach. Thus, cells in G_0 and those in G_1 both contain the same amount of DNA, and so contribute to the same peak of fluorescence as measured by FCM. To discriminate actively dividing cells in G_1 from those in G_0 that are not undergoing cell division, as well as for studies of cell cycle traverse rates and phase transition times, bivariate analysis is required.[115–117] In such analyses, two fluorescent signals are detected simultaneously. The first, obtained after staining with a standard DNA-specific fluorescent dye (e.g., PI or DAPI), quantifies nuclear DNA content. The second is obtained after immunofluorescent staining of 5-bromo-2-deoxyuridine (BrdU) or halogenated 5-ethynyl-2′-deoxyuridine (EdU), both analogues of thymidine. This quantifies the amount of BrdU/EdU incorporated by a previous period of in vivo labelling into DNA during replication (i.e., thereby identifying nuclei in S-phase; Figure 7A).

The mitotic cycle is particularly active in fast growing plant tissues, such as young and germinating embryos or meristems. The continuous proliferative activity of the latter provides cells for growth and maintains the meristem throughout the life of the plant.[118] High mitotic activity is manifest by an increased proportion of 4C cells, and therefore its identification by FCM can provide information on the physiological state of the plant/plant part.

This can be illustrated by the example of following the distribution of nuclei with different DNA contents during seed development and maturation. At the beginning of seed development, the proportion of 4C cells in the embryo can reach 30%–45%,[119] whereas during maturation, DNA synthesis and mitosis usually cease and the overall proportion of 2C nuclei increases; in some species, in the dry mature seeds 100% of the nuclei have 2C DNA amounts (e.g., lettuce or pepper). This proportion does not change during Phase I of germination, when water uptake (imbibition) occurs, because during this time only DNA repair processes occur without DNA synthesis. However, in Phase II of germination (i.e., the phase that precedes radicle protrusion), when metabolic events such as reserve mobilization, protein synthesis, and DNA replication (without

mitosis) are initiated, the proportion of 4C nuclei increases again, indicative of the completion of germination. Knowledge of the pattern of DNA synthesis can be used by seed producers for establishing the optimal harvest time and conditions for seed priming (methods to physiologically enhance seed performance using hydration techniques followed usually by drying[114]). In addition to the proportion of 4C nuclei, the 4C/2C (G_2/G_1) ratio can be used as a measure of cell cycle activity.[120] However, there are no universal values of these parameters for distinguishing the particular stages of maturation or germination; indeed, the developmental/germination pattern should be established individually for each species.[119] In addition, the different vigor of seed lots of the same species can bias the dynamics of DNA synthesis since seeds with low vigor require more time for DNA repair. In polysomatic species (characterized by endoreduplication of differentiating cells; see Section 7) the 4C/2C ratio can be biased by the presence of 4C nuclei entering an endocycle, which cannot be distinguished from the 4C nuclei in the G_2 phase of the mitotic cycle. For such species, especially for the later stages of seed germination or seedling growth, the usefulness of this ratio is often limited, and alternative measures should be applied (see Section 7).

Another application of establishing cell cycle activity by FCM involves studying the effect of antimitotic, allelopathic, and genotoxic compounds on plants and cell growth. Together with microscopic observations of the mitotic stages and detection of the occurrence of chromosomal aberrations during mitosis, FCM can provide information on the mode of action of specific compounds which inhibit plant growth.[1] The anticancer properties of phytochemicals can also be established by treatment of human or other animal cells, followed by FCM analysis of the cell cycle and of apoptosis.[121,122] McMurphy and Rayburn[123] reported an increase in the CV values of G_1 peaks and the frequency of G_2 nuclei following coal fly ash treatments compared with the control group. In addition, DNA damage was detected in lettuce (*Lactuca sativa*) plants exposed to cadmium,[124] while cell cycle arrest was observed in *Allium cepa* exposed to X-ray radiation,[125] and in *P. sativum* plants exposed to chromium (VI).[126] In such experiments, the high proportion of 4C nuclei should not be confused with nuclei entering endoreduplication (see Section 7).

FCM analysis of the cell cycle is also used in chromosome analysis and sorting, which requires the cell cycle to be chemically synchronized in the plant tissue to produce the highest possible proportion of metaphase cells (for details see reference 127). FCM allows easy treatment optimization and the tracking of cell cycle progression of the synchronized populations.

Technical details for univariate FCM analysis of the cell cycle are the same as for ploidy and endoreduplication analysis (see Sections 2 and 7). Internal standardization is not recommended in FCM cell cycle analysis since the nuclei/background of the standard can bias estimations of the proportions of the cell cycle phases in the sample.

6.1 | Best practices

- Sample preparation and measurement should follow the rules for ploidy estimation (see Section 2); however, including an internal standard is not recommended.
- Be sure to select the appropriate tissue (containing meristematic cells); analysis of the cell cycle in endopolyploid tissues is limited because FCM cannot distinguish 4C cells in the G_2 phase of the cell cycle and those entering endocycles.
- Use of linear as well as logarithmic amplification of the FCM signal is acceptable.
- Employ the 4C/2C (G_2/G_1) ratio or the proportion of 4C nuclei to compare the cell cycle activity in specific tissues, at developmental stages, or at various times following the application of cytotoxic compounds.
- Do not confuse a high proportion of 4C nuclei being an effect of an antimitotic drug treatment with cells commencing endoreduplication.
- Use bivariate analysis to distinguish G_0 and G_1 cells, as well as to determine cell cycle traverse rates and phase transition times.

7 | ESTIMATION OF THE EXTENT OF ENDOREDUPLICATION

In many plant species, somatic cells with nuclear DNA contents higher than 4C are present as a consequence of endoreplication.[128] While it has been commonly observed in angiosperm families, algae, and most mosses, it appears to be rare or absent in hornworts, liverworts, ferns, and gymnosperms.[129] During endoreplication, the nuclei undergo endocycles (repeated rounds of DNA synthesis without an intervening mitosis), resulting in endopolyploid cells (4C → 8C → 16C → 32C → 64C, and so on; Figure 8A).[130,131] The presence of cells with different ploidy levels, including endopolyploid ones, in the same organ/tissue is termed polysomaty. Two types of endoreplication have been defined: (i) endomitosis, in which changes occur in chromatin structure comparable to those which take place during mitosis, resulting in a doubling of the number of chromosomes in the cell (for a review see reference 132), and (ii) endoreduplication, in which any mitosis-like stage is omitted (Figure 7). Since endomitosis is rare in plants,[133] only endoreduplication is considered here.

Endoreduplication occurs when differentiating cells that have left the mitotic cycle continue to synthesize DNA during the endocycles, an endocycle consisting of only endo-G and endo-S phases. Endopolyploid cells can constitute up to 80% of the cells in some tissues, and typically occur in specialized cell types such as trichomes, vascular elements, endosperm cells, root hairs, and root parenchyma cells.[130,134]

Endoreduplication is common in large cells (e.g., endosperm cells[135]) and/or for cells within rapidly growing plant tissues (e.g., in the transition zone between the radicle and hypocotyl during germination)[136]. Nevertheless, there are many species in which only 2C and 4C cells are present (e.g., *Plantago asiatica*[137] and *Helianthus annuus*[138]), which probably grow only by mitotic proliferation. The angiosperm families with predominantly polysomatic species and predominantly nonpolysomatic species are listed by Barow and Jovtchev.[139]

Barow and Meister[140] determined that taxonomic position is the major factor determining whether endoreduplication occurs and, if so, the level of endopolyploidy reached. For example, families such as

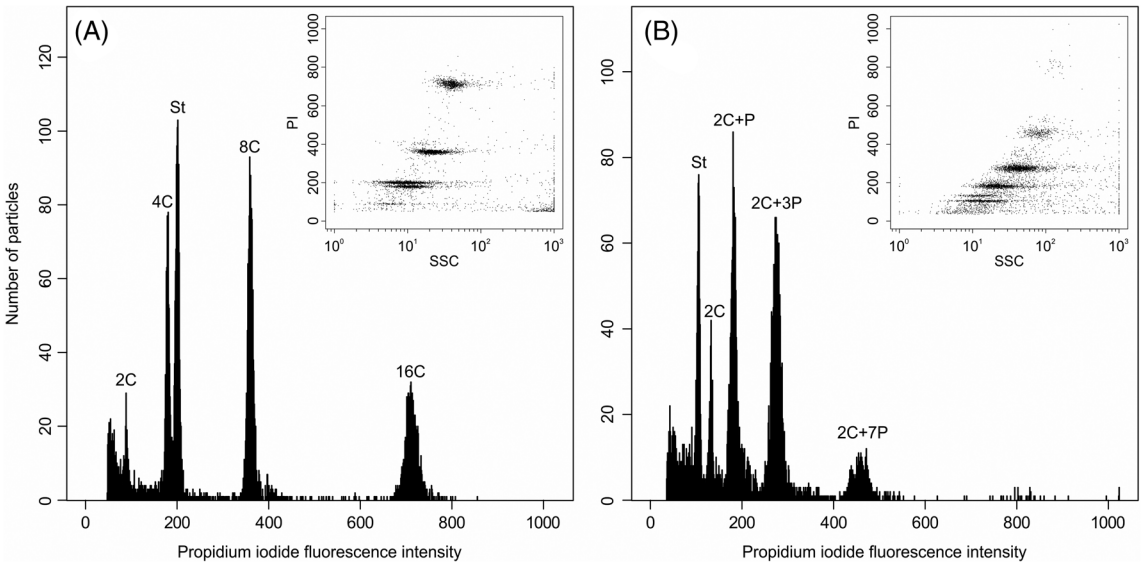

FIGURE 8 Histograms and dot-plots of two orchid species demonstrating two different types of endopolyploidy, endoreduplication (A, *Acianthera hystrix*) and partial endoreplication (B, *Porroglossum schramii*). Whereas nuclei that have undergone endoreduplication exhibit an approximately two-fold increase of fluorescence intensity (A) nuclei that have undergone only partial endoreplication exhibit a substantial departure from this (B). The peaks are labeled according to the convention used for endoreduplication (2C → 4C → 8C → etc.) as well as partial endoreplication (2C → 2C + P → 2C + 3P → etc.). *Solanum pseudocapsicum* was used as the internal standard (labeled as St) in both analyses and its relative genome size is given as unit. The relative peak positions (against the standard) for *A. hystrix* (A) are 0.44: 0.90: 1.80: 3.55 and for *P. schramii* (B) are 1.26: 1.73: 2.62: 4.37. The between peak ratios for (A) follows the pattern 2C: 4C = 2.02, 4C: 8C = 2.01 and 8C: 16C = 1.98, and for (B) 2C: 2C + P = 1.38, 2C + P: 2C + 3P = 1.51 and 2C + 3P: 2C + 7P = 1.67

Brassicaceae, Cucurbitaceae, and Amaranthaceae are characterized by high endoreduplication, whereas it appears to be absent in Asteraceae and Liliaceae (see[139] for a summary of family-level endoreduplication patterns). However, the assumption that the degree of endoreduplication is characteristic for a given family does not apply to all families. For example, Fabaceae includes species with high endoreduplication (e.g., *Phaseolus vulgaris* and *Vicia sativa* with endopolyploidy up to 128C and 64C, respectively, in the cotyledons of mature seeds), as well as species with no endoreduplication such as *Sophora japonica*.[140–142] In addition, the degree of endopolyploidy may differ between ecotypes and varieties/lines of the same species such as in the endosperm of *Zea mays*.[143]

7.1 | Factors that contribute to variation in endopolyploidy

The maximum endopolyploidy level varies in different species and cell types, but typically reaches no more than 64C. However, cells with much higher C-values have been reported, including 8192C in the embryo suspensor of *Phaseolus coccineus* (Fabaceae) or even 24,576C in the endosperm haustorium of *Arum maculatum* (Araceae)[130]. Most polysomatic plants exhibit a characteristic pattern of tissue- and organ-specific endoreduplication that is systemic in nature, and endoreduplication is usually absent from the inflorescence and very young leaves.[140,144,145] Endoreduplication is developmentally regulated, being higher in older tissues than younger ones within the same plant.[132,145] It is also found to be cell-type-specific within tissues of the same developmental stage.[146]

Other factors that correlate with the presence and amount of endoreduplication in plants include life history strategy and growth habit, as endoreduplication is observed more frequently in annual and biennial herbs than in perennials, and is usually absent from woody species.[140,141] Environmental factors, such as light, temperature, water supply, or salinity can also influence the extent of endoreduplication.[147,148] In addition, endoreduplication can be affected by certain biotic interactions, including arbuscular mycorrhizal colonization,[149] powdery mildew infection,[150] parasitism by nematodes,[151] pathogen attack (e.g., *Phytophthora*[152]), and herbivory[153].

Given the many factors that can contribute to changes in endoreduplication, these relationships and limitations must be carefully considered when endoreduplication-focused studies are performed. Where comparisons are being made between or within species, it is therefore important to limit external sources of variation between individuals.

A special type of endoreplication is found in orchids, termed 'progressive partial endoreplication'[57,154,155] or 'strict partial endoreplication'.[156] For the sake of simplicity, we prefer to use the term 'partial endoreplication' since it highlights the crucial difference with conventional endoreplication, namely that, during cell differentiation, only part of the genome (usually denoted as 'P') is replicated, and subsequent fractions of nuclei have a genome size equal to 2C + P, 2C + 3P, 2C + 7P, and so forth.[154] This gives rise to FCM histograms

with an increasing ratio between the relative fluorescence of subsequent peaks (Figure 8B) theoretically approaching the ratio 1:2. The size of the replicated fraction of the genome, that is 'P', is species-specific and can vary from approximately 20% to nearly 100%.[57] The study of Hřibová et al.[155] supports the hypothesis that during the S-phase of the cell cycle, only the DNA comprising the 'P' fraction of the genome is replicated, and rejects the alternative hypothesis that the whole genome is replicated during the S phase followed by a post-synthetic elimination of DNA from the genome. The existence of partial endoreplication challenges the estimation of genome size in orchids and highlights the need for the careful selection of appropriate plant tissues.[57]

7.2 | Quantification of endoreduplication using FCM

The investigation of endoreduplication using FCM requires the analysis of histograms containing multiple fluorescent peaks, which correspond to the different ploidy levels within the plant/plant part. Depending on the degree of endoreduplication (the number of endocycles), it is often necessary to use a logarithmic scale to observe all peaks on the same histogram (Figure 8). For regular endoreduplication, the peaks should appear at consistent intervals, with each peak occupying an arithmetical-doubling series of increasing fluorescence corresponding to a doubling of the DNA amount compared to the previous peak. Usually no internal standard is needed (and may be undesired, as for cell cycle analysis; see Section 6). However, in highly endopolyploid material, the 2C peak can be very small or even absent, and so the addition of a reference standard, even if only used for a daily setting of the cytometer, may be useful to locate the 2C peak. Alternatively, young leaves of the sample species can be used as an external standard to define the position of the 2C peak. Using an internal standard is recommended for species with very small genomes such as *Arabidopsis thaliana* (2C = 0.32 pg), where the 2C peak can fall in the range of fluorescing debris or when debris forms a peak of its own and may be mistaken for the 2C peak.[26,139]

The total number of nuclei measured will vary depending on the number of peaks in the histogram; however, it should not be less than 1000 nuclei in the highest peak. Usually, the measurement of fluorescence of 5000–10,000 nuclei is considered adequate. After locating all nuclear peaks in the histogram, the next step is to label them according to ploidy level. It is recognized that most plant tissues/organs with endopolyploid cells will also contain at least some mitotically-active cells and/or cells in the G_0 phase (see Section 6). Evaluation of the FCM histogram requires marking all peaks to determine the number/proportion of nuclei with different DNA contents as well as the number of endocycles. As it is not possible to use FCM to distinguish 4C nuclei that have just entered endoreduplication (i.e., being in the G phase of the first endocycle) from those in the G_2 phase of the mitotic cycle, the occurrence of an 8C peak is needed to provide unequivocal evidence of endoreduplication. Once the peaks have been identified, gating should be applied (if available in the flow cytometer software) to eliminate fluorescent signals from debris, and the total number of events in each peak should be recorded.

There are various metrics to describe endopolyploidy using the information collected from FCM histograms:

- The number of endocycles, established based on the number of peaks corresponding to nuclei with DNA contents >4C.
- The percentage of nuclei corresponding to each ploidy level.
- The mean C-value (mean ploidy) according to the equation[157]:

$$Mean\ C\text{-}value = \sum_{i=1}^{n} \frac{C_i \times N_i}{N_{sample}},$$

where n, number of peaks with different DNA contents of the sample; C_i, C-value of the nuclei of the peak n_i; N_i, number of nuclei in the peak n_i; N_{sample}, number of nuclei in all the peaks of the sample.

- Cycle value (also called endoreduplication index, EI), reports the mean number of endocycles per nucleus, according to the equation[140]:

$$Cycle\ value = ((n_{2C} \times 0) + (n_{4C} \times 1) + (n_{8C} \times 2) + (n_{16C} \times 3) + \ldots) \div (n_{2C} + n_{4C} + n_{8C} + n_{16C} + \ldots),$$

where n, number of counts per given C-value content.

This value is used to compare the degree of endoreduplication between species. According to Barow and Meister,[140] a cycle value lower than 0.1 is indicative of a lack of endoreduplication; the authors identified this threshold by determining the proportion of doublet nuclei in a series of plants. This threshold also accounts for nuclei in G_2.

- Super cycle value (*SCV*), where 8C nuclei are considered as the first endoreduplicated level[158]:

$$SCV = ((n_{2C} \times 0) + (n_{4C} \times 0) + (n_{8C} \times 1) + (n_{16C} \times 2) + \ldots) \div (n_{2C} + n_{4C} + n_{8C} + n_{16C} + \ldots),$$

where n, number of counts per given C-value content.

- The $\sum$ >2C/2C ratio (i.e., the ratio between the number of all nuclei with a DNA content higher than the 2C amount and the number of nuclei possessing a 2C DNA content[142]) is most often used to describe the dynamics of endoreduplication changes over time, such as during seed development or germination.

Depending on the aim of the study, a combination of different measures can be used to describe the degree of endoreduplication in a plant or its part/organ.

7.3 | Best practices

- Sample preparation and measurement should follow the rules for ploidy estimation (see Section 2). However, including an internal

standard is generally not recommended except when analyzing highly endopolyploid material, where the 2C peak can be very small or absent. Alternatively, external standardization can be used for such material to define the 2C peak position on the histogram.

- When preparing plant material for analysis of endoreduplication, the influence of environmental factors (especially light) as well as organ/tissue age on the level of endopolyploidy should be considered.
- The number of analyzed nuclei should be increased to 5000–10,000 for highly endopolyploid material.
- The logarithmic scale should be used to ensure visualization of all peaks in the histogram and hence ensure higher endopolyploidy levels are not missed.
- Be aware that 4C nuclei at the G_2 phase of the cell cycle cannot be distinguished from those in the first G phase of the endocycle.
- Be aware that the 2C peak may contain very few nuclei (or even be absent) and therefore can be easily overlooked.
- Be aware of the occurrence of partial endoreplication in some orchids.

8 | FLOW CYTOMETRIC SEED SCREENING FOR EXPLORING REPRODUCTIVE PATHWAYS

Flow cytometric seed screening (FCSS) was originally developed to provide a simple and efficient method to recognize different reproductive pathways reported in angiosperms.[159] The method is based on using FCM to distinguish between different ploidies of the embryo and endosperm nuclei that are present in seeds. It has been widely used, for example, in plant breeding, where FCSS allows identification of apomictic cultivars that deviate from the normal sexual reproduction pathway. In evolutionary studies, it can be applied to address questions concerning the frequency of apomixis in natural populations[160–162] or to search for inter-ploidy hybrids in polyploid complexes.[163–165]

Although there are a few exceptions,[166,167] the vast majority of sexually reproducing angiosperms produce seeds which have a diploid ($2n$) embryo and triploid ($3n$) endosperm. These arise from the production of reduced (n) gametes (i.e., sperm and egg cells) via meiosis, followed by a double fertilization involving: (i) fertilization of the haploid egg cell (n) with one of the haploid sperm nuclei (n) of the mature pollen grain to form the diploid ($2n$) embryo and (ii) fusion of the second sperm nucleus (n) with the two haploid polar nuclei ($n + n$) in the central cell of the embryo sac to form the triploid ($3n$) endosperm.

8.1 | Identifying seeds originating from sexual reproduction

Seeds that have arisen from sexual reproduction can typically be identified by analyzing mature seeds with standard FCM methods. The FCM histogram contains a 2C peak (originating from G_0/G_1 nuclei of the embryo) and a 3C peak (derived from G_0/G_1 nuclei of the endosperm) (Figure 9A, B). If additional peaks are present, they typically contain less nuclei than the 2C and 3C ones, corresponding to nuclei from the embryo and/or the endosperm that are either in the G_2 phase of the cell cycle (e.g., 4C and 6C, respectively; Figure 9A) or have become endopolyploid (see Section 7). If endopolyploidy is present, this may require viewing the peaks using the logarithmic scale.

Deviations from this 2C:3C (embryo: endosperm) peak ratio can occur when fertilization is effected by unreduced ($2n$) gametes, which are taxonomically widespread in angiosperms,[168] or is due to heteroploid hybridization. If an unreduced sperm cell fertilizes a reduced egg cell, this will give rise to a FCM histogram with a 3C peak corresponding to nuclei originating from the fusion of the $2n$ sperm with the reduced egg cell (n), and a 4C peak from the fusion of the two reduced polar nuclei with the unreduced sperm nucleus (i.e., $n + n + 2n$). Alternatively, if it is the embryo sac that is unreduced, this will give rise to a 3C peak corresponding to embryo nuclei arising from fertilization of the unreduced egg cell with a reduced sperm nucleus (i.e., $2n + n$), and a 5C peak comprising endosperm nuclei derived from the fusion of two unreduced polar nuclei and one reduced sperm nucleus (i.e., $2n + 2n + n$). Similarly, hybridization between sexual diploids and tetraploids with reduced gametes leads to 3C:4C or 3C:5C embryo: endosperm ploidy ratios for when the diploid is the ovule- or pollen-donor, respectively. If other ploidy levels and/or unreduced gametes are involved, other ratios are possible, but they always follow the $f + m$ (embryo) and $2f + m$ (endosperm) rule (where f and m are the female and male gamete ploidy levels, respectively). For analysis of such complex data, knowledge of the ploidy of parental plants is essential.

8.2 | Identifying reproductive pathways in seeds from asexual reproduction

Despite the prevalence of sexual reproduction in plants, seeds can also be formed from a variety of asexual (apomictic) pathways, with some species being obligate apomicts whereas others are facultative.[159] At least 10 different pathways of seed formation via apomixis can be reconstructed by analyzing the ratio between peaks observed in the FCM histogram. The occurrence of nuclei with different ploidies depends on whether: (i) the embryo sac is reduced or unreduced, (ii) the sperm nuclei are reduced or unreduced, and (iii) the egg and/or polar nuclei are fertilized.

Gametophytic apomixis is characterized by three components[169]: (i) the formation of an unreduced embryo sac from cells which have not undergone complete meiosis (= apomeiosis), either from sporophytic cells such as the nucellus (= apospory) or from megaspore mother cells, which have not undergone full meiotic reduction (= diplospory), (ii) the development of an unreduced embryo without fertilization (parthenogenesis), (iii) the development of functional endosperm (i.e., autonomous endosperm formation or pseudogamy).

These components can occur in different combinations and can be detected by determining the ratio between the peaks in the FCM histograms, which will differ from the typical sexual embryo to endosperm ratio (i.e., 2C:3C) found in most angiosperms (see above; Figure 9). The ratio is calculated by determining the peak index, defined as the DNA

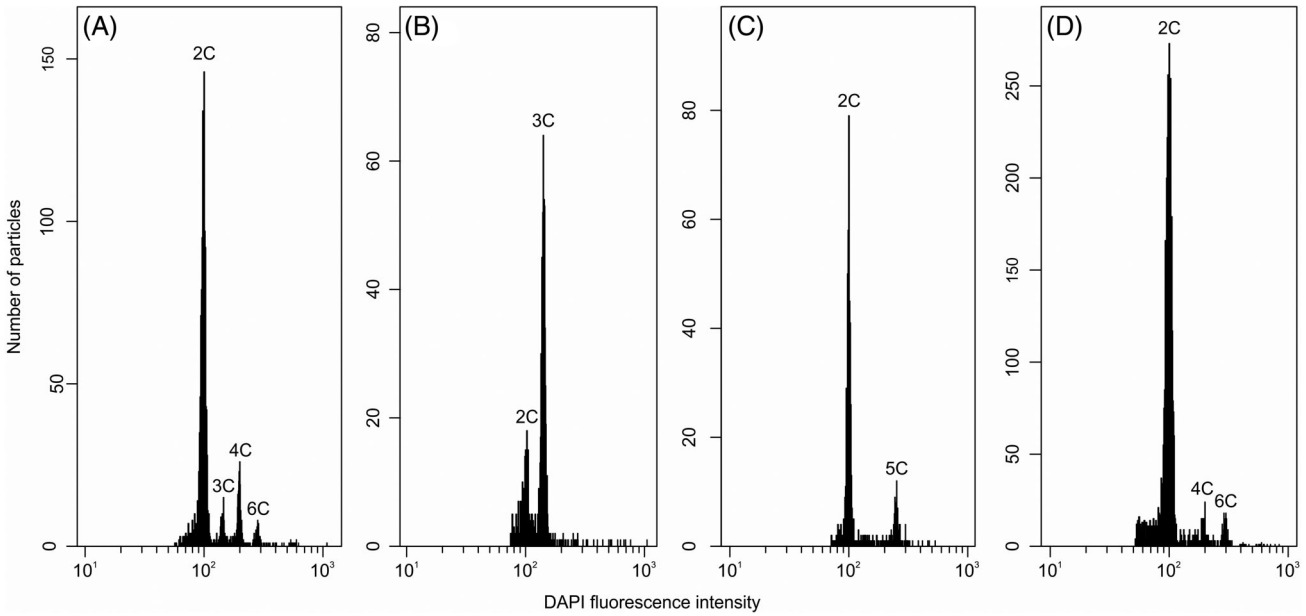

FIGURE 9 Histograms of nuclei isolated from seeds of four species demonstrating their origin from sexual (A and B) and asexual (C and D) reproduction. A, *Beta vulgaris*; B, *Anethum graveolens*; C, *Poa pratensis*; D, *Boechera stricta*. Different embryo to endosperm ploidy ratios are demonstrated, which depend on the reproduction pathway: (i) 2C:3C for sexual reproduction (A and B; peaks 4C and 6C in (A) correspond to nuclei either in G_2 or entering the first endocycle in the embryo and endosperm, respectively, and the small unmarked peaks with higher DNA content correspond to endopolyploid nuclei: 8C in the embryo and 12C in the endosperm); (ii) 2C:5C for apomictic reproduction, when an unreduced embryo has not been fertilized and the endosperm has been formed by the fusion of the two unreduced polar nuclei and a reduced sperm nucleus (C); (iii) 2C:6C for apomictic reproduction, when the embryo has developed from an unreduced and unfertilized egg cell, and was formed by the fusion of the unreduced polar nuclei and an unreduced sperm nucleus (D), the small 4C peak corresponds to the G_2 nuclei of the embryo). The relative peak positions are 1.00: 1.47: 2.01: 2.91 (A), 1.00: 1.46 (B), 1.00: 2.52 (C), 1.00: 1.97: 2.97 (D)

content of the embryo nuclei relative to that of endosperm nuclei. For example, the presence of peaks that occur in the ratio of 2C:6C (Figure 9D) is typical of a seed that has arisen from apomeiosis followed by parthenogenesis; the 2C nuclei have arisen from an embryo that has developed from an unreduced and unfertilized egg cell while the 6C peak corresponds to endosperm nuclei from the fusion of unreduced polar nuclei and an unreduced sperm nucleus (but see below). The other frequent ratios are 2C:5C (unreduced unfertilized embryo and fusion of the two unreduced polar nuclei and a reduced sperm nucleus; Figure 9C) and 2C:4C (unreduced unfertilized embryo and autonomous endosperm development, that is, fusion of the two unreduced polar nuclei without contribution of pollen).

8.3 | Potential pitfalls when using FCSS

While FCSS has been used successfully to explore the diversity and evolution of reproductive pathways in many plant lineages,[170–173] users should be aware of the following potential pitfalls:

8.3.1 | Analyzing sexually reproducing angiosperms which deviate from the 2C:3C embryo:endosperm ratio

As noted above, there are a few angiosperm lineages whose female gametophyte development deviates from the typical '*Polygonum*'-type, and hence give rise to seeds, which deviate from the 2C:3C ratio despite undergoing normal sexual reproduction. These include: (i) species in some of the early diverging angiosperm lineages such Nymphaeales and Austrobaileyales, which produce diploid endosperms, (ii) species within the eudicots and monocots such as those in *Penaea* (Penaeaceae, Myrtales), *Plumbagella* (Plumbaginaceae, Caryophyllales), and *Fritillaria* (Liliaceae, Liliales), which produce pentaploid endosperms, and (iii) some *Peperomia* species (Piperaceae, within the magnoliids) which produce endosperms of even higher ploidy.[166,174] FCM users should therefore be aware of these potential issues and are encouraged to thoroughly check the literature prior to embarking on FCSS to ensure that the flow histograms are correctly interpreted.

8.3.2 | Analyzing species with endopolyploid endosperm

Some species may contain endosperm cells that have undergone endoreduplication (see Section 7). In such cases the 3C peak may be much smaller than the successive peaks (i.e., 6C, 12C, or higher), or even almost undetectable, as seen for example, in studies of *Leptorhabdos* and *Pedicularis* (Orobanchaceae).[175] If endopolyploid cells are present in the endosperm, users should take particular care in analyzing the FCM histogram to avoid misidentifying peaks. It may be advantageous to adjust the flow cytometer settings to allow a wider

range of fluorescence intensities to be displayed (i.e., including additional peaks representing higher ploidy levels). This may be achieved either by locating the embryo peak near the left-hand side limit of the fluorescence axis (i.e., the X-axis) or using a logarithmic scale (preferable).

8.3.3 | Analyzing mature seeds

The prerequisite for effective FCSS is being able to access a sufficient number of endosperm nuclei in the seed. However, for some species, there is only a narrow developmental window when both embryo and endosperm nuclei are present and clearly distinguishable in FCM histograms (Suda et al., pers. comm., reference[176]). After this, only embryo nuclei can be detected by FCM, and hence insights into reproductive pathways cannot be determined. For example, Brožová et al.[175] were unable to detect the endosperm peak in 37 out of 257 species (ca. 15%) of the Himalayan flora.

A failure to detect endosperm nuclei can arise especially when analyzing some species with starchy endosperm (including cereals), as endosperm cells typically undergo programmed cell death (PCD) after the phase of accumulating storage compounds is completed. Once this has happened, endosperm nuclei cannot be detected by FCM. This is further complicated by the observation that the pathway of PCD progression during seed development varies between species (e.g., in maize, PCD starts ca. 16 days after pollination (DAP) and is completed by 40 DAP, whereas in wheat, while PCD starts at a similar time as maize, it spreads randomly throughout the embryo by 30 DAP[135]). Successful FCSS analysis of starchy seeds therefore depends on using immature seeds that still contain endosperm cells (see below). This introduces a developmental monitoring component to the protocol.

Typically, in mature seeds the proportion of endosperm nuclei is much lower than that of the embryo nuclei (e.g., Figure 9A). Hence, in addition to the impact of the age of the seed, there is a risk that endosperm nuclei will not be detected. For example, in some species of Rosaceae, the endosperm is only present as a thin layer just inside the seed coat, thus care must be taken not to exclude this layer during FCSS sample preparation, particularly if, prior to chopping, the seed coat is removed. In such endosperm-poor species, users should consider exploring the option of increasing the total number of nuclei analyzed to allow the endosperm nuclei to produce a detectable peak.

In contrast, the opposite has been reported for some species, that is, the endosperm peak is much larger than the embryo peak. Examples of this include *Anethum graveolens* (Apiaceae, Figure 9B),[177] and more generally in some families (e.g., Apiaceae, Papaveraceae, and Ranunculaceae[175]). Given these observations, comparing the position of the 2C peak in leaf tissue with the putative 2C embryo peak obtained from analyzing seeds is strongly recommended when analyzing a new species for the first time. However, researchers should be aware that the sample/standard ratio may be slightly different between the seed and leaf tissue due to different levels of DNA condensation, and, consequently, DNA accessibility[178] and/or different amounts of staining inhibitors.[38] Usually, external standardization (i.e., analyzing leaf tissue and seeds separately with the same settings of the cytometer) is sufficient, but should be used with caution as slightly different fluorescence intensities can be expected, as noted above. However, if possible, internal standardization (see Section 3) is also recommended.

8.3.4 | Analyzing immature and aborted seeds

As for mature seeds, the analysis of immature seeds can lead to misinterpretation if care is not taken. For example, the seed coat of immature seeds, which typically dies during maturation, can still contain cells from the maternal parent having 2C nuclei. These nuclei might be confused with 2C nuclei from the embryo or could give rise to an additional peak if the embryo has a different ploidy from the maternal parent. Approaches to obtain reliable results in such cases involve dissecting out the embryo and endosperm of each seed, running each separately (to correctly identify each peak), then together (to determine the embryo:endosperm ratio). This procedure is rather time-consuming and requires considerable skill and hence may simply be impractical, especially when one considers that with a typical high-throughput FCSS, 96 samples can be prepared and analyzed using an auto-sampler within just 3 h.[169]

Embryos may abort at a stage in development where significant quantities of endosperm have already developed, potentially resulting in partially filled seeds containing mostly endosperm and a very small, degraded embryo. In such cases the endosperm peak may be interpreted as the embryo peak. Caution should therefore be taken when the sample comprises seeds that vary in size as this might indicate the presence of late aborted seed. Removing the seed coat or dissecting a seed can help to identify this situation, since a seed filled with endosperm is quite different from one containing a mature embryo. As long as these are correctly identified, and if the reproductive pathway can be inferred from endosperm ploidy alone, aborted seeds may provide valuable information on reproductive pathways associated with seed abortion.

8.3.5 | Analyzing species which deviate from regular gametophyte development or produce unreduced pollen

Deviations from the regular female gametophyte development may occur at a low but constant frequency within an individual, such as one extra nucleus contributing to the endosperm,[179,180] which will increase the endosperm ploidy by 1C or 2C (reduced or unreduced embryo sac, respectively). In contrast, the presence of only one polar nucleus can decrease the endosperm ploidy, and lead to a sexual-like peak index of 1.5 (i.e., 2C:3C) even in apomictic seeds (i.e., parthenogenesis with pseudogamous endosperm development in an unreduced embryo sac), as reported in some Poaceae[181] and suspected in *Potentilla*[182].

Different numbers of polar and sperm nuclei may yield the same peak index.[182] Most importantly, recognizing the contribution of unreduced pollen from FCSS is not straightforward. For example, if both reduced sperm nuclei fuse with the unreduced polar nuclei this will give rise to the same ploidy level of the endosperm (i.e., (C + C) + (2C + 2C) = 6C) as would be found if fertilization involved a single unreduced sperm fusing with two unreduced polar nuclei (i.e., 2C + (2C + 2C) = 6C). Available data from species in the well-studied Rosaceae family show that the frequency of both sperm nuclei contributing to the endosperm is species- or genus-specific, ranging from moderate, as in, for example, *Crataegus*[179] to nearly exclusive, as found in *Sorbus*[180].

8.3.6 | Analyzing polyploid complexes

In polyploid complexes involving natural mixed populations of several species and several ploidy levels (which is the rule rather than the exception in genera rich in apomictic micro-species), different combinations of parents and different seed formation pathways can lead to the same peak index value that cannot be distinguished using FCSS alone.

In studies involving multiple species, or when variation in the reproductive system is expected, as in facultative apomicts, it is helpful to record the genome sizes (either as absolute values or as a ratio to a fixed standard) of the embryo and endosperm instead of just the peak index. It is often possible to use the leaf of an internal standard after some optimization of the sample preparation protocol (e.g., seed nuclei, as obtained from dry cells, should be incubated in a buffer for 15–30 min prior to FCM analysis, while a fresh leaf sample can be analyzed directly after preparation). In addition to estimating the ploidy of the embryo, the data enable calculation of the nuclear DNA content of the gametes under a given mode of fertilization.[180] However, researchers should be aware that part of the observed variation may be attributable to a lower precision in recording the endosperm peak position (due to the lower number of nuclei in the peak).

Despite such words of caution, the FCSS approach, when used with appropriate care and attention, has been shown to be widely applicable for analyzing a diversity of species, regardless of the size and constitution of the seeds, and hence for addressing a wide range of biological questions.[170,172,182] FCSS can be applied to analyze single seeds (e.g., as demonstrated for *Poa pratensis*, *Triticum aestivum*, *Boechera stricta*, and *Z. mays*), or bulked samples comprising up to 50 seeds in a single sample.[159,173] However, bulking of seeds, which is usually done if seeds are very small, should be carried out with caution if there is a possibility that a small admixture of seeds with different reproductive pathways exists. Certainly, the peak corresponding to a single or a few seeds in a bulk of 50 seeds can be easily missed. Recent developments in the application of FCSS have included optimization of the method to enable screening of reproductive pathways in two asymmetrically compensating allopolyploid species *Onosma arenaria* and *Rosa canina*, highlighting the precision of FCSS in being able to dissect different reproductive modes in complex biological systems.[183]

8.4 | Best practices

- For screening of seeds by FCM, it is essential that their biology and stage of development is known to avoid misidentifying the embryo and endosperm peaks in the fluorescence histogram.
- Sample preparation and measurement should follow basic rules for ploidy estimation (see Section 2).
- Use a reference standard, at least at the start of an experiment, to calibrate peak positions and ensure the position of the 2C peak in the fluorescence histogram is correctly identified, using internal standardization whenever possible.
- If the seed tissues are polysomatic (containing endopolyploid cells), the flow cytometer settings should be adjusted to ensure all peaks are detected. This can either be achieved by locating the position of the embryo peak near the left-hand side limit of the fluorescence axis (making sure that the 2C peak is not lost among signal from debris and is fully recorded, that is, there are a few 'empty' channels before the peak), or by using a logarithmic scale (preferable).
- For the analysis of starchy seeds, FCSS should be performed on immature seeds before the endosperm undergoes PCD.
- If using immature seeds, users should be aware of the potential presence of nuclei from the 2C maternal tissue, which could lead to the misinterpretation of peaks in the flow histogram.
- If the proportion of nuclei between the embryo and endosperm is skewed and hence one of the peaks contains a much lower number of nuclei than the other (e.g., as seen when analyzing the embryo peak of a species with endosperm-rich seeds), researchers should be aware of the lower precision in recording its position, and hence the lower accuracy of the calculated peak index or genome size.

9 | ESTIMATION OF GENOME SIZE STABILITY IN PLANT MATERIAL CULTURED IN VITRO

Plant tissue cultures (in vitro cultures) are frequently employed in both fundamental and applied research as well as in agricultural/horticultural production.[1] For example, they are used for plantlet production, polyploid cultivar breeding, and the generation of (doubled) haploid plants. In addition, in the medical field they are used, for example, for the production of bioactive compounds and for micropropagating rare or endangered plants of medicinal value. Finally, in genetic engineering, regeneration of euploid plants following transformation is a critical step in integrating desired transgenes into breeding material.

One of the major problems limiting the use of in vitro cultures is the occurrence of somaclonal variation (i.e., genomic variation originating during cell and tissue culture), which changes the genotype of the produced material. It arises from the impact of mutations, chromosome rearrangements and aberrations, transposable element activity, gene silencing, and/or polyploidization, all of which can be triggered by the stressful conditions experienced during in vitro culture.[184]

These processes lead to the generation of individuals (= somaclonal variants) that differ from the plants used as the starting material. While such changes are usually not desirable and somaclonal variants should be routinely eliminated, in some particular cases, useful new variants can also be generated.[185,186] For example, some somaclones of wheat were shown to display improved yield traits and resistance to leaf rust and spot blotch diseases.[186] Moreover, somaclonal variation generated by in vitro culture conditions may have potential application in ex situ conservation programmes of rare plant species as it can provide a novel source of genetic diversity.[187] The commercial implications of eliminating noneuploid individuals from greenhouse populations of commercial crops at the earliest possible stage are profound.

The extent of somaclonal variation generated during in vitro culture has been shown to depend on a variety of factors, including: (i) the type of starting material, (ii) the method of propagation in vitro, (iii) the type and concentration of plant growth regulators (PGRs) added to the growth medium, (iv) the number of subcultures and their duration, and (v) the degree of resistance of specific genotypes to stress caused by culturing in vitro.[184] In addition, it should be noted that culturing in vitro can induce and/or reveal variation between cells, tissues, and organs. Thus, variation can occur within cultures or between somaclones.

Heritable changes arising from somaclonal variation can be detected by cytological and molecular methods.[184] Since stable genome size is one of the indicators that is often used to assess the genetic stability of plant material produced in vitro, FCM is commonly used for detecting quantitative genomic changes generated by tissue culture. Generally, the procedures recommended for the estimation of genome size (see Sections 3 and 4) or ploidy (see Section 2), depending on the type of genomic variation expected, should be followed. Nevertheless, the following specific recommendations are suggested to help obtain in vitro plant material with a stable nuclear DNA content:

9.1 | Choice of starting material

Theoretically, almost any type of plant material can be used as a source of plantlets (explants) (e.g., meristems, leaves, stems, roots, hypocotyls, protoplasts, or cell suspensions). However, the presence of endopolyploid cells (see Section 7) in the starting material will increase the likelihood that regenerants will have nondesirable (usually increased) DNA content/ploidy. Given these observations, it is recommended that prior to selecting which material to use as the explant for in vitro culture, the available choices of starting materials should be analyzed by FCM to determine the extent of endopolyploidy. The most suitable material for minimizing somaclonal variation should contain a majority of cells in the G_0/G_1 phase of the cell cycle (see Section 6), the proportion of cells in the G_2 phase (i.e., with 4C DNA amount) should be not too high, and, if possible, no cells with a DNA content greater than 4C should be present.

When somatic hybrids are produced via protoplast fusion, one should be aware that, in addition to the desired heterokaryons, homokaryons and/or the products of multiple fusions can also be generated - typically these plants can be distinguished based on their genome size and removed.

9.2 | Indirect plantlet regeneration versus direct somatic embryogenesis

DNA content instability is especially likely to occur when plantlets have been regenerated from callus. In contrast, the direct formation of somatic embryos from cultured plant tissue minimizes the chance of genome size changes arising from somaclonal variation. Given these observations, it is strongly recommended that in vitro cultured plant material that has been generated by methods involving an intermediate callus stage should be analyzed by FCM to check that the ploidy is maintained.

9.3 | Composition of growth medium

PGRs, such as auxins and cytokinins, are usually incorporated into the growth medium to induce morphogenesis during in vitro propagation. However, in many cases, they have been shown to increase the genetic variability of plant material growing on such media.[184,185] Therefore, if PGRs have been used, especially at high concentrations, the genome size of regenerants should be checked by FCM.

9.4 | The number and duration of subcultures in vitro

The rate of production of somaclonal variants is increased by both a longer duration of culturing and an increased number of subcultures. This especially applies to cell suspension and callus cultures. FCM analysis of such cultures can be used to provide information on their totipotency based on the number of G_0/G_1 cells present (i.e., if G_0/G_1 cells are not present or their proportion is low, then the organogenic potential of such cultures is usually low).

9.5 | Detecting genome size changes in plant material originating from tissue culture

Generally, the chopping method can be applied for all kinds of in vitro-derived plant material. Explants, regenerants, and calli do not need any special preparation and can be used directly for sample preparation. Cells from a suspension must first be isolated, for example, by filtration, to prevent dilution of the FCM nuclei isolation buffer. For the estimation of DNA content in protoplasts, see the recent review by Antoniadi et al.[188]

Ideally, the source material (usually originating from the field or greenhouse) should be used as the control (without somaclonal variation) for comparison with the micropropagated material. However, if this is not possible, then seedlings derived from the seeds collected from the source material can also be used as a control. It is also recommended that the genome size of the source material is checked before starting in vitro culture to make sure that the correct material is being used for micropropagation.

If the micropropagated species contains staining inhibitors,[38] then it is important to select plant parts that are free of secondary metabolites (e.g., very young leaves) for FCM analysis. If this is not possible, at least organs of the same age from control and micropropagated material should be compared, as the concentration of secondary metabolites can depend on the organ age and growth conditions. Indeed, if inhibitors are suspected, then their presence should be tested[38,60,81,189] and appropriate steps to overcome these issues taken. If this is not done, there is a risk that differences in the FCM-estimated DNA content may be misinterpreted as genuine somaclonal variation rather than arising from a technical artifact.

9.6 | Best practices

- The best practice rules for ploidy and genome size estimation (see Sections 2, 3, and 4) also apply for the detection of genome size instability in tissue cultures.
- Polysomatic explants should be avoided for generating plantlets; the most suitable is a tissue/plant part where most cells are in the G_0/G_1 phase of the cell cycle. Therefore, FCM analysis of explants is recommended before starting tissue culture.
- Ploidy/genome size estimation is strongly recommended especially (i) when using plant material that has been through a callus stage or obtained from protoplast fusion, and (ii) when PGRs have been added to the growth medium.
- Source plant material originating from the field or greenhouse should be used as a control without somaclonal variation.

10 | SPECIFIC INFORMATION THAT SHOULD BE INCLUDED IN FCM PAPERS

- Description of plant material origin and growth/cultivation conditions as well as developmental stage and plant tissue/part used for FCM sample preparation.
- Description of experimental design, especially the number of biological and/or technical replications, measurement randomization method used, if appropriate (i.e., random order of measurements if samples from several groups to be compared are analyzed, and/or when repeated analyses of the same individual are carried out), number of nuclei analyzed in a sample, and method of standardization (internal or external).
- Information on species and cultivar/variety/line/clone of the standard used to convert relative fluorescence into absolute DNA amounts, together with its assumed genome size and the reference for this.
- Information on the location of the herbarium voucher if prepared, to allow the identification of the measured plant to be checked.
- Details of the protocol used for sample preparation, including method of nuclei isolation, buffer composition, applied fluorochrome, sample incubation time.
- Internet location and confirmed availability of archived primary cytometric data sets and associated metadata.
- Name/producer of flow cytometer(s) as well as light source(s) used to make the measurements.
- Statistical measures/variation of the measurements.
- Representative histogram(s), gating procedure, and the range of peak CV values.
- Sample/standard species fluorescence ratios when absolute genome size is reported in absolute units.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

AUTHOR CONTRIBUTIONS

Elwira Sliwinska: Conceptualization (equal); writing – original draft (lead); writing – review and editing (lead). **João Loureiro:** Conceptualization (equal); writing – original draft (equal); writing – review and editing (equal). **Ilia Leitch:** Writing – original draft (equal); writing – review and editing (equal). **Petr Šmarda:** Writing – original draft (equal); writing – review and editing (equal). **Jillian Bainard:** Writing – original draft (supporting); writing – review and editing (supporting). **Petr Bureš:** Writing – original draft (supporting); writing – review and editing (supporting). **Zuzana Chumová:** Writing – original draft (supporting); writing – review and editing (supporting). **Lucie Horová:** Writing – original draft (supporting); writing – review and editing (supporting). **Petr Koutecký:** Writing – original draft (supporting); writing – review and editing (supporting). **Magdalena Lučanová:** Writing – original draft (supporting); writing – review and editing (supporting). **Pavel Trávníček:** Writing – original draft (supporting); writing – review and editing (supporting). **David Galbraith:** Conceptualization (equal); writing – original draft (supporting); writing – review and editing (equal).

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/cyto.a.24499.

ORCID

Elwira Sliwinska https://orcid.org/0000-0001-6597-5309
João Loureiro https://orcid.org/0000-0002-9068-3954
Ilia J. Leitch https://orcid.org/0000-0002-3837-8186
Jillian Bainard https://orcid.org/0000-0001-6501-361X
Petr Bureš https://orcid.org/0000-0002-0146-956X
Zuzana Chumová https://orcid.org/0000-0002-6865-6183
Petr Koutecký https://orcid.org/0000-0002-3455-850X
Magdalena Lučanová https://orcid.org/0000-0002-4612-3693

Pavel Trávníček https://orcid.org/0000-0002-5806-1645

David W. Galbraith https://orcid.org/0000-0003-4020-1635

REFERENCES

1. Sliwinska E. Flow cytometry – a modern method for exploring genome size and nuclear DNA synthesis in horticultural and medicinal plant species. Folia Horticulturae. 2018;30:103–28.
2. Keeler K, Kwankin B, Barnes P, Galbraith D. Polyploid polymorphism in *Andropogon gerardii* Vitman (Poaceae). Genome. 2011;29:374–9.
3. Hanušová K, Čertner M, Urfus T, Koutecký P, Košnar J, Rothfels CJ, et al. Widespread co-occurrence of multiple ploidy levels in fragile ferns (*Cystopteris fragilis* complex; Cystopteridaceae) probably stems from similar ecology of cytotypes, their efficient dispersal and interploidy hybridization. Ann Bot. 2019;123:845–55.
4. Chumová Z, Krejčíková J, Mandáková T, Suda J, Trávníček P. Evolutionary and taxonomic implications of variation in nuclear genome size: lesson from the grass genus *Anthoxanthum* (Poaceae). PLoS ONE. 2015;10:e0133748.
5. Zozomová-Lihová J, Malánová-Krásná I, Vít P, Urfus T, Senko D, Svitok M, et al. Cytotype distribution patterns, ecological differentiation, and genetic structure in a diploid–tetraploid contact zone of *Cardamine Amara*. Am J Bot. 2015;102:1380–95.
6. Sabara HA, Kron P, Husband BC. Cytotype coexistence leads to triploid hybrid production in a diploid–tetraploid contact zone of *Chamerion angustifolium* (Onagraceae). Am J Bot. 2013;100:962–70.
7. Castro M, Castro S, Figueiredo A, Husband B, Loureiro J. Complex cytogeographical patterns reveal a dynamic tetraploid–octoploid contact zone. AoB PLANTS. 2018;10:ply012.
8. Čertner M, Fenclová E, Kúr P, Kolář F, Koutecký P, Krahulcová A, et al. Evolutionary dynamics of mixed-ploidy populations in an annual herb: dispersal, local persistence and recurrent origins of polyploids. Ann Bot. 2017;120:303–15.
9. Krejčíková J, Sudová R, Lučanová M, Trávníček P, Urfus T, Vít P, et al. High ploidy diversity and distinct patterns of cytotype distribution in a widespread species of *oxalis* in the greater cape floristic region. Ann Bot. 2013;111:641–9.
10. Trávníček P, Jersáková J, Kubátová B, Krejčíková J, Bateman RM, Lučanová M, et al. Minority cytotypes in European populations of the *Gymnadenia conopsea* complex (Orchidaceae) greatly increase intraspecific and intrapopulation diversity. Ann Bot. 2012;110:977–86.
11. Vít P, Douda J, Krak K, Havrdová A, Mandák B. Two new polyploid species closely related to *Alnus glutinosa* in Europe and North Africa – an analysis based on morphometry, karyology, flow cytometry and microsatellites. Taxon. 2017;66:567–83.
12. Koutecký P, Štěpánek J, Baďurová T. Differentiation between diploid and tetraploid *Centaurea phrygia*: mating barriers, morphology and geographic distribution Preslia. 2012;84:1–32.
13. Prančl J, Koutecký P, Trávníček P, Jarolímová V, Lučanová M, Koutecká E, et al. Cytotype variation, cryptic diversity and hybridization in *ranunculus* sect. *Batrachium* revealed by flow cytometry and chromosome numbers. Preslia. 2018;90:195–223.
14. Te Beest M, Le Roux JJ, Richardson DM, Brysting AK, Suda J, Kubešová M, et al. The more the better? The role of polyploidy in facilitating plant invasions. Ann Bot. 2012;109:19–45.
15. Nagy DU, Stranczinger S, Godi A, Weisz A, Rosche C, Suda J, et al. Does higher ploidy level increase the risk of invasion? A case study with two geo-cytotypes of *Solidago gigantea* Aiton (Asteraceae). J Plant Ecol. 2018;11:317–27.
16. Pandit MK, White SM, Pocock MJO. The contrasting effects of genome size, chromosome number and ploidy level on plant invasiveness: a global analysis. New Phytol. 2014;203:697–703.
17. Čertner M, Sudová R, Weiser M, Suda J, Kolář F. Ploidy-altered phenotype interacts with local environment and may enhance polyploid establishment in *Knautia serpentinicola* (Caprifoliaceae). New Phytol. 2019;221:1117–27.
18. Rey PJ, Manzaneda AJ, Alcántara JM. The interplay between aridity and competition determines colonization ability, exclusion and ecological segregation in the heteroploid *Brachypodium distachyon* species complex. New Phytol. 2017;215:85–96.
19. Pavlíková Z, Holá D, Vlasáková B, Procházka T, Münzbergová Z. Physiological and fitness differences between cytotypes vary with stress in a grassland perennial herb. PLoS ONE. 2017;12:e0188795.
20. Maherali H, Walden AE, Husband BC. Genome duplication and the evolution of physiological responses to water stress. New Phytol. 2009;184:721–31.
21. Šmarda P, Knápek O, Březinová A, Horová L, Grulich V, Danihelka J, et al. Genome sizes and genomic guanine+cytosine (GC) contents of the Czech vascular flora with new estimates for 1700 species. Preslia. 2019;91:117–42.
22. Čertner M, Lučanová M, Sliwinska E, Kolář F, Loureiro J. Plant material selection, collection, preservation and storage for nuclear DNA content estimation. Cytometry Part A. 2021. https://doi.org/10.1002/cyto.a.24482
23. Lipnerová I, Bureš P, Horová L, Šmarda P. Evolution of genome size in *Carex* (Cyperaceae) in relation to chromosome number and genomic base composition. Ann Bot. 2013;111:79–94.
24. Suda J, Krahulcová A, Trávníček P, Krahulec F. Ploidy level versus DNA ploidy level: an appeal for consistent terminology. Taxon. 2006;55:447–50.
25. Johnson PG, Riordan TP, Arumuganathan K. Ploidy level determinations in Buffalograss clones and populations. Crop Sci. 1998;38:478–82.
26. Galbraith DW. Simultaneous flow cytometric quantification of plant nuclear DNA contents over the full range of described angiosperm 2C values. Cytometry A. 2009;75A:692–8.
27. Bharathan G, Lambert G, Galbraith DW. Nuclear DNA content of monocotyledons and related taxa. Am J Bot. 1994;81:381–6.
28. Temsch EM, Koutecký P, Urfus T, Šmarda P, Doležel J. Reference standards for flow cytometric estimation of absolute nuclear DNA content in plants. Cytometry Part A. 2021;99. https://doi.org/10.1002/cyto.a.24495
29. Galbraith DW, Lambert GM, Macas J, Doležel J. Analysis of nuclear DNA content and ploidy in higher plants. Curr Protoc Cytom. 1997;2:7.6.1–7.6.22.
30. Sliwinska E, Jansen R. Estimation of the share of components of different ploidy in anisoploid sugar beet (*Beta vulgaris* L.) populations by flow cytometry. J Appl Genet. 1997;38:151–60.
31. Trávníček P, Eliášová A, Suda J. The distribution of cytotypes of *Vicia cracca* in Central Europe: the changes that have occurred over the last four decades. Preslia. 2010;82:149–63.
32. Suda J, Trávníček P. Reliable DNA ploidy determination in dehydrated tissues of vascular plants by DAPI flow cytometry - new prospects for plant research. Cytometry A. 2006;69A:273–80.
33. Suda J, Kron P, Husband BC, Trávnícek P. Flow cytometry and ploidy: applications in plant systematics, ecology and evolutionary biology. In: Doležel J, Greilhuber J, Suda J, editors. Flow cytometry with plants cells. Weinheim: Wiley-VCH; 2007. p. 103–30.
34. Leitch IJ, Bennett MD. Genome downsizing in polyploid plants. Biol J Linn Soc. 2004;82:651–63.
35. Šmarda P, Bureš P, Horová L, Foggi B, Rossi G. Genome size and GC content evolution of *Festuca*: ancestral expansion and subsequent reduction. Ann Bot. 2008;101:421–33.
36. Temsch EM, Greilhuber J. Genome size in Dipsacaceae and *Morina longifolia* (Morinaceae). Plant Sys Evol. 2010;289:45–56.
37. Chumová Z, Mandáková T, Trávníček P. Are B-chromosomes responsible for the extraordinary genome size variation in selected *Anthoxanthum* annuals? Plant Sys Evol\. 2016;302:731–8.

38. Loureiro J, Kron P, Temsch EM, Koutecký P, Lopes S, Castro M, et al. Isolation of plant nuclei for estimation of nuclear DNA content: overview and best practices. Cytometry A. 2021;99:318–27.
39. Doležel J, Bartoš J, Voglmayr H, Greilhuber J. Nuclear DNA content and genome size of trout and human. Cytometry A. 2003;51A: 127–8.
40. Bennett MD, Leitch IJ. Nuclear DNA amounts in angiosperms - progress, problems and prospects. Ann Bot. 2005;95:45–90.
41. Greilhuber J, Doležel J, Lysak MA, Bennett MD. The origin, evolution and proposed stabilization of the terms 'Genome size' and 'C-value' to describe nuclear DNA contents. Ann Bot. 2005;95: 255–60.
42. Greilhuber J, Doležel J. 2C or not 2C: a closer look at cell nuclei and their DNA content. Chromosoma. 2009;118:391–400.
43. Pellicer J, Leitch IJ. The plant DNA C-values database (release 7.1): an updated online repository of plant genome size data for comparative studies. New Phytol. 2019;226:301–5.
44. Pellicer J, Hidalgo O, Dodsworth S, Leitch IJ. Genome size diversity and its impact on the evolution of land plants. Genes. 2018;9:88.
45. Loureiro J, Trávníček P, Rauchova J, Urfus T, Vit P, Stech M, et al. The use of flow cytometry in the biosystematics, ecology and population biology of homoploid plants. Preslia. 2010;82:3–21.
46. Greilhuber J, Leitch IJ. Genome size and the phenotype, Physical structure, behaviour and evolution of plant genomes. In: Leitch IJ, Greilhuber J, Doležel J, Wendel JF, editors. Plant Genome Diversity. Volume 2. Wien: Springer-Verlag; 2013. p. 323–44.
47. Simonin KA, Roddy AB. Genome downsizing, physiological novelty, and the global dominance of flowering plants. PLoS Biol. 2018;16: e2003706.
48. Roddy AB, Théroux-Rancourt G, Abbo T, Benedetti JW, Brodersen CR, Castro M, et al. The scaling of genome size and cell size limits maximum rates of photosynthesis with implications for ecological strategies. Int J Plant Sci. 2020;181:75–87.
49. Novák P, Guignard MS, Neumann P, Kelly LJ, Mlinarec J, Koblížková A, et al. Repeat-sequence turnover shifts fundamentally in species with large genomes. Nat Plants. 2020;6:1325–9.
50. Doležel J, Čížková J, Šimková H, Bartoš J. One major challenge of sequencing large plant genomes is to know how big they really are. Int J Mol Sci. 2018;19:3554.
51. Bainard JD, Husband BC, Baldwin SJ, Fazekas A, Gregory TR, Newmaster SG, et al. The effects of rapid desiccation on estimates of plant genome size. Chromosome Res. 2011;19:825–42.
52. Šmarda P, Stančík D. Ploidy level variability in south American fescues (*Festuca* L., Poaceae): use of flow cytometry in up to 5 1/-2-year-old caryopses and herbarium specimens. Plant Biol. 2006;8: 73–80.
53. Viruel J, Conejero M, Hidalgo O, Pokorny L, Powell RF, Forest F, et al. A target capture-based method to estimate ploidy from herbarium specimens. Front Plant Sci. 2019;10:937.
54. Šmarda P, Bureš P, Horová L, Leitch IJ, Mucina L, Pacini E, et al. Ecological and evolutionary significance of genomic GC content diversity in monocots. Proc Nat Acad Sci USA. 2014;111:E4096–102.
55. Doležel J, Greilhuber J, Lucretti S, Meister A, Lysák MA, Nardi L, et al. Plant genome size estimation by flow cytometry: inter-laboratory comparison. Ann Botany. 1998;82(Suppl. A):17–26.
56. Galbraith DW. Endoreduplicative standards for calibration of flow cytometric C-value measurements. Cytometry A. 2014;85:368–74.
57. Trávníček P, Ponert J, Urfus T, Jersáková J, Vrána J, Hřibová E, et al. Challenges of flow-cytometric estimation of nuclear genome size in orchids, a plant group with both whole-genome and progressively partial endoreplication. Cytometry A. 2015;87:958–66.
58. Doležel J, Sgorbati S, Lucretti S. Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. Physiol Plant. 1992;85:625–31.
59. Cossarizza A, Chang H-D, Radbruch A, Acs A, Adam D, Adam-Klages S, et al. Guidelines for the use of flow cytometry and cell sorting in immunological studies (second edition). Eur J Immunol. 2019;49:1457–973.
60. Doležel J, Greilhuber J, Suda J. Estimation of nuclear DNA content in plants using flow cytometry. Nat Protoc. 2007;2:2233–44.
61. Shapiro HM. Practical Flow Cytometry. New Jersey: John Wiley & Sons, Inc.; 2003.
62. Pellicer J, Powell RF, Leitch IJ. The application of flow cytometry for estimating genome size, ploidy level endopolyploidy, and reproductive modes in plants. In: Besse P, editor. Molecular plant taxonomy. Methods in molecular biology. Volume 2222. New York, NY: Humana; 2021. p. 325–62.
63. Šmarda P, Bureš P. Understanding intraspecific variation in genome size in plants. Preslia. 2010;82:41–61.
64. Vindeløv LL, Christensen IJ, Nissen NI. Standardization of high-resolution flow cytometric DNA analysis by the simultaneous use of chicken and trout red blood cells as internal reference standards. Cytometry. 1983;3:328–31.
65. Šmarda P, Bureš P. The variation of base composition in plant genomes, Plant genomes, their residents, and their evolutionary dynamics. In: Wendel JF, Greilhuber J, Doležel J, Leitch IJ, editors. Plant genome diversity. Wien: Springer-Verlag; 2012. p. 209–36.
66. Trávníček P, Čertner M, Ponert J, Chumová Z, Jersáková J, Suda J. Diversity in genome size and GC content shows adaptive potential in orchids and is closely linked to partial endoreplication, plant life-history traits and climatic conditions. New Phytol. 2019;224: 1642–56.
67. Suda J, Leitch IJ. The quest for suitable reference standards in genome size research. Cytometry A. 2010;77A:717–20.
68. Osborne DJ, Boubriak II. DNA and desiccation tolerance. Seed Sci Res. 2008;4:175–85.
69. Leprince O, Colson P, Houssier C, Deltour R. Changes in chromatin structure associated with germination of maize and their relation with desiccation tolerance. Plant Cell Environ. 1995;18:619–29.
70. Leprince O, Buitink J. Introduction to desiccation biology: from old borders to new frontiers. Planta. 2015;242:369–78.
71. Leitch IJ, Johnston E, Pellicer J, Hidalgo O, Bennett MD. Plant DNA C-values Database (release 7.1, April 2019). 2019.
72. Greilhuber J. Cytochemistry and C-values - the less well known world of nuclear DNA amounts. Ann Bot. 2008;101:791–804.
73. Galbraith DW, Villalobos-Menuey E. Analysis of plant genome sizes using flow cytometry: a case study demonstrating dynamic range and measurement linearity. Beckman Coulter App Note. 2020;1–13. https://labplan.ie/content/uploads/2018/02/Plant-Genome-Analysis-by-Flow-Cytometry.pdf
74. Rice A, Glick L, Abadi S, Einhorn M, Kopelman NM, Salman-Minkov A, et al. The chromosome counts database (CCDB) – a community resource of plant chromosome numbers. New Phytol. 2015; 206:19–26.
75. Piegu B, Guyot R, Picault N, Roulin A, Saniyal A, Kim H, et al. Doubling genome size without polyploidization: dynamics of retrotransposition-driven genomic expansions in *Oryza australiensis*, a wild relative of rice. Genome Res. 2006;16:1262–9.
76. Kelly LJ, Renny-Byfield S, Pellicer J, Macas J, Novák P, Neumann P, et al. Analysis of the giant genomes of *Fritillaria* (Liliaceae) indicates that a lack of DNA removal characterizes extreme expansions in genome size. New Phytol. 2015;208:596–607.
77. Greilhuber J. Intraspecific variation in genome size: a critical reassessment. Annals of Botany. 1998;82(supplement A):27–35.
78. Greilhuber J. Intraspecific variation in genome size in angiosperms - identifying its existence. Ann Bot. 2005;95:91–8.
79. Noirot M, Barre P, Duperray C, Louarn J, Hamon S. Effects of caffeine and chlorogenic acid on propidium iodide accessibility to DNA:

consequences on genome size evaluation in coffee tree. Ann Bot. 2003;92:259–64.
80. Noirot M, Barre P, Louarn J, Duperray C, Hamon S. Nucleus-cytosol interactions - a source of stoichiometric error in flow cytometric estimation of nuclear DNA content in plants. Ann Bot. 2000;86:309–16.
81. Price HJ, Hodnett G, Johnston JS. Sunflower (*Helianthus annuus*) leaves contain compounds that reduce nuclear propidium iodide fluorescence. Ann Bot. 2000;86:929–34.
82. Loureiro J, Rodriguez E, Doležel J, Santos C. Flow cytometric and microscopic analysis of the effect of tannic acid on plant nuclei and estimation of DNA content. Ann Bot. 2006;98:515–27.
83. Wood JCS. Establishing and maintaining system linearity. Curr Protoc Cytom. 2009;47:1.4.1–1.4.14.
84. Vindeløv LL, Christensen IJ, Jensen G, Nissen NI. Limits of detection of nuclear DNA abnormalities by flow cytometric DNA analysis. Results obtained by a set of methods for sample-storage, staining and internal standardization. Cytometry. 1983;3:332–9.
85. Doležel J, Gohde W. Sex determination in dioecious plants *Melandrium album* and *M. rubrum* using high-resolution flow-cytometry. Cytometry. 1995;19:103–6.
86. Šmarda P, Horová L, Knápek O, Dieck H, Dieck M, Ražná K, et al. Multiple haploids, triploids, and tetraploids found in modern-day "living fossil" *Ginkgo biloba*. Horticulture Res. 2018;5:55.
87. Čertnerová D, Škaloud P. Substantial intraspecific genome size variation in golden-brown algae and its phenotypic consequences. Ann Bot. 2020;126:1077–87.
88. Melichárková A, Šlenker M, Zozomová-Lihová J, Skokanová K, Šingliarová B, Kačmárová T, et al. So closely related and yet so different: strong contrasts between the evolutionary histories of species of the *Cardamine pratensis* polyploid complex in Central Europe. Front Plant Sci. 2020;11:1988.
89. Šmarda P, Bureš P, Hovrová L, Rotreklová O. Intrapopulation genome size dynamics in *Festuca pallens*. Ann Bot. 2008;102:599–607.
90. Trávníček P, Kirschner J, Chudáčková H, Rooks F, Štěpánek J. Substantial genome size variation in *Taraxacum stenocephalum* (Asteraceae, Lactuceae). Folia Geobot. 2013;48:271–84.
91. Šmarda P, Bureš P. Intraspecific DNA content variability in *Festuca pallens* on different geographical scales and ploidy levels. Ann Bot. 2006;98:665–78.
92. Barow M, Meister A. Lack of correlation between AT frequency and genome size in higher plants and the effect of non-randomness of base sequences on dye binding. Cytometry. 2002;47:1–7.
93. Meister A, Barow M. DNA base composition of plant genomes. In: Doležel J, Greilhuber J, Suda J, editors. Flow cytometry with plant cells. Weinheim: Viley-VCH; 2007. p. 177–215.
94. Šmarda P, Bureš P, Šmerda J, Horová L. Measurements of genomic GC content in plant genomes with flow cytometry: a test for reliability. New Phytol. 2012;193:513–21.
95. Meister A. Calculation of binding length of base-specific DNA dyes by comparison of sequence and flow cytometric data. Application to *Oryza sativa* and *Arabidopsis thaliana*. J Theor Biol. 2005;232:93–7.
96. Stackebrandt E, Liesack W. Nucleic acids and classification. In: Goodfellow M, O'Donnell AG, editors. Handbook of new bacterial systematics. Academic Press: London; 1993. p. 151–94.
97. Mann S, Chen Y-PP. Bacterial genomic G+C composition-eliciting environmental adaptation. Genomics. 2010;95:7–15.
98. Almpanis A, Swain M, Gatherer D, McEwan N. Correlation between bacterial G+C content, genome size and the G+C content of associated plasmids and bacteriophages. Microbial Genomic. 2018;4:4.
99. Rocha EPC, Danchin A. Base composition bias might result from competition for metabolic resources. Trends Genet. 2002;18:291–4.
100. Vinogradov AE. DNA helix: the importance of being GC-rich. Nucleic Acids Res. 2003;31:1838–44.
101. Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science. 1983;220:1049–51.
102. Otto F, Tsou KC. A comparative study of DAPI, DIPI, and Hoechst 33258 and 33342 as chromosomal DNA stains. Stain Technol. 1985;60:7–11.
103. Lewalski H, Otto FJ, Kranert T, Wassmuth R. Flow cytometric detection of unbalanced ram spermatozoa from heterozygous 1;20 translocation carriers. Cytogenet Genome Res. 1993;64:286–91.
104. Kapuscinski J. DAPI: a DNA-specific fluorescent probe. Biotech Histochem. 1995;70:220–33.
105. Wilson WD, Tanious FA, Barton HJ, Jones RL, Strekowski L, Boykin DW. Binding of 4′,6-diamidino-2-phenylindole (DAPI) to GC and mixed sequences in DNA: intercalation of a classical groove-binding molecule. J Am Chem Soc. 1989;111:5008–10.
106. Reis LA, Rocha MS. DNA interaction with DAPI fluorescent dye: force spectroscopy decouples two different binding modes. Biopolymers. 2017;107:e23015.
107. Bernardi G. Isochores and the evolutionary genomics of vertebrates. Gene. 2000;241:3–17.
108. Langlois RG, Carrano AV, Gray JW, Van Dilla MA. Cytochemical studies of metaphase chromosomes by flow cytometry. Chromosoma. 1980;77:229–51.
109. Godelle B, Cartier D, Marie D, Brown SC, Siljak-Yakovlev S. Heterochromatin study demonstrating the nonlinearity of fluorometry useful for calculating genomic base composition. Cytometry. 1993;14:618–26.
110. Abu-Daya A, Brown PM, Fox KR. DNA sequence preferences of several AT-selective minor groove binding ligands. Nucleic Acids Res. 1995;23:3385–92.
111. Breusegem SY, Clegg RM, Loontiens FG. Base-sequence specificity of Hoechst 33258 and DAPI binding to five $(a/T)_4$ DNA sites with kinetic evidence for more than one high-affinity Hoechst 33258-AATT complex. J Mol Biol. 2002;315:1049–61.
112. Veleba A, Bureš P, Adamec L, Šmarda P, Lipnerová I, Horová L. Genome size and genomic GC content evolution in the miniature genome-sized family Lentibulariaceae. New Phytol. 2014;203:22–8.
113. Ibarra-Laclette E, Lyons E, Hernandez-Guzman G, Perez-Torres CA, Carretero-Paulet L, Chang T-H, et al. Architecture and evolution of a minute plant genome. Nature. 2013;498:94–8.
114. Black M, Bewley JD, Halmer P. The encyclopedia of seeds: science, technology and uses. Wallingford: CABI; 2006.p. 828.
115. Mickelson-Young L, Wear E, Mulvaney P, Lee T-J, Szymanski ES, Allen G, et al. A flow cytometric method for estimating S-phase duration in plants. J Exp Bot. 2016;67:6077–87.
116. Naill MC, Roberts SC. Cell cycle analysis of *Taxus* suspension cultures at the single cell level as an indicator of culture heterogeneity. Biotechnol Bioeng. 2005;90:491–500.
117. Lucretti S, Nardi L, Nisini PT, Moretti F, Gualberti G, Doležel J. Bivariate flow cytometry DNA/BrdUrd analysis of plant cell cycle. Methods Cell Sci. 1999;21:155–66.
118. Doonan JH. The plant cell cycle. In: Humphrey T, Brooks G, editors. Cell cycle control: mechanisms and protocols. Volume 296 methods in molecular biology. Totowa: Humana Press Inc; 2010. p. 31–50.
119. Sliwinska E. Nuclear DNA replication and seed quality. Seed Sci Res. 2009;19:15–25.
120. Bino RJ, Bergervoet JHW, De Vos CHR, Kraak HL, Lanteri S, Van Der Burg WJ, et al. Comparison of nuclear replication activity and protein expression patterns during tomato seed germination. Field Crop Res. 1996;45:71–7.
121. Yan LL, Zhang YJ, Gao WY, Man SL, Wang Y. *In vitro* and *in vivo* anticancer activity of steroid saponins of *Paris polyphylla* var. *yunnanensis*. Exp Oncol. 2009;31:27–32.
122. Parajuli P, Joshee N, Rimando AM, Mittal S, Yadav AK. *In vitro* antitumor mechanisms of various *Scutellaria* extracts and constituent flavonoids. Planta Med. 2009;75:41–8.

123. McMurphy LM, Rayburn AL. Nuclear alterations of maize plants grown in soil contaminated with coal fly ash. Arch Environ Contam Toxicol. 1993;25:520–4.
124. Monteiro MS, Rodriguez E, Loureiro J, Mann RM, Soares AMVM, Santos C. Flow cytometric assessment of cd genotoxicity in three plants with different metal accumulation and detoxification capacities. Ecotoxicol Environ Saf. 2010;73:1231–7.
125. Carballo JA, Pincheira J, de la Torre C. The G_2 checkpoint activated by DNA damage does not prevent genome instability in plant cells. Biol Res. 2006;39:331–40.
126. Rodriguez E, Santos C, Azevedo R, Moutinho-Pereira J, Correia C, Dias MC. Chromium (VI) induces toxicity at different photosynthetic levels in pea. Plant Physiol Biochem. 2012;53:94–100.
127. Doležel J, Lucretti S, Molnár I, Cápal P, Giorgi D. Chromosome analysis and sorting. Cytometry A. 2021;99:328–42.
128. D'Amato F. Role of polyploidy in reproductive organs and tissues. In: Jori BM, editor. Embryology of angiosperms. New York: Springer Verlag; 1984. p. 519–66.
129. Leitch IJ, Dodsworth S. Endopolyploidy in plants. eLS. 2017. 1–10. https://doi.org/10.1002/9780470015902.a0020097.pub2
130. Nagl W. Nuclear organization. Annu Rev Plant Physiol. 1976;27:39–69.
131. Nagl W. DNA endoreduplication and polyteny understood as evolutionary strategies. Nature. 1976;261:614–5.
132. Joubès J, Chevalier C. Endoreduplication in higher plants. Plant Mol Biol. 2000;43:735–45.
133. Maluszynska J, Kolano B, Sas-Nowosielska H. Endopolyploidy in plants, Physical structure, behaviour and evolution of plant genomes. In: Leitch IJ, Greilhuber J, Doležel J, Wendel JF, editors. Plant genome diversity. Volume 2. Wien: Springer-Verlag; 2013. p. 99–119.
134. Sugimoto-Shirasu K, Roberts K. "Big it up": endoreduplication and cell-size control in plants. Curr Opin Plant Biol. 2003;6:544–53.
135. Sabelli PA. Replicate and die for your own good: Endoreduplication and cell death in the cereal endosperm. J Cereal Sci. 2012;56:9–20.
136. Sliwinska E, Bassel GW, Bewley JD. Germination of *Arabidopsis thaliana* seeds is not completed as a result of elongation of the radicle but of the adjacent transition zone and lower hypocotyl. J Exp Bot. 2009;60:3587–94.
137. Makowczyńska J, Andrzejewska-Golec E, Sliwinska E. Nuclear DNA content in different plant materials of *Plantago asiatica* L. cultured *in vitro*. Plant Cell Tissue Organ Cult. 2008;94:65.
138. Rewers M, Sliwinska E. Endoreduplication in the germinating embryo and young seedling is related to the type of seedling establishment but is not coupled with superoxide radical accumulation. J Exp Bot. 2014;65:4385–96.
139. Barow M, Jovtchev G. Endopolyploidy in plants and its analysis by flow cytometry. In: Doležel J, Greilhuber J, Suda J, editors. Flow cytometry with plant cells. Weinheim: Wiley-VCH Verlag GmbH & Co.; 2007. p. 349–72.
140. Barow M, Meister A. Endopolyploidy in seed plants is differently correlated to systematics, organ, life strategy and genome size. Plant Cell Environ. 2003;26:571–84.
141. Bainard JD, Bainard LD, Henry TA, Fazekas AJ, Newmaster SG. A multivariate analysis of variation in genome size and endoreduplication in angiosperms reveals strong phylogenetic signal and association with phenotypic traits. New Phytol. 2012;196:1240–50.
142. Rewers M, Sliwinska E. Endoreduplication intensity as a marker of seed developmental stage in the Fabaceae. Cytometry A. 2012;81A:1067–75.
143. Larkins BA, Dilkes BP, Dante RA, Coelho CM, Ym W, Liu Y. Investigating the hows and whys of DNA endoreduplication. J Exp Bot. 2001;52:183–92.
144. Lukaszewska E, Sliwinska E. Most organs of sugar-beet (*Beta vulgaris* L.) plants at the vegetative and reproductive stages of development are polysomatic. Sex Plant Reprod. 2007;20:99–107.
145. Galbraith DW, Harkins KR, Knapp S. Systemic endopolyploidy in *Arabidopsis thaliana*. Plant Physiol. 1991;96:985–9.
146. Zhang C, Gong FC, Lambert GM, Galbraith DW. Cell type-specific characterization of nuclear DNA contents within complex tissues and organs. Plant Methods. 2005;1:7.
147. Gendreau E, Höfte H, Grandjean O, Brown S, Traas J. Phytochrome controls the number of endoreduplication cycles in the *Arabidopsis thaliana* hypocotyl. Plant J. 1998;13:221–30.
148. Chen CT, Setter TL. Response of potato tuber cell division and growth to shade and elevated CO_2. Ann Bot. 2003;91:373–81.
149. Bainard LD, Bainard JD, Newmaster SG, Klironomos JN. Mycorrhizal symbiosis stimulates endoreduplication in angiosperms. Plant Cell Environ. 2011;34:1577–85.
150. Chandran D, Inada N, Hather G, Kleindt CK, Wildermuth MC. Laser microdissection of *Arabidopsis* cells at the powdery mildew infection site reveals site-specific processes and regulators. Proc Nat Acad Sci USA. 2010;107:460–5.
151. Caillaud M-C, Dubreuil G, Quentin M, Perfus-Barbeoch L, Lecomte P, de Almeida EJ, et al. Root-knot nematodes manipulate plant cell functions during a compatible interaction. J Plant Physiol. 2008;165:104–13.
152. Lingua G, D'Agostino G, Fusconi A, Berta G. Nuclear changes in pathogen-infected tomato roots. Eur J Histochem. 2001;45:21–30.
153. Scholes DR, Paige KN. Chromosomal plasticity: mitigating the impacts of herbivory. Ecology. 2011;92:1691–8.
154. Bory S, Catrice O, Brown SC, Leitch IJ, Gigant R, Chiroleu F, et al. Natural polyploidy in *Vanilla planifolia* (Orchidaceae). Genome. 2008;51:816–26.
155. Hřibová E, Holušová K, Trávníček P, Petrovská B, Ponert J, Šimková H, et al. The enigma of progressively partial endoreplication: new insights provided by flow cytometry and next-generation sequencing. Genome Biol Evol. 2016;8:1996–2005.
156. Brown SC, Bourge M, Maunoury N, Wong M, Wolfe Bianchi M, Lepers-Andrzejewski S, et al. DNA remodeling by strict partial endoreplication in orchids, an original process in the plant kingdom. Genome Biol Evol. 2017;9:1051–71.
157. Lemontey C, Mousset-Déclas C, Munier-Jolain N, Boutin JP. Maternal genotype influences pea seed size by controlling both mitotic activity during early embryogenesis and final endoreduplication level/cotyledon cell size in mature seed. J Exp Bot. 2000;51:167–75.
158. Nowicka A, Kovacik M, Tokarz B, Vrána J, Zhang Y, Weigt D, et al. Dynamics of endoreduplication in developing barley seeds. J Exp Bot. 2021;72:268–82.
159. Matzk F, Meister A, Schubert I. An efficient screen for reproductive pathways using mature seeds of monocots and dicots. Plant J. 2000;21:97–108.
160. Barcaccia G, Arzenton F, Sharbel TF, Varotto S, Parrini P, Lucchin M. Genetic diversity and reproductive biology in ecotypes of the facultative apomict *Hypericum perforatum* L. Heredity. 2006;96:322–34.
161. Delgado L, Galdeano F, Sartor ME, Quarin CL, Espinoza F, Ortiz JPA. Analysis of variation for apomictic reproduction in diploid *Paspalum rufum*. Ann Bot. 2014;113:1211–8.
162. Samaniego F, Kolář F, Urfus T, Barragán Á, Romoleroux K. Determination of apomixis by flow cytometry in two species of *Lachemilla* (Rosaceae) in Ecuador. Neotropical Biodiversity. 2018;4:152–63.
163. Koutecký P, Baurová T, Štech M, Kosnar J, Karásek J. Hybridization between diploid *Centaurea pseudophrygia* and tetraploid *C. jacea* (Asteraceae): the role of mixed pollination, unreduced gametes, and mentor effects. Biol J Linn Soc. 2011;104:93–106.
164. Krahulcová A, Krahulec F, Vladimirov V. Impact of interspecific hybridization within a polyploid agamic complex of *Pilosella* (Asteraceae, Cichorieae) in Bulgaria compared with Central Europe. Willdenowia. 2018;48:345–62.
165. Chrtek J, Herben T, Rosenbaumová R, Münzbergová Z, Dočkalová Z, Zahradníček J, et al. Cytotype coexistence in the field

cannot be explained by inter-cytotype hybridization alone: linking experiments and computer simulations in the sexual species *Pilosella echioides* (Asteraceae). BMC Evol Biol. 2017;17:87.
166. Friedman WE, Madrid EN, Williams JH. Origin of the fittest and survival of the fittest: relating female gametophyte development to endosperm genetics. Int J Plant Sci. 2008;169:79–92.
167. Friedman WE, Ryerson KC. Reconstructing the ancestral female gametophyte of angiosperms: insights from *Amborella* and other ancient lineages of flowering plants. Am J Bot. 2009;96:129–43.
168. Kreiner JM, Kron P, Husband BC. Evolutionary dynamics of unreduced gametes. Trends Genet. 2017;33:583–93.
169. Aliyu O, Schranz E, Sharbel T. Quantitative variation for apomictic reproduction in the genus *Boechera* (Brassicaceae). Am J Bot. 2010; 97:1719–31.
170. Kaushal P, Dwivedi KK, Radhakrishna A, Srivastava MK, Kumar V, Roy AK, et al. Partitioning apomixis components to understand and utilize gametophytic apomixis. Front Plant Sci. 2019;10:256.
171. Matzk F. Reproductive mode screening. In: Doležel J, Greilhuber J, Suda J, editors. Flow cytometry with plant cells. Weinheim: Wiley-VCH Verlag GmbH & Co.; 2007. p. 131–51.
172. Krahulcová A, Rotreklová O. Use of flow cytometry in research on apomictic plants. Preslia. 2010;82:23–39.
173. Rojek J, Kapusta M, Kozieradzka-Kiszkurno M, Majcher D, Górniak M, Sliwinska E, et al. Establishing the cell biology of apomictic reproduction in diploid *Boechera stricta* (Brassicaceae). Ann Bot. 2018;122:513–39.
174. Madrid EN, Friedman WE. Female gametophyte and early seed development in *Peperomia* (Piperaceae). Am J Bot. 2010;97:1–14.
175. Brožová V, Koutecký P, Doležal J. Plant apomixis is rare in Himalayan high-alpine flora. Sci Rep. 2019;9:14386.
176. Sabelli PA. Seed development: a comparative overview on biology of morphology, physiology, and biochemistry between monocot and dicot plants. In: Agrawal G, Rakwal R, editors. Seed development: OMICS technologies toward improvement of seed quality and crop yield. Dordrecht: Springer; 2012. p. 3–25.
177. Sliwinska E, Zielinska E, Jedrzejczyk I. Are seeds suitable for flow cytometric estimation of plant genome size? Cytometry A. 2005; 64A:72–9.
178. Doležel J, Bartoš J. Plant DNA flow cytometry and estimation of nuclear genome size. Ann Bot. 2005;95:99–110.
179. Talent N, Dickinson TA. Endosperm formation in aposporous *Crataegus* (Rosaceae, Spiraeoideae, tribe Pyreae): parallels to Ranunculaceae and Poaceae. New Phytol. 2007;173:231–49.
180. Lepší M, Koutecký P, Nosková J, Lepší P, Urfus T, Rich TCG. Versatility of reproductive modes and ploidy level interactions in *Sorbus* s.l. (Malinae, Rosaceae). Botanic J Linnean Soc. 2019;191:502–22.
181. Crane CF. Classification of apomictic mechanisms. In: Savidan Y, Carman JG, Dresselhaus T, editors. The flowering of apomixis: from mechanisms to genetic engineering. Mexico City, Mexico: CIMMYT, IRD & European Commission; 2001. p. 24–43.
182. Dobeš C, Lückl A, Hülber K, Paule J. Prospects and limits of the flow cytometric seed screen–insights from *Potentilla sensu lato* (Potentilleae, Rosaceae). New Phytol. 2013;198:605–16.
183. Kolarčik V, Kocová V, Vašková D. Flow cytometric seed screen data are consistent with models of chromosome inheritance in asymmetrically compensating allopolyploids. Cytometry A. 2018;93:737–48.
184. Bairu MW, Aremu AO, Van Staden J. Somaclonal variation in plants: causes and detection methods. Plant Growth Regul. 2011;63:147–73.
185. Miler N, Kulus D, Sliwinska E. Nuclear DNA content as an indicator of inflorescence colour stability of *in vitro* propagated solid and chimera mutants of *chrysanthemum*. Plant Cell Tiss Org Cult. 2020;143:421–30.
186. Arun B, Singh BD, Sharma S, Paliwal R, Joshi AK. Development of somaclonal variants of wheat (*Triticum aestivum* L.) for yield traits and disease resistance suitable for heat stressed and zero-till conditions. Field Crop Res. 2007;103:62–9.
187. Żabicki P, Sliwinska E, Mitka J, Sutkowska A, Tuleja M, Migdałek G, et al. Does somaclonal variation play advantageous role in conservation practice of endangered species?: comprehensive genetic studies of *in vitro* propagated plantlets of *Viola stagnina* kit. (Violaceae). Plant cell. Tissue Org Cult. 2019;136:339–52.
188. Antoniadi I, Skalický V, Sun G, Ma W, Galbraith DW, Novák O, et al. Fluorescence activated cell sorting—a selective tool for plant cell isolation and analysis. Cytometry Part A. 2021. https://doi.org/10.1002/cyto.a.24461
189. Greilhuber J, Temsch EM, Loureiro J. Nuclear DNA content measurement. In: Doležel J, Greilhuber J, Suda J, editors. Flow cytometry with plant cells. Weinheim: Wiley-VCH; 2007. p. 67–102.

How to cite this article: Sliwinska E, Loureiro J, Leitch IJ, Šmarda P, Bainard J, Bureš P, et al. Application-based guidelines for best practices in plant flow cytometry. Cytometry. 2022;101:749–81. https://doi.org/10.1002/cyto.a.24499