Discover the beauty of standardized flow cytometry

Workflow standardization and automation made easy with the MACSQuant® Analyzer 16

Benefit from our best-in-class solutions for standardized flow cytometry. Find the perfect synergy of the right tools and be rewarded with reliable and reproducible results

- **Highly specific and consistent antibodies**
- **Flawless instrumentation**
- **Automated sample labeling and acquisition**
- **Automated analysis and reporting software**

▸ **miltenyibiotec.com/mq16**

Miltenyi Biotec B.V. & Co. KG | Friedrich-Ebert-Straße 68 | 51429 Bergisch Gladbach | Germany
Phone +49 2204 8306-0 | Fax +49 2204 85197 | macsde@miltenyi.com | **www.miltenyibiotec.com**

Miltenyi Biotec provides products and services worldwide.
Visit **www.miltenyibiotec.com/local** to find your nearest Miltenyi Biotec contact.

Unless otherwise specifically indicated, Miltenyi Biotec products and services are for research use only and not for therapeutic or diagnostic use. MACSQuant, Tyto, the Tyto logo, and the Miltenyi Biotec logo are registered trademarks or trademarks of Miltenyi Biotec and/or its affiliates in various countries worldwide. Copyright © 2022 Miltenyi Biotec and/or its affiliates. All rights reserved.

Miltenyi Biotec

Check for updates

Received: 28 April 2022 | Revised: 26 May 2023 | Accepted: 12 September 2023

DOI: 10.1002/cyto.a.24798

REVIEW

MIFlowCyt

ISAC INTERNATIONAL SOCIETY FOR ADVANCEMENT OF CYTOMETRY

CYTOMETRY Journal of Quantitative Cell Science PART A

Best practices for instrument settings and raw data analysis in plant flow cytometry

Petr Koutecký[1] | Tyler Smith[2] | João Loureiro[3] | Paul Kron[4]

[1]Department of Botany, Faculty of Science, University of South Bohemia, České Budějovice, Czechia

[2]Agriculture and Agri-Food Canada (AAFC), Ottawa, Ontario, Canada

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

[4]Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada

Correspondence
Petr Koutecký, Department of Botany, Faculty of Science, University of South Bohemia, Branišovská 1760, České Budějovice 37005, Czechia.
Email: kouta@prf.jcu.cz

Funding information
Centro de Ecologia Funcional, Grant/Award Number: FCT/UIDB/04004/2020; Natural Sciences and Engineering Research Council of Canada, Grant/Award Number: RGPIN-2020-05652; Programa Operacional Regional do Centro, Grant/Award Number: CENTRO-01-0145-FEDER-000020

Abstract

Flow cytometry (FCM) is now the most widely used method to determine ploidy levels and genome size of plants. To get reliable estimates and allow reproducibility of measurements, the methodology should be standardized and follow the best practices in the field. In this article, we discuss instrument calibration and quality control and various instrument and acquisition settings (parameters, flow rate, number of events, scales, use of discriminators, peak positions). These settings must be decided before measurements because they determine the amount and quality of the data and thus influence all downstream analyses. We describe the two main approaches to raw data analysis (gating and histogram modeling), and we discuss their advantages and disadvantages. Finally, we provide a summary of best practice recommendations for data acquisition and raw data analysis in plant FCM.

KEYWORDS
best practices, calibration, flow cytometry, gating, histogram modeling, instrument settings, plant sciences

1 | INTRODUCTION

Recently, a series of articles on best practices in flow cytometry (FCM) with plants [1] has been initiated. Most of the applications are based on estimating nuclear DNA content (genome size) [2]. Best practices concerning material selection, sample preparation, and specific measurement precision and data interpretation requirements are summarized in other articles of this series [2–5]. With some particular modifications, the same methodology is also used with other organisms treated under botany in the broadest sense, such as fungi [6] or algae [7]. All the cited articles, however, consider mainly pre-measurement (i.e., sample collection and preparation) and post-measurement steps (i.e., analysis of processed data) and do not address how to collect the measurements themselves or how to process the raw data (i.e., histogram analysis). The present article aims to fill this gap: we will focus on instrument settings and measurement per se (which determines the precision of results and some features that cannot be improved during the subsequent analyses) and the analysis of raw data.

2 | INSTRUMENT CALIBRATION AND QUALITY CONTROL

Standard biomedical flow cytometers are routinely used in plant biology, although instruments optimized for plant samples may provide superior results. This may mean specialized instruments designed for specific organisms and tissues, such as ship-borne instruments for phytoplankton sampling [8], instruments for sampling air-borne pollen and fungi [9], and instruments that can be optimized for fragile

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2023 The Authors. *Cytometry Part A* published by Wiley Periodicals LLC on behalf of International Society for Advancement of Cytometry.

tissues such as protoplasts [10]. However, in most cases, all that is required is an instrument with light sources and optics appropriate for the fluorochromes and applications common in plant FCM. The optical filter set should be compatible with the spectral properties (i.e., excitation and emission spectra) of the DNA fluorochrome to be used [4], propidium iodide (PI), and 4′,6-diamidino-2-phenylindole dihydrochloride (DAPI) being the most frequent ones. PI-stained samples are optimally analyzed on instruments equipped with a green solid-state laser (532 nm). This wavelength is usually not included with standard biomedical flow cytometers but is often an option with new instruments. Alternatively, an argon-ion turquoise laser (488 nm) is more commonly available and a 561 nm solid-state laser can also be used; still, they provide slightly less efficient excitation than the green laser. DAPI-stained samples are excited mainly by UV bands emitted from mercury arc lamps or, in more recent instruments, by UV light-emitting diodes (LEDs) or UV lasers.

Regardless of the flow cytometer used and the specific application, the instrument should be in optimal condition in terms of fluidics and optics alignment to achieve accurate results (most often displayed as a histogram of fluorescence, in which nuclei of particular DNA contents should form narrow symmetric peaks with minimum background; see Chapter 5 for details on histogram components). Fluidics must be free of air bubbles and in-line filters free of any blocking particles, such as microorganisms. As sheath fluid, one may use commercial solutions sold by cytometer manufacturers. As an alternative suitable for most applications, users can opt for double sterile distilled water. To prevent microbial growth, especially if the instrument is not used daily, various biocides may be used, such as sodium azide (0.02% w/v, [11]) or ProClin™ (Sigma-Aldrich; concentrations 0.03%–0.1% are recommended by the manufacturer for particular formulations of this biocide). Plant tissues include various secondary metabolites, such as mucilage, that not only distort sample preparation and staining [4] but may adhere to tubes and flow chamber walls and contribute to clogging and distortion of the laminar flow. Therefore, for plant cytometers, it is necessary to use cleaning solutions regularly and replace the tube delivering the sample to the flow chamber quite often (when any impurities or discoloration become visible). For flow system maintenance, we refer users to technical manuals specific to the particular cytometers they use. Regular servicing by the manufacturer is also an option but can represent a substantial recurring cost. In laser instruments, it is crucial to ensure an optimal alignment of the light source with the core stream in the flow chamber because the light intensity decreases toward the margins of the laser beam [12]. If the laser is not properly aligned, part of nuclei may receive incomplete illumination, which may result in shift of the mean fluorescence, higher coefficients of variation (CVs; a measure of peak "width"), and asymmetric peaks.

For measurements of genome size and derived parameters (such as ploidy levels), which are the most common applications in plant FCM [2], measured fluorescence (i.e., the *x*-axis of the histogram) must be proportional to the actual fluorescence (number of photons emitted), because the fluorescence of the particles is expected to be proportional to their DNA content. In other words, we expect a linear relationship between DNA content of nuclei, fluorescence of (properly stained) nuclei, and the reported fluorescence measure. Failure to have a linear relationship between DNA content and fluorescence is a sample preparation (i.e., nuclei isolation and staining) issue, but failure to have a linear relationship between actual and measured fluorescence is an instrument linearity problem. Instruments from different manufacturers differ in the linearity of their response (see Wood [13] for details on sources of non-linearity), and the users should be aware of the performance of their instruments.

Quality control tests for fluidics and optical alignment involve testing of peak CVs, while linearity tests compare peak means for particles of known relative fluorescence. We recommend using calibration particles to check all three features. Fluorescent beads (usually latex or polystyrene microspheres with a size ranging from 0.5 to 6 µm, covered with the selected fluorochrome) can be used for such purposes. Their consistent size ensures that their measurements reflect instrument behavior rather than variation in the sample. Single-type beads can be used to check fluidics stability and optics alignment. When the flow cytometer is well-calibrated, there should be low background and the CV (which is the standard deviation [SD] divided by mean and expressed as %) of the fluorescent peak of the beads should be below 2.0%. The use of so-called rainbow calibration particles that contain a mixture of beads of the same diameter (e.g., 3 µm) covered with different amounts of the fluorochrome (i.e., yielding different fluorescent intensities) enables the operator to check the stability and resolution of a flow cytometer as well as the linearity of the measurement. Concerning linearity, regression of fluorescence intensity (channel number on a linear scale) and expected intensity (from the bead characteristics) should produce a straight line. An alternative test for non-linearity was proposed by Bagwell et al. [14]. Two types of beads of different fluorescence intensities are measured simultaneously using different amplifications of the signal (different settings of photomultipliers or amplifiers; the respective parameters are usually called voltage or gain by different cytometer manufacturers) and regression of the differences between their mean fluorescence versus mean fluorescence of the lower fluorescence beads should produce a straight line. For precise measurements, only a limited range of fluorescence intensities (in the most typical case, of a sample and a standard) should be used, within which the linear response is assured by the manufacturer for a particular cytometer. For most instruments, the recommended range is less than 3- or 4-fold [15]. This linearity limitation means that an array of standards (of different genome sizes) is required to cover the wide range of plant DNA contents [15].

Due to the high costs of the calibration beads, suitable alternatives (though never so uniform and standardized as beads) include the use of fixed cells, such as chicken or trout red blood cells (where CVs <3% are expected in well-tuned instruments), or even plant standards with known superior quality. A good example of the latter is *Pisum sativum*, one of the gold reference standards that should provide fluorescence peaks with a CV below 2.0% when the cytometer is perfectly calibrated. Chicken and trout erythrocytes are known to form small fractions of cells stuck together (doublets, triplets, etc.), which can be used to evaluate linearity in the same way as bead mixtures

15524930, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.24798 by Cochrane Portugal, Wiley Online Library on [03/11/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(the ratio of the mean fluorescence of doublets to singlets is expected to be 2.0 and acceptable in the range 1.95–2.05). However, some caution is needed, as performance may depend on the cytometer construction, especially when the pulse height is used as a measure of fluorescence intensity. Unlike uniformly sized beads, clusters of particles are of larger size and, thus, may be less uniformly illuminated in laser instruments [13]. When using biological standards, if additional peaks are present (nuclei in G_2 phase or due to endopolyploidy) and complete replication of DNA is assumed, these can also be used to assess instrument linearity. In our experience working with standards, the ratio of mean fluorescence of 4C/2C peak should be in the range of 1.95–2.05. Use of endopolyploid standards with a series of peaks (2C, 4C, 8C, 16C, etc.) to control the linearity over wide fluorescence range has also been proposed [16]; however, when using such biological standards, one must be aware that deviations from linearity may reflect staining problems (i.e., non-proportional staining or fluorescence of nuclei of different DNA content or condensation states) rather than instrument-based linearity problems [2].

Daily quality control checks are recommended. In the event a quality control check fails, appropriate remedies are provided by the instrument manufacturer. Typically, problems with peak CVs are addressed by cleaning or priming (fluidics problems) or by laser alignment (optics). Linearity problems within the 3- to 4-fold range of fluorescence intensities may indicate problems with photodetectors that may require service.

3 | INSTRUMENT SETTINGS

Individual types of cytometers differ in design, including the light source, operation speed (both the pressure of sheath fluid and sample release), the diameter of the flow chamber, and light collection optics. These parameters can influence the results and should be considered when comparing data from different instruments, as it may be responsible (along with factors such as different buffers and different sample preparation procedures) for minor differences found among laboratories when analyzing identical material [17, 18]. In the following paragraphs, we will consider several technical aspects related to instrument settings.

When a particle passes through a light beam, it produces a pulse of fluorescence. Three basic characteristics of the pulse may be recorded: pulse height (the maximum intensity), pulse width (the time duration of the pulse), and pulse area (the integral of the fluorescence intensity over time). The exact shape of the pulse depends on the particle shape and diameter, as well as on the light beam diameter [19]. Moreover, in laser instruments, the light intensity decreases from the center to the margins of the beam, depending on the way the laser beam is focused. When the light spot size is much bigger than the particle diameter, and the light intensity over the core stream is more or less homogenous, the whole particle receives even illumination, and peak height is a good measure of the amount of the fluorescent dye. In other cases (bigger particles, smaller spot), peak area ("fluorescence area") is a more reliable parameter. Individual cytometers differ in which of these two parameters is used. A combination of the pulse parameters (i.e., pulse analysis) can be used to discriminate between single nuclei and doublets or triplets (two or three particles coming to the measuring point together, respectively), which may be mistaken for higher ploidy nuclei. This is essential when searching for less numerous cells of different ploidy levels, such as in studies of the cell cycle, endopolyploidy, and particularly unreduced gametes [5, 20]. In summary, researchers should always check the parameters and performance of their particular cytometers and provide details about the cytometer used and methods applied (parameters, gating strategy, doublet discrimination if applied) when publishing the resulting analyses.

Apart from fluorescence, two scatter parameters are available in laser-based instruments. Forward scatter (FSC), a measure of light scattered from the particle in the forward direction, which reflects particle size, is of little use with plant nuclei due to their small size of several micrometers and only slight differences between them (doubling particle volume leads to only a 1.26-fold increase in the particle diameter). In contrast, it is advisable to record side scatter (SSC), a measure of light scattered in a sideways direction, which reflects particle shape and internal structure (complexity). Availability of the SSC parameter might be one of the important factors when choosing a new instrument or light sources. Gating fluorescence versus SSC allows discriminating between intact and damaged nuclei, fluorescent debris, among other particles (e.g., fig. 3 in Sliwinska et al. [2]).

Another parameter that is worth considering is time. In most flow cytometers, it is possible to display how a given parameter varies with time. In most applications, it is important to guarantee that fluorescence is stable throughout the duration of measurement. If there are fluctuations of fluorescence due to disturbed flow, cleaning and realigning the instrument may be necessary. If the user does not let the sample incubate (stain) long enough, one can notice a slight increase of fluorescence in the first minutes of analysis until the stain is saturated and peak means stabilize. This results in a left-skewed fluorescence peak with a large CV, which could be erroneously attributed to other causes (e.g., degraded material or staining inhibitors). Also, if there is some deterioration of the sample over time, causing decreased fluorescence (due to some cytosolic compounds), it can be detected with a cytogram of fluorescence versus time. Lastly, fluctuations visible on the fluorescence versus time cytogram may indicate other technical issues, such as instability of the laser beam due to overheating or other laser malfunctions.

The diameter of the core stream increases with the square root of the sample flow rate, expressed as volume of sample per time (µL/s or µL/min) [12]. Therefore, the sample flow rate should be kept as low as possible to achieve high precision, because the smaller core stream keeps nuclei centered in the laser beam where they receive even illumination. However, a few factors that may compromise measurements with extremely slow flow rates should be considered. First, biological samples may deteriorate over time, especially if the species under study contains high amounts of secondary metabolites or if the tissue is in suboptimal condition (e.g., withered, senescent). Second, long measurement time (in the order of tens of minutes) may lead to

15524930, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.24798 by Cochrane Portugal, Wiley Online Library on [03/11/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

deposition of the nuclei in the sample suspension and/or to nuclei clumping together. Third, depending on a particular cytometer design, an extremely low sample flow rate may lead to irregular sample supply or disturbance of the laminar flow; as with many other details in plant FCM, researchers should test their instrument performance and find the optimal running conditions. Finally, one of the main advantages of FCM for ploidy level/genome size estimations, among other applications, is its high-throughput nature; long running times undermine this. A balance between precision and acquisition time should be found unless extra-high precision is needed (such as simultaneous analyses of samples with slightly different DNA content, as required in analyses of intraspecific genome size variation [21]). Setting the flow rate to analyze lower tens of particles per second typically allows for reasonably low CV values while thousands of nuclei are measured within a few minutes, allowing the researcher to analyze dozens to hundreds of samples per working day.

4 | ACQUISITION SETTINGS (SOFTWARE SET UP)

Several parameters must be set appropriately before acquisition, and those should be reported in publications.

4.1 | Scale

The data can be acquired using either linear or logarithmic scales (Figure 1). The light intensity is "binned," that is, the more or less continuous variable is recorded and digitalized into a discrete number of intervals ("bins"), which are in FCM traditionally called "channels." The *x*-axis of histograms or both axes of scatterplots consist of these channels as units, frequently 256 or 1024. Typically, fluorescence is measured on a linear scale in plant studies, especially when the peak index (i.e., the ratio of the mean fluorescence of two peaks) is the focal variable (e.g., genome sizes and ploidy level inference). Logarithmic scales are useful when processing samples with a large range of peak positions, that is, when minimum and maximum fluorescence intensity within a sample set is very different. These cases may include studies of endopolyploidy when multiple peaks are present, each having twice the fluorescence of the previous one [22], or studies of phytoplankton [8], when very different types of organisms are present. Scatter parameters (FSC, SSC) are typically recorded on a logarithmic scale.

The appearance of histograms is very different on logarithmic and linear scales. On a linear scale, the location of peaks directly reflects the means of those peaks, that is, a tetraploid peak is placed twice as far along the *x*-axis as a diploid peak, and an octoploid peak will be twice as far as the tetraploid peak (Figure 1A). In contrast, on a logarithmic scale, each successive doubling of the peak mean is represented as the same visual increment along the *x*-axis. As a result, the distance between a diploid and tetraploid peak is the same as the distance between the tetraploid and octoploid peaks (Figure 1B). Scaling also influences the shape of a peak. When plotted on a logarithmic scale, diploid, tetraploid, and octoploid peaks with the same number of events and identical CVs will have the same height and width. On a linear scale, the height and width of each peak will vary with their mean: as the mean increases, height decreases, and width increases. Because CVs and event counts are hard to estimate visually from histograms of either scaling, it is important to report these values in text or tables.

The use of either scale types has its advantages and disadvantages. The data can be converted between scales only with some limitations. On a linear scale, some data may not be recorded (low

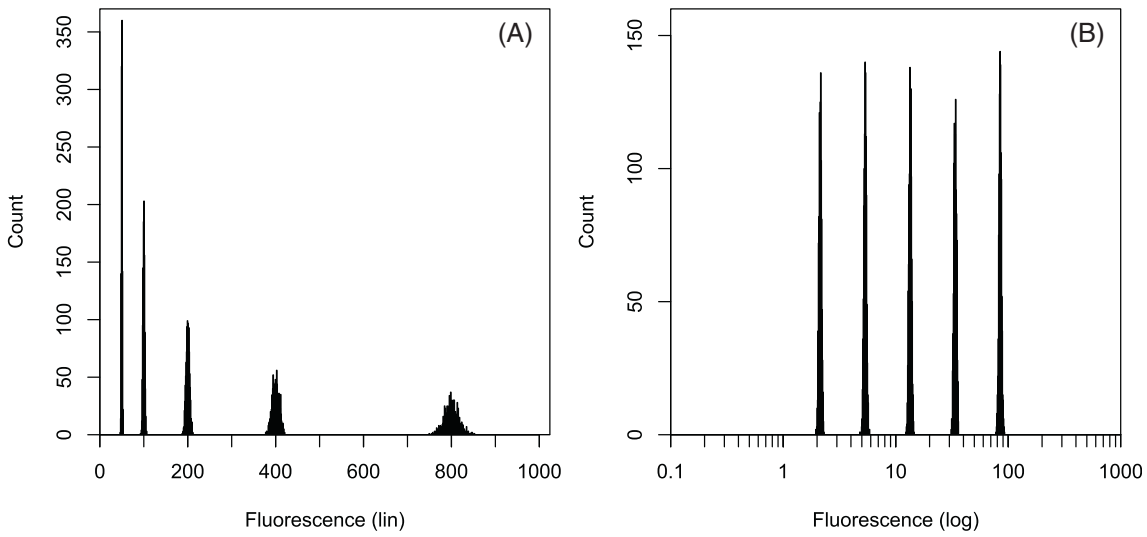

FIGURE 1 Simulated fluorescence histogram on linear (A) and four-decade logarithmic scale (B), 1024 channels each. Each peak reflects a random draw of 1000 particles from a normal distribution having CV = 2.0% and means 50, 100, 200, 400, and 800, respectively (i.e., the subsequent means differing two-fold, similarly to what would be obtained in an endopolyploid sample having nuclei with 2C, 4C, 8C, and so on DNA content). On the linear scale, the peaks on the right appear wider and lower. On the logarithmic scale, height of the peaks is proportional to the number of particles and peaks have the same appearance.

15524930, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.24798 by Cochrane Portugal, Wiley Online Library on [03/11/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

fluorescence intensities below a threshold or high "off scale" intensities). Also, each channel differs from the previous one by *x* while it is *x*-fold on a logarithmic scale. Therefore, at low fluorescence intensities (left side of the axis), one channel on a linear scale spans over several channels on a logarithmic scale, while at higher intensities, the opposite occurs. During recalculation, "missing" data cannot be adequately recovered. This is especially important when converting logarithmic to linear data. While peak means can be recalculated, it is impossible to correctly calculate the SD and thus peak CV. If the CV values are computed from channel numbers on a logarithmic scale, that is, linear-like (as done by some software), the values are usually very low but do not correspond to values that would be obtained on a linear scale. It is especially pronounced toward the right side of the axis where each channel spans a much higher range of light intensities than on the left side. Last, log-amplifiers are known to suffer from deviation from linearity more than linear amplifiers, especially in the first and last decade of the logarithmic axis [13] (in this case, nonlinearity means that the difference between channels is not proportional to the ratio of the actual values). For these reasons, the use of the scale type must be decided before data acquisition, according to the focal variables. If the instrument uses log-amplifiers to record data on a logarithmic scale (as is true in most of older and still in some recent instruments), the use of logarithmic scale is limited to scatter parameters and with fluorescence only when the exact fluorescence intensities or peak index are not the main focus. Genome size measurements and related studies (such as ploidy estimation or base composition) must be done on a linear scale. However, there are some newer instruments that do not use log-amplifiers and acquire all data over wide dynamic range on a linear scale. In this case, either scale type can be used, as logarithmic scale is only used for better data display while all computations are actually done on a linear scale.

4.2 | Number of nuclei

Measurements of fluorescence intensity are subject to errors. This random variation between particles of the same type is usually modeled as a normal distribution (Gaussian curve) characterized by the mean and SD. As acquisition proceeds, peak mean position varies by chance (sampling error), but with an increasing number of particles, it approaches the true value. The corresponding statistics are called standard error of mean (SE). In most analyses, we are interested in the ratio of the mean fluorescence of the sample and the standard. Thus, we have two peaks, both subjected to sampling errors. It is, therefore, crucial to measure enough particles to minimize these errors. Moreover, as the histogram contains not only the G_0/G_1 nuclei that we are most often interested in, but also other particles (e.g., debris, nuclei in other phases of the cell cycle, endopolyploid nuclei), the total number of events needed is higher than the theoretical sum of the peaks of interest.

Only a few studies provide recommendations on the number of events to acquire. Greilhuber et al. [23] suggested that at least 1300 nuclei should be scored for a peak to keep the relative SE (expressed as % of mean) below 0.1% when peak CV is 3%. However, the choice of 0.1% relative SE seems arbitrary, and no rationale is given in the cited book (although it may reflect the standard in other fields). Furthermore, the 0.1% relative SE value may be an overly strict limit, as it is much below the empirical differences between repeated measurement of the same material (i.e., technical replicates, see below). The relationship of the relative SE and the number of nuclei is shown in Figure 2 for different peak CVs. From this and the SE definition, the 0.1% relative SE is achieved at 900 events for CV 3%. However, the SE drops quickly over the first 100 events, after which its rate of decline slows dramatically.

Jan Suda in Greilhuber et al. [23] presented a graph showing an empirical estimation of the required number of events. Several species in several replicates were analyzed with an internal standard. The peak index (i.e., the ratio of sample and standard mean fluorescence) was recorded for each analysis after every 200 events until a total of 20,000 events were acquired. The mean % difference from the final value (based on 20,000 events) was below arbitrary limits of 0.2% and 0.1% after 3000 and 7000 total events, respectively, and in general, the difference decreases quickly at the beginning. In contrast, after ca. 3000, the decrease becomes moderate. However, no details on individual samples are available. The total count of 5000 events for analyses with two peaks of interest (the sample and the internal standard) is also mentioned by Doležel et al. [11], and at least 1000 nuclei per peak are recommended in Sliwinska et al. [2], without any details.

Due to a lack of empirical detail, we repeated the experiment of Suda [23] with a slightly different setup. Leaf tissue of two species was co-chopped in 0.5 mL Otto I buffer, incubated for 10 min, stained with DAPI (4 µg/mL) in 1 mL of Otto II buffer supplemented with 2-mercaptoethanol (2 µL/mL), and analyzed using a Sysmex (formerly Partec) CyFlow Space cytometer, with 365 nm UV-LED as the light source and using the flow rate 0.6–0.8 µL/s (constant within each sample), which delivered 20–30 events per second. The sample was run for ca. 1 min to check the stability of the flow and staining, the initial data were then discarded, and the fluorescence of 15,000 events

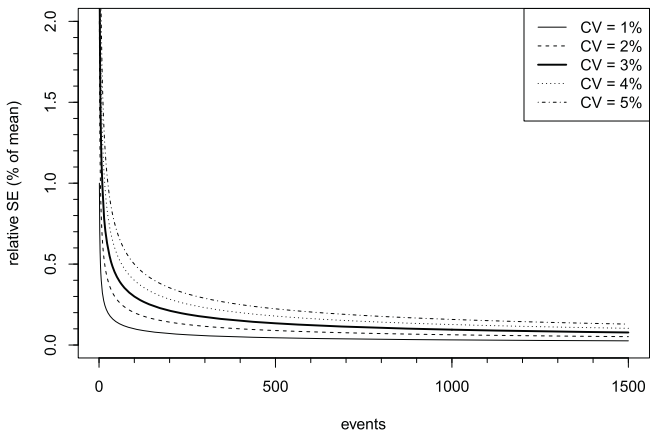

FIGURE 2 Standard error of the mean (expressed as % of mean) as a function of number of events for a given coefficient of variation (CV) when a peak is modeled using the normal distribution.

15524930, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.24798 by Cochrane Portugal, Wiley Online Library on [03/11/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

was recorded. We analyzed the resulting histograms starting at 500 events, and every 50 events thereafter using flowPloidy package [24] in R software [25]. This allowed us to quantify how the peak indices varied over the course of a run, and relative to the number of events (nuclei) recorded within each peak.

Two species pairs were analyzed, 10 replicates (newly chopped samples) for each pair, analyzed within the same day: (i) *Solanum pseudocapsicum* and *Bellis perennis* as the high-quality dataset (both species are widely used as internal standards [15] that produce high-quality peaks and low background; Figure 3A) and (ii) *Sorbus intermedia* and *B. perennis* as the low-quality dataset (the former species belongs to the Rosaceae, a family that is known to contain staining inhibitors [3] which cause higher peak CVs and background; Figure 3B). We also prepared samples containing only *S. pseudocapsicum*, (i.e., without a co-chopped standard), to assess the impact of event count on a single peak mean, as in analyses with an external standard. For the external standard tests, we collected a total of 10,000 events, and calculated the mean peak position over the course of the run, starting at 500 events.

When the full data set (i.e., all 15,000 events) was analyzed, the high-quality samples had a mean peak index of 1.283, and all samples were within 0.5% of this value (range 1.279–1.289, Table 1). The variation among individual samples was quite low, even with only 500 events recorded: only one sample was ever more than 0.5% different from its final value, and it dropped below this threshold once 1200 events had been recorded (Figure 4A). At this point, samples had between 308 and 509 nuclei recorded per peak. Quality continued to improve as events accumulated, with all samples within 0.2% of their final values by 4200 events. CVs stabilized to within 0.5% of their final value by 2500 events.

The low-quality samples had a mean peak index of 1.30, with individuals varying by as much as 1.76% (range 1.277–1.321, Table 1). Variation among samples was somewhat higher than in the high-quality data set. Nine of 10 samples were within 0.5% of their final

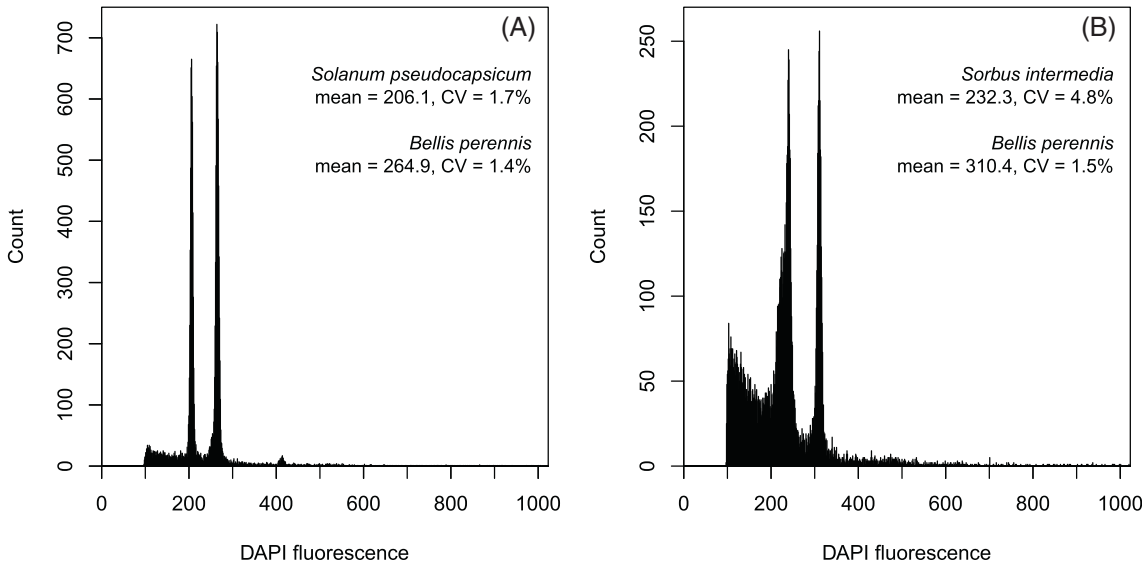

FIGURE 3 Examples of histograms of DAPI fluorescence. (A) High-quality dataset (*Solanum pseudocapsicum* + *Bellis perennis*). (B) Low-quality dataset (*Sorbus intermedia* + *B. perennis*).

TABLE 1 Histogram analysis summary for 10 replicates.

Experiment	Peak index	Events per peak	% of total events	Peak CVs (%)
High quality	1.28 ± 0.003	4424.7 ± 774.0	29.5 ± 5.2	1.5 ± 0.1
		5439.8 ± 852.0	36.3 ± 5.7	1.4 ± 0.1
Low quality	1.30 ± 0.01	3328.3 ± 691.5	22.2 ± 4.6	3.2 ± 1.0
		2729.0 ± 525.9	18.2 ± 3.5	1.4 ± 0.1
	Peak mean			
External standard	51.36 ± 3.50	5691.4 ± 762.6	56.9 ± 7.6	1.8 ± 0.4

Note: Values are mean ± standard deviation. The high-quality experiment included *Solanum pseudocapsicum* (first row) and *Bellis perennis* (second row); the low-quality experiment included *Sorbus intermedia* (first row) and *B. perennis* (second row); the external standard experiment included only *S. pseudocapsicum*. Peak counts and coefficients of variation (CVs) are reported for each species separately. The results from the high and low quality experiment are based on 15,000 total events; the external standard experiment is based on 10,000 total events.

15524930, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.24798 by Cochrane Portugal, Wiley Online Library on [03/11/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

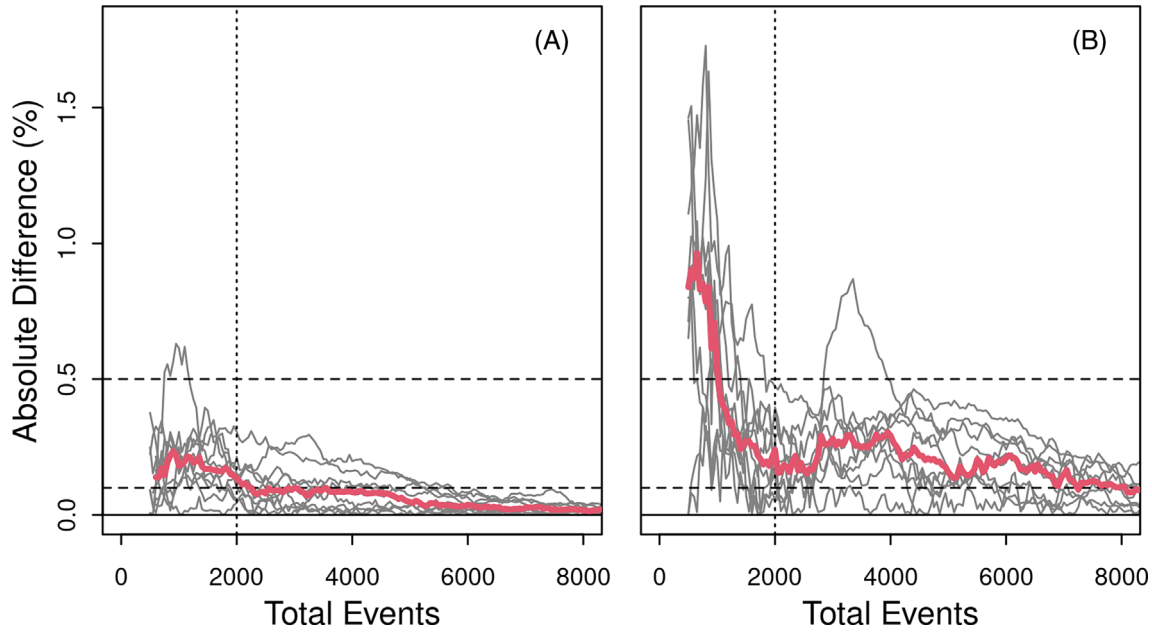

FIGURE 4 Accuracy of peak index as a function of the number of events recorded. The *y*-axis shows the percent difference between the peak index calculated using a given number of events (shown on the *x*-axis) compared to the peak index calculated from 15,000 events. Gray lines show individual samples; the red line is the median value from all 10 samples. Horizontal lines show 0.1% and 0.5% absolute difference. The vertical line indicates 2000 events, our recommended minimum number of events needed for low-stringency applications like ploidy analysis. (A) High-quality data (*Solanum pseudocapsicum* and *Bellis perennis*). (B) Low-quality data (*Sorbus intermedia* and *B. perennis*). [Color figure can be viewed at wileyonlinelibrary.com]

values by 1800 events (corresponding to 173–541 nuclei per peak), with the last sample reaching that threshold at 4000 events, at which point it had 1053 and 550 nuclei in the first and second peaks respectively (Figure 3B). CVs stabilized within 0.5% of their final value by 2800 events.

The absolute position of a single peak in the external standard experiment is more variable than the peak indices in the internal standard experiments. For our single-sample, external standard experiment, the mean peak position for 10 replicates was 51.35, with individual sample means as much as 12% different (range 45.1–56.1, Table 1). Eight of 10 samples were within 1% of their final peak mean value after 2000 events were recorded, at which point they had 931–1389 nuclei in the peak (Figure 5). At this point, the CVs of all samples were below 1.75%, and 7 of 10 replicates remained stable within 0.2% of their final value for the rest of the run. The CVs of the other three replicates slightly increased gradually over time, but remained below 2.4%. The slower convergence of the single peak mean compared to the peak index may be surprising. Closer inspection of fluorescence versus time scatterplots shows that there are slight fluctuations of fluorescence during acquisition, which influence peak means and CVs but little influence the peak index, because the fluctuations of both peaks are synchronized and the index remains stable.

This experiment shows that peak index (for internal standardization) and peak mean (for external standardization) are estimated quite well even from a relatively low number of events. Previous recommendations to acquire at least 1000 events per peak [5] or 5000 events in total [5, 9] (when there are two peaks and a moderate level of debris) are somewhat conservative. In fact, the differences in the peak index estimated over the course of acquisition within a single replicate were almost always less than the variation among final values (after 15,000) of individual replicates. Our results show that lower counts are sufficient for high-quality histograms, or when high precision is not required (e.g., ploidy level analysis). In such cases, we suggest that a relaxed guideline of 2000 total events and 600 events per peak could be adopted without marked loss of precision.

Multiple individuals are often analyzed simultaneously to increase sample throughput, for example when screening a population for rare cytotypes. When using bulked samples, researchers must again consider the number of events. For instance, in a bulked sample of five individuals, every individual contributes with about 1/5 of the sample peak count (not the total count); however, due to a slightly different amount of tissue and random sampling error, the real contribution might be smaller. Therefore, it is advisable to prepare artificial mixed samples at the beginning of the screening, containing the desired number of individuals and one individual different from the rest. Several such samples should be tested to establish a total count value for which it is always possible to identify the signal of the different individual, and for which the number of nuclei in the smallest peak is sufficient to achieve the required precision.

The experiment described above and all considerations on the number of events are independent from the cytometer type, the sample's genome size, or the dye used. Although the samples were stained

15524930, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.24798 by Cochrane Portugal, Wiley Online Library on [03/11/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

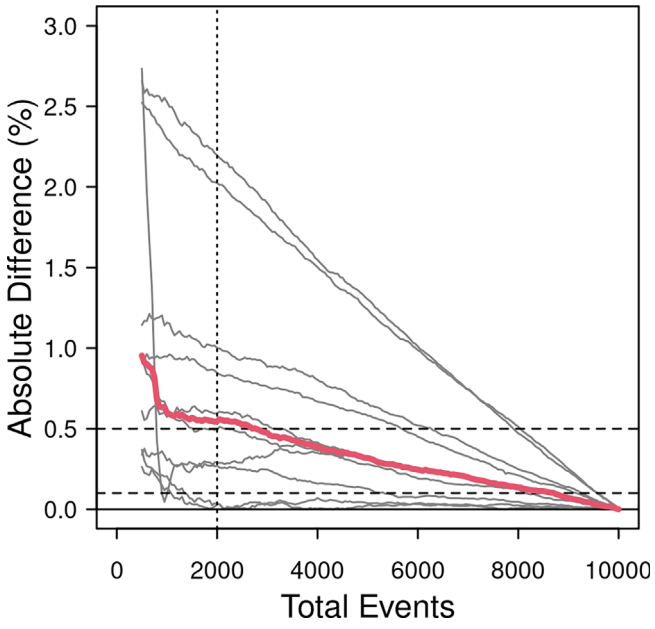

FIGURE 5 Accuracy of peak mean as a function of the number of events recorded. The *y*-axis shows the percent difference between the peak mean calculated using a given number of events (shown on the *x*-axis) compared to the peak mean calculated from 10,000 events. Gray lines show individual samples; the red line is the median value from all 10 samples. Horizontal lines show 0.1% and 0.5% absolute difference. The vertical line indicates 2000 events, our recommended minimum number of events needed for low-stringency applications like ploidy analysis. [Color figure can be viewed at wileyonlinelibrary.com]

with DAPI, the results can be generalized to all dyes, including PI. The peak index is usually different when the same species are analyzed with DAPI and PI, and this difference can even be used to calculate AT/GC content of the sample [2]. However, the peak index was not the focal variable here. The experiment studied variation of the peak index observed within a sample, that is, the increasing precision in estimating peak means and peak index as the acquisition proceeds. A cytometer does not record molecules of the dye directly, but only the light intensity converted by photomultipliers to electric current. Thus, any combination of the study species, the dye, and the cytometer settings leading to peaks with the same CV and position, will result in the same relationship of the number of events and precision of the estimate.

There are other sources of more or less random variation in plant FCM, such as different peak index (i.e., genome size estimate) observed between samples of the same taxonomic unit or between the repeated measurements of the same individual, in the typical case based on the same number of events in all samples. This variation is not covered by the experiment described above and is not a matter of acquisition settings. To cope with these types of variation, the researchers must plan their experiments to ensure randomization between groups of samples (relative to the question asked) and including repeated measurements when needed, as described by Sliwinska et al. [2].

4.3 | Discriminators

For correctly visualizing the data to be acquired, one of the parameters must be set as discriminator (sometimes called leading trigger). Only particles having values above the minimum or below the maximum threshold (the latter not available in all cytometers) are recorded while the others are "invisible," that is, not recorded. The detection capacity of the flow cytometer is so high that without discriminators such background noise from low fluorescence signals (such as autofluorescent debris) would constitute the majority of data. This would be especially undesirable if a pre-set total count of events is used to define run time. The abundant low-fluorescence signal could also make nuclei peaks inconspicuous on a histogram if the *y*-axis is scaled to capture total events. In genome size studies, fluorescence is typically used for triggering, and low fluorescence signals are discarded. However, the limits should be chosen with caution. First, it should leave at least several "empty" (i.e., containing only debris) channels before the lowest-fluorescence (leftmost) peak, so that the calculation of the peak mean and CV are not affected by triggering. If histogram modeling is used for histogram analysis (see Chapter 5), some debris in this range is also required for proper model fitting. Second, when working with yet unknown material, the limit should be set so that half of the fluorescence of the putative leftmost sample peak is displayed to ensure that there is no other peak. This is important in groups that exhibit some level of endopolyploidy or a high number of cells in G_2 phase, in order not to overlook the leftmost peak; note that the number of nuclei in the first peak can be smaller than in the subsequent peaks.

4.4 | Peak position and number of channels

Due to huge variation in plant genome sizes (ca. 2400-fold; [26]) and, consequently, the need to use several different standards, peak positions must be adjusted before acquisition according to the genome size of the material under study. This is done by the parameter called voltage or gain by different manufacturers (but note that "gain" may have also a more specific meaning, which is post-acquisition amplification of the signal). Setting higher values means higher sensitivity of the photomultipliers for the given parameter, resulting in peaks shifted more to the right along the *x*-axis (but this may also cause higher background noise). In combination with the number of channels and type of axis, peak position influences the appearance of graphical outputs (Figure 1A), as described above.

The acquired data are binned, that is, divided into channels. The relative difference of mean fluorescence between the neighboring channels decreases with the channel number (for example, it is 2% between channels 51 and 50 while it is 0.2% between channels 501 and 500). Peaks of constant CV (i.e., constant SD, which defines peak "width") are concentrated in fewer channels toward the left margin of the linear scale (Figure 1A). For data comparability, it is necessary to use the same settings during one experiment. Depending on the instrument, 256–1024 channels are usually used for histograms. A

15524930, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.24798 by Cochrane Portugal, Wiley Online Library on [03/11/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

higher number of channels improves resolution, but a lower number may be advantageous for statistical analysis of histograms (histogram modeling, see below) because it produces a smoother curve. For 2D plots, recalculation to a lower number of channels is sufficient, for example, 256 × 256, as is often used.

For genome size measurements and ploidy level analysis, a linear scale should be used, and the peak of the standard positioned (by adjusting the voltage/gain) to ca. 1/5 of the distance from the left end of the *x*-axis. This setting and thus the approximate position of the standard peak should be fixed for all samples within an experiment. The appearance of standard peaks provides a visual confirmation that run conditions are stable, which would be impossible if the peaks were located in very different positions (see Figure 1A). We expect the standard peak to be similar among samples, while the quality of the test samples may vary. In case the genome size of the test sample is much lower than that of the standard, the peak of the standard can be localized at a higher channel (e.g., closer to the midpoint of the histogram) to keep the lowest expected sample peak approximately 1/5 of the way from the left margin, so long as it is kept constant throughout the whole study. Peak positions in lowermost ca. 10% of the *x*-axis (e.g., below 100 on a 1024-channel scale) should be avoided. First, some events at the low end of the peak might be lost (they might be cut off by the threshold, see above), and second, the exaggerated tall, narrow shape of peaks with a low mean can make them difficult to compare to peaks with higher means (which appear wider and must consist of many more events to get similar height, Figure 1A).

5 | RAW DATA ANALYSES

Flow cytometric studies of plant nuclei are dominated by applications that involve estimating nuclear DNA content based on the fluorescence of DNA-selective fluorochromes [2]. In this context, a key component of data analysis is the fluorescence histogram. The fluorescence of particles is plotted along the *x*-axis, and the number of particles on the *y*-axis. Histogram peaks represent groups of particles with similar fluorescence. Ideally, these are groups of nuclei with the same genome size at the same cell cycle stage (i.e., G_0/G_1 and G_2). However, the histogram also includes other classes of particles, such as: S-phase nuclei, with DNA content ranging from 2C to 4C; damaged and aggregated nuclei, with irregular shapes and fluorescence corresponding to DNA content ranging from less than 1C to 6C or more; and cellular debris, with various physical and fluorescence properties (Figure 6) [27].

The role of the analyst is to accurately estimate population parameters for the particles of interest. The mean fluorescence for a group reflects the DNA content of individual particles: that is, the genome size of nuclei. If a peak is visualized as a curve, the area under the curve determines the number of particles in a group, providing estimates of the number of nuclei in different cell cycle stages [2]. The primary challenge in histogram analysis is that the different groups of particles overlap on the plot. Therefore, the analyst needs to isolate the nuclei of the group of interest from the other groups to get accurate estimates. There are two general approaches to address this: gating and model-fitting.

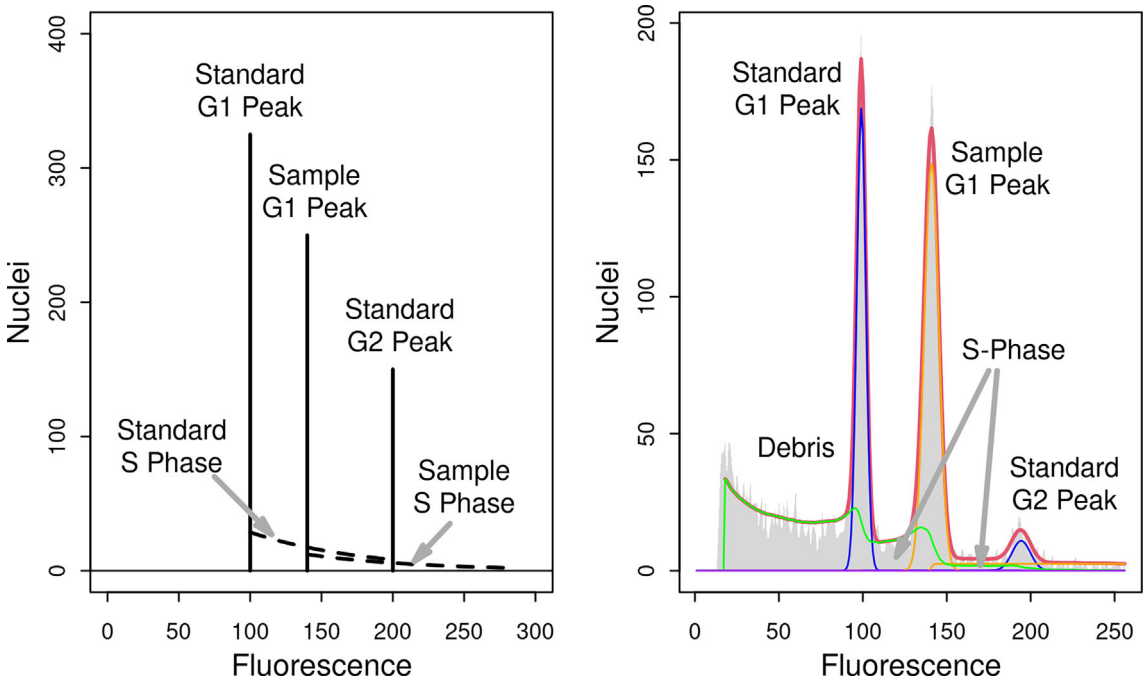

FIGURE 6 Left: an ideal cell cycle histogram for a sample and co-chopped standard. Right: an empirical histogram analyzed by flowPloidy software [24]. [Color figure can be viewed at wileyonlinelibrary.com]

15524930, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.24798 by Cochrane Portugal, Wiley Online Library on [03/11/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

5.1 | Gating: General concepts

Gating is broadly defined as the process of defining a region on a scatterplot or histogram that either includes or excludes unwanted events (debris, aggregates, etc.) so that the subsequent analysis uses mostly events of interest (e.g., nuclei). Defining a region on a scatterplot for gating purposes is very common in FCM analysis. It is often used as a preliminary step in histogram analysis for DNA content (see below). Other examples of its use include signal pulse analysis, that is, the separation of single particles from aggregates (Figure 7), and defining a set of particles to be sorted (although this is rarely used with plant samples, except for specialized analyses such as sorting chromosomes [28], nuclei types within pollen grains [29], or protoplasts [10]). The parameters used depend on the application.

The isolation of events of interest can, in some applications, involve complex analytical approaches involving pattern recognition, clustering analysis, and multiple parameters [30]. In the context of measuring plant nuclei fluorescence, the process is typically more straightforward. Using relatively few parameters, the analyst defines one or more "gates": sets of rules based on measured traits that place events either within or outside the group of interest. In practice, this can involve multiple steps, in each of which a subset of events is identified for exclusion or inclusion based on a single parameter (on a histogram) or by a pair of parameters (on a scatterplot); the subsets defined in each step are then combined using Boolean operators, leading to a final set comprising the events of interest.

Steps in the gating process can be manual or automated. Most people are familiar with manual gating: polygons are drawn "by hand" on scatterplots based on the user's judgment about the appropriate limits for event clusters (Figure 7C) or bars are placed around peaks on a histogram, representing best-guess limits of a presumably Gaussian curve. A simple common example is the process of selecting a region of interest containing nuclei on a scatterplot of two parameters (e.g., fluorescence and SSC) to "gate out" debris particles. Once debris is gated out, the fluorescence data for the remaining particles are displayed on a histogram on which peaks of interest are delimited and

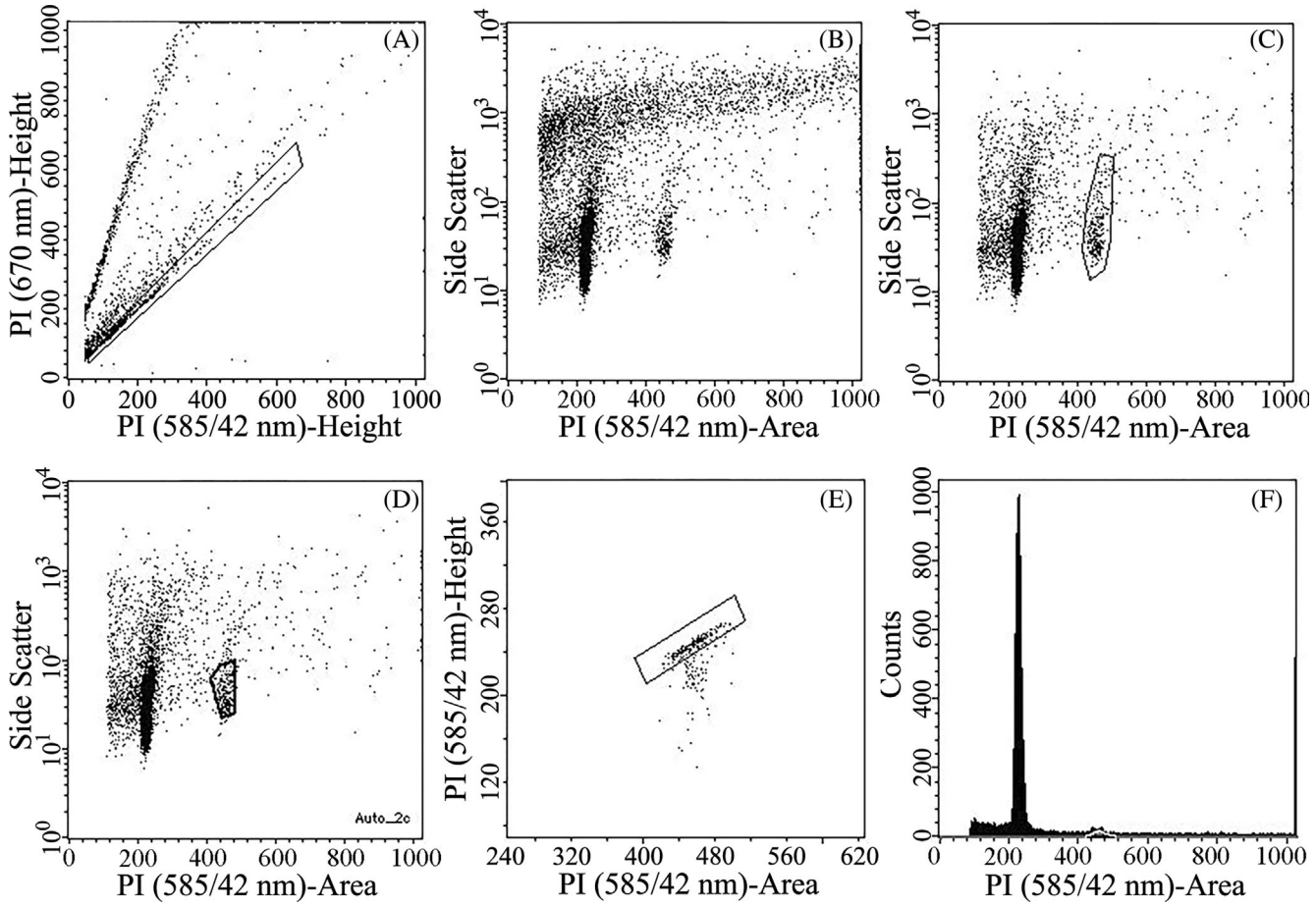

FIGURE 7 An example of multi-step gating used to isolate 2*n* nuclei in a pollen sample. (A) Selection of nuclei region on a FL2-H (585 nm) by FL3-H (670 nm) scatterplot, used to gate out debris. (B) Scatterplot of FL2-A versus SSC for the same sample prior to gating. (C) and (D) show the same data after gating as in (A), with a region draw around the 2C nuclei either (C) manually or (D) using an automated cluster identification tool ("Snap-to" regions, CellQuest Pro, BD Biosciences). (E) 2C nuclei data as selected by the combination of gates in (A) and (C), displayed on a height versus area plot for pulse analysis (aggregate removal). The rectangular region contains only single nuclei. (F) Histogram of ungated data with the small peak outline in white showing the events of interest, single 2C nuclei, which are only a part of the 2C peak.

15524930, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.24798 by Cochrane Portugal, Wiley Online Library on [03/11/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

measured. An example of a more elaborate process involving debris gating using two fluorescence wavelength ranges and aggregate gating is illustrated in Figure 7.

Scatter plots of fluorescence versus SSC are often used to isolate stained nuclei from debris. While fluorescence provides an estimate of DNA content, SSC is related to the geometric complexity of the particle: smooth spheres will produce lower SSC values than comparably-sized particles with rough surfaces and complex shapes. Plotted together, these two variables allow the analyst to distinguish between nuclei, which have relatively simple shapes, and irregular debris particles with similar fluorescence but higher SSC values.

The use of two fluorescence detection ranges (e.g., 585 versus 670 nm; Figure 7A) can also be very useful in excluding some debris types, taking advantage of differences in autofluorescence versus stain fluorescence properties [31, 32]. However, not all flow cytometers support this approach. In samples stained with more than one fluorochrome and in the special case of fluorescently tagged lines, appropriate parameters are the fluorescence emission ranges. For unstained particles, as well as stained, morphology-based characteristics such as density (reflected in SSC) and size (FSC and pulse width) are often used because the morphology of most biological structures tends to be fairly constrained relative to other particles (Figure 8). Signal pulse analysis for aggregate discrimination [19, 33], also discussed in [5], uses plots of fluorescence area versus height (intensity) or width (time of flight) (Figure 7E). Combinations are as varied as the applications.

As mentioned above, most users are familiar with manual gating, in which polygonal regions are drawn around clusters of events on scatterplots (Figure 7A,C,E). The main criticisms of manual gating are that it can be subjective and difficult to apply in a repeatable way, either by the same lab over time or by different labs analyzing the same samples. The effects of this are likely to be less important in high-quality samples and can be reduced by having a single user do all gating, striving for consistency. The use of density plots rather than simple scatterplots may also be helpful for poorly defined clusters (Figure 8). Despite these concerns, manual gating allows for some role of experience and judgment, often necessary when dealing with problematic samples with high levels of debris.

Methods for automated cluster gating exist, ranging from simple two-parameter density gating (Figure 7D) to more complex options, many of which are freely available [34]. This still appears to be little used in plant FCM, perhaps because traditional manual gating seems straightforward for many applications. However, some consideration should be given to such methods, because of speeding up the gating process, and especially because they reduce subjectivity and make results more reproducible [34], although gate definition is still limited by both the effectiveness of the algorithm and the quality of the data.

Gates can also be defined using a single variable, using a histogram. In the most basic approach, users draw boundaries (limits) around peaks or regions on a histogram based on a subjective assessment of those limits. The boundaries are then used to select subsets of events. This approach is often preceded by a "debris gating" step on a scatterplot so that the data displayed on the histogram is already a subset of events. This method has the advantage that it can be done with the most basic software, such as the one provided with the cytometer or general-purpose statistical or plotting programs. It has the disadvantage of subjectivity, as already discussed, but the impact of this depends on the application. For example, measures of relative means are robust to variations in gating, as are the ratios between means (peak index): ploidy classification is unlikely to depend on gating variation. Even in the case of absolute DNA content (genome size) estimation, acceptable results can be obtained from manual gates, although model fitting may provide greater accuracy in some samples (see next section).

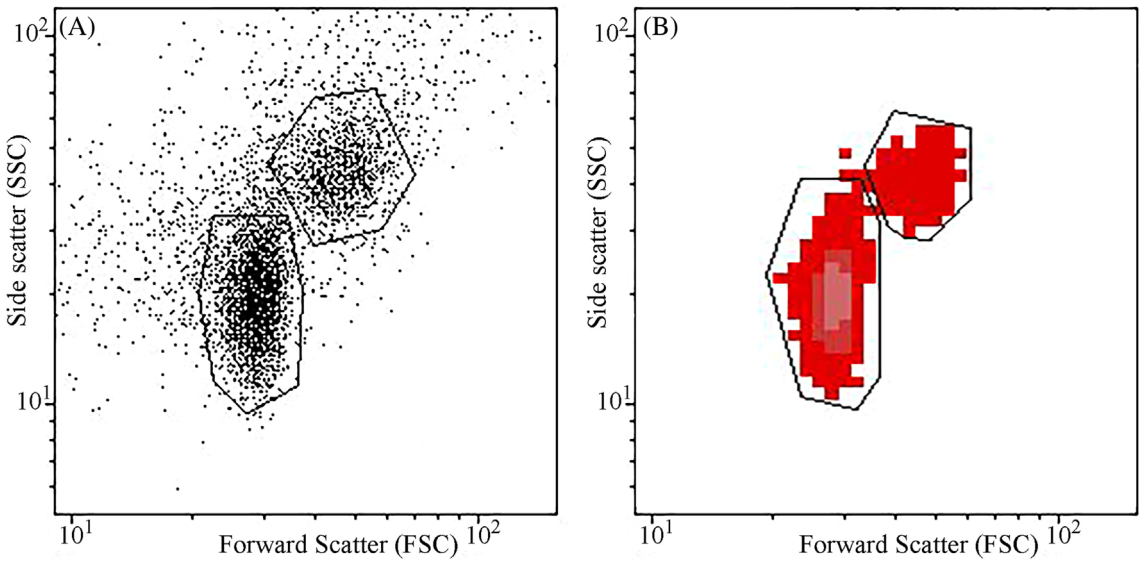

FIGURE 8 Gating of two nuclei types based on scatter properties (SSC reflecting optical density, FSC reflecting size). (A) Shows gating on a simple scatterplot, (B) on a density plot. [Color figure can be viewed at wileyonlinelibrary.com]

15524930, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.24798 by Cochrane Portugal, Wiley Online Library on [03/11/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

In contrast, event counts and proportions are more affected by subjective gating choices, especially in the case of rare events. Finally, while this is not a concern for the conscientious experimenter, it is worth stating that there is a potential for abuse of manual peak gating as it relates to quality control measures: event numbers per peak can be inflated, and CVs underestimated by either widening or narrowing, respectively, the peak limits. Therefore, for gating on scatterplots (cytograms) and on histograms, the strategy should be consistent within an experiment and described in detail in manuscripts, preferably with figures.

5.2 | Histogram modeling

The second approach to isolating nuclei for analysis is based on fitting statistical models to histogram data. This approach uses theoretical expectations of the characteristics of different event categories (e.g., nuclei, debris, aggregates) to create nonlinear models [35]. These models can then be fit to the histogram data; the model fitting process produces estimates of the parameters of interest. The model-based approach provides more objective, repeatable, and in some cases more precise parameter estimates than gate fitting. However, it can be challenging to apply to particularly noisy samples. Last but not least, the modeling approach may be faster than manual gating due to automation of some steps.

Effective histogram modeling builds on theoretical models describing the distribution of the different kinds of particles present in a sample [35]. The DNA content of nuclei, including the G_0/G_1 nuclei of the sample and internal standard, and the G_2 and endopolyploid nuclei if present, are straightforward in this regard. Variation in the actual DNA content between cells is small, and observed differences are due mainly to the accumulation of errors in fluorescence detection. According to the central limit theorem, we expect this situation to produce normally distributed observations. As such, only two parameters are necessary to fit models to these data: the mean and SD of the peaks. Furthermore, the means of G_2 and endopolyploid nuclei peaks are constrained to be a multiple of the G_0/G_1 mean (although some error is permitted to account for minor deviations from strict linearity). However, in some samples, peaks may be skewed (non-Gaussian) due to uneven staining of nuclei, or the presence of aneuploidy [27]. This can impact the estimation of means and proportions of cells in G_0/G_1, G_2, and S-phase, and users must be aware of some inaccuracy of the results (although the same challenge is present with manual gating). If the problem arises from uneven staining, this is ideally dealt with through sample preparation optimization (i.e., buffer selection, tissue homogenization method, stain duration, etc. needs to be validated for each species and tissue type, as discussed in Loureiro et al. [4]).

The other groups of particles require more esoteric functions to model them. For example, S-phase nuclei, spanning the range between G_0/G_1 and G_2 peaks, can be fit with a variety of broadened polygons. Aggregates, clusters of two or more nuclei stuck together, are fit with equilibrium equations, functions of the abundance of smaller particles. This approach generally assumes that aggregation is a random process where all particles are equally likely to aggregate [27, 35]. For systems in which certain types of aggregates may be more common, for example, pollen in which male germ units form [5], this assumption is violated, and pulse analysis may be more appropriate for aggregate correction [20].

Several models are available for debris, including phenomenological exponential curves and mechanistic models built on the assumption that debris is composed primarily of variously fractured nuclei. Models that assume that debris consists primarily of cut and damaged nuclei may not perform well for samples with large amounts of exponentially distributed debris [27]. As a general rule, estimates obtained from samples with high debris are expected to have a higher error associated with them [27], and debris reduction at the sample preparation stage should always be a priority.

Despite the variety of options available for modeling some of the histogram components and keeping in mind the caveats in the preceding paragraph, in practice, the resulting parameter estimates are not very sensitive to these decisions. The S-phase component is usually relatively small, and estimates derived from simple rectangular models are usually very close to the results from using more complex polygons. The different debris models do not usually alter mean estimates, although the differences in cell counts can be more marked.

Once the necessary model components are identified, fitting the model to the data is a routine statistical procedure (nonlinear least-squares regression). The details are described thoroughly by Bagwell [35]. With model-fitting software such as flowPloidy [24], FloJo (BD Biosciences), or ModFit (Verity Software), the model-fitting procedure is completed by the program, and the researcher is not required to conduct the statistical calculations themselves. In histogram modeling, the Reduced Chi-Square Statistic (RCS) can be used to indicate how well a model fits the data for the particular sample and can be helpful in making model choices [27, 35]. Still, it is not a statistical test, only a guide to interpretation. Histogram models that fall within the expected RCS range of 0.7–4.0 are generally acceptable [35]. Histograms with higher RCS values (which indicates a poor fit) should be inspected to ensure the model provides a sensible representation of the data, and whether the model may be improved by adjusting its parameters (such as which debris model and linearity constraints to use). However, a high RCS value on its own does not invalidate an otherwise reasonable analysis. It is particularly important not to misinterpret the RCS statistic as a measure of the accuracy of the DNA content (peak position or peak index, in general) estimate. It only estimates how well the model fits the data, it does not reveal anything about the quality of those data. Quality assessment of the DNA content estimate should still be based on the parameters discussed earlier (CV, nuclei number).

While model fitting provides an objective, repeatable, and precise method for analyzing histogram data, it cannot completely eliminate the need for manual gating. Prior manual gating out of some debris may be required in some low-quality plant samples. One should be aware that this will impact the model fitting procedure. This is more likely to affect nuclei counts and proportions than mean estimates. Using signal pulse analysis to exclude some doublets before histogram modeling will bias the modeling of aggregates [27], and the use of

15524930, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.24798 by Cochrane Portugal, Wiley Online Library on [03/11/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

both approaches in combination is not recommended. Gating that excludes too much low fluorescence debris, such as setting a threshold too close to the G_0/G_1 peak, will affect debris modeling by depriving the algorithm of information about the distribution of debris in the low range. Similarly, truncating the high end of the fluorescence range just above the G_2 peak may affect aggregate modeling [27]. In general, when gating data prior to model fitting, it is best to take a light touch. Whereas in a gating-only approach, one wants to remove as much debris as possible with gates, when fitting a model to the histogram, one only needs to gate out enough debris to allow the algorithm to fit the data.

An important consideration in applying these models is that they were developed from a single mammal species (*Homo sapiens*) and do not account for the variation encountered in tissue samples from different organs of many different species. This is most evident in the large quantity of debris encountered in some plant species, much of which may be of the exponentially distributed type, not modeled well as cut nuclei. In addition, models developed for a species without cell walls or secondary compounds are not always sufficient for application to plant samples. Consequently, while model-fitting can provide superior accuracy and objectivity in many cases, manual gating is necessary when dealing with particularly challenging tissue preparations.

5.3 | Software

We do not propose to evaluate the various software options available for sample analysis. However, there are some considerations to keep in mind. Software that can analyze data from many sources (instruments) can be a powerful tool for review purposes in this age of data transparency and open-access databases. Some software packages associated with specific instrument manufacturers may lack this flexibility. Open-source software packages, such as those using the R programming platform [25], are free and available to anyone and therefore can be used by those who may wish to analyze and examine FCM data files but who do not have access to a cytometer and its associated software. Open-source programs can also be customized and adapted by users to better suit their requirements. We are currently only aware of one open-source program for histogram modeling, flowPloidy [24], which was designed with analysis of plant nuclei histograms in mind.

6 | CONCLUSIONS AND BEST PRACTICE RECOMMENDATIONS

- Keep your instrument in perfect condition (fluidics, optics alignment) and monitor its performance daily using calibration beads or standardized biological material with known properties (fixed chicken or trout erythrocytes, plant standards of superior quality, such as *P. sativum*).
- Apart from fluorescence, also record time parameter (to monitor fluorescence stability over the course of a run) and SSC, if available (to allow discrimination of intact nuclei from damaged ones, fluorescent debris, etc.). If doublet discrimination (i.e., pulse analysis) is intended based on combination of peak area, height, and width, ensure that the respective parameters are recorded.
- Use a linear scale for fluorescence, especially for genome size/ploidy level analysis. Use a limited range of fluorescence (differences among peaks not more than 4-fold) to ensure linearity. Check the linear response of your cytometer in the selected range using calibration beads or other material known to contain clusters of particles (e.g., chicken and trout erythrocytes) or plants showing prominent G_2 phase peaks; use of endopolyploid plants showing multiple peaks to monitor linearity over the wide range of genome sizes is a matter of debate and is not recommended.
- For genome size measurements of samples containing two major peaks (i.e., a sample and a standard) and low to moderate level of debris, acquire at least 600 nuclei per peak and/or 2000 events in total. When the very high precision is needed, or to be in conformity with older recommendations in the literature, 1000 nuclei per peak or 5000 events should be acquired. The total number of events (especially if this is used to define the run time) should be increased in particularly noisy samples or when multiple peaks are present (such as in highly endopolyploid species) to ensure that enough events for peaks of interest are recorded.
- Use discriminators to discard the low fluorescence debris, especially if the total number of events is used to define the run time. However, always keep at least several "empty" (i.e., containing only debris) channels before the leftmost peak to allow accurate peak detection and debris modeling.
- The approximate position of the standard peak on fluorescence axis should be fixed during the experiment to allow visual comparison between samples.
- No peak should be located in the lowest 10% of the fluorescence axis (using linear scale). Peaks at the left margin always appear tall and narrow, which may mask their relatively high CVs and/or relatively low nuclei counts, compared to peaks more to the right.
- Consider using histogram modeling for analysis of raw data. If some manual gating is applied (e.g., due to overly high level of debris or in specialized cases such as search for unreduced gametes or other cells of different ploidy levels), the gating strategy should be clearly described in articles, preferably supported with figures illustrating the typical gate positions.

AUTHOR CONTRIBUTIONS

Petr Koutecký: Conceptualization; writing – original draft; writing – review and editing. **Tyler Smith:** Writing – original draft; writing – review and editing; formal analysis. **João Loureiro:** Conceptualization; writing – original draft; writing – review and editing. **Paul Kron:** Writing – original draft; writing – review and editing.

ACKNOWLEDGMENTS

We are grateful to Petr Šmarda for discussion and comments on the earlier version of the manuscript. The study was supported by the Natural Sciences and Engineering Research Council of Canada grant RGPIN-2020-05652 (P. Kron), by Project CULTIVAR financed by the Programa Operacional Regional do Centro 2014–2020

15524930, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.24798 by Cochrane Portugal, Wiley Online Library on [03/11/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(Centro2020)–CENTRO-01-0145-FEDER-000020, and by Centro de Ecologia Funcional (FCT/UIDB/04004/2020) (J. L.).

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The list-mode data files are available at http://flowrepository.org/id/FR-FCM-Z5BC.

ORCID

Petr Koutecký https://orcid.org/0000-0002-3455-850X
Tyler Smith https://orcid.org/0000-0001-7683-2653
João Loureiro https://orcid.org/0000-0002-9068-3954
Paul Kron https://orcid.org/0000-0002-1734-5019

REFERENCES

1. Galbraith DW, Loureiro J, Antoniadi I, Bainard J, Bureš P, Cápal P, et al. Best practices in plant cytometry. Cytom Part A. 2021;99:311–7.
2. Sliwinska E, Loureiro J, Leitch IJ, Šmarda P, Bainard JD, Bureš P, et al. Application-based guidelines for best practices in plant flow cytometry. Cytom Part A. 2022;101:749–81.
3. Čertner M, Lučanová M, Sliwinska E, Kolář F, Loureiro J. Plant material selection, collection, preservation, and storage for nuclear DNA content estimation. Cytom Part A. 2022;101:737–48.
4. 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. Cytom Part A. 2021;99:318–27.
5. Kron P, Loureiro J, Castro S, Čertner M. Flow cytometric analysis of pollen and spores: an overview of applications and methodology. Cytom Part A. 2021;99:348–58.
6. Talhinhas P, Carvalho R, Loureiro J. The use of flow cytometry for fungal nuclear DNA quantification. Cytom Part A. 2021;99:343–7.
7. Čertnerová D, Galbraith DW. Best practices in the flow cytometry of microalgae. Cytom Part A. 2021;99:359–64.
8. Petersen TW, Brent Harrison C, Horner DN, van den Engh G. Flow cytometric characterization of marine microbes. Methods. 2012;57:350–8.
9. Mitsumoto K, Yabusaki K, Kobayashi K, Aoyagi H. Development of a novel real-time pollen-sorting counter using species-specific pollen autofluorescence. Aerobiologia (Bologna). 2010;26:99–111.
10. 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. Cytom Part A. 2022;101:725–36.
11. Doležel J, Greilhuber J, Suda J. Estimation of nuclear DNA content in plants using flow cytometry. Nat Protoc. 2007;2:2233–44.
12. Shapiro H. Practical flow cytometry. Ed. 4. Hoboken, NJ: John Wiley & Sons; 2003. p. 681.
13. Wood JCS. Establishing and maintaining system linearity. Curr Protoc Cytom. 2009;47(1):4.1-1.4.14.
14. Bagwell BC, Baker D, Whetstone S, Munson M, Hitchcox S, Ault KA, et al. A simple and rapid method for determining the linearity of a flow cytometer amplification system. Cytometry. 1989;10:689–94.
15. Temsch EM, Koutecký P, Urfus T, Šmarda P, Doležel J. Reference standards for flow cytometric estimation of absolute nuclear DNA content in plants. Cytom Part A. 2022;101:710–24.
16. Galbraith DW. Endoreduplicative standards for calibration of flow cytometric C-value measurements. Cytom Part A. 2014;85:368–74.
17. 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 Bot. 1998;82:17–26.
18. Bainard JD, Husband BC, Baldwin SJ, Fazekas AJ, Gregory TR, Newmaster SG, et al. The effects of rapid desiccation on estimates of plant genome size. Chromosome Res. 2011;19:825–42.
19. Wersto RP, Chrest FJ, Leary JF, Morris C, Stetler-Stevenson M, Gabrielson E. Doublet discrimination in DNA cell-cycle analysis. Cytom Part A. 2001;46:296–306.
20. Kron P, Husband BC. Distinguishing 2N gamete nuclei from doublets in pollen using flow cytometry and pulse analysis. Cytom Part A. 2015;87:943–57.
21. Šmarda P, Bureš P, Horová L, Rotreklová O. Intrapopulation genome size dynamics in *Festuca pallens*. Ann Bot. 2008;102:599–607.
22. 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.
23. 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. 67–101.
24. Smith TW, Kron P, Martin SL. flowPloidy: an R package for genome size and ploidy assessment of flow cytometry data. Appl Plant Sci. 2018;6:e1164.
25. R Core Team. R: A language and environment for statistical computing. 2021 Available from: https://www.r-project.org/
26. Pellicer J, Hidalgo O, Dodsworth S, Leitch IJ. Genome size diversity and its impact on the evolution of land plants. Genes (Basel). 2018;9:88.
27. Rabinovitch PS. Practical considerations for DNA content and cell cycle analysis. In: Bauer KD, Duque RE, Shankey TV, editors. Clinical flow cytometry: principles and application. Baltimore: Williams & Wilkins; 1993. p. 117–42.
28. Doležel J, Lucretti S, Molnár I, Cápal P, Giorgi D. Chromosome analysis and sorting. Cytom Part A. 2021;99:328–42.
29. Santos MR, Bispo C, Becker JD. Isolation of *Arabidopsis* pollen, sperm cells, and vegetative nuclei by fluorescence-activated cell sorting (FACS). In: Schmidt A, editor. Plant germline development: methods and protocols. New York: Humana Press; 2017. p. 193–210.
30. Boddy L, Wilkins MF, Morris CW. Pattern recognition in flow cytometry. Cytometry. 2001;44:195–209.
31. Galbraith DW. Simultaneous flow cytometric quantification of plant nuclear DNA contents over the full range of described angiosperm 2C values. Cytom Part A. 2009;75:692–8.
32. Stehlik I, Kron P, Barrett SCH, Husband BC. Sexing pollen reveals female bias in a dioecious plant. New Phytol. 2005;175:185–94.
33. Sharpless TK, Melamed MR. Estimation of cell size from pulse shape in flow cytofluorometry. J Histochem Cytochem. 1976;24:257–64.
34. Verschoor CP, Lelic A, Bramson JL, Bowdish DME. An introduction to automated flow cytometry gating tools and their implementation. Front Immunol. 2015;6:380.
35. Bagwell B. Theoretical aspects of flow cytometry data analysis. In: Bauer KD, Duque RE, Shankey TV, editors. Clinical flow cytometry: principles and application. Baltimore: Williams & Wilkins; 1993. p. 41–61.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Koutecký P, Smith T, Loureiro J, Kron P. Best practices for instrument settings and raw data analysis in plant flow cytometry. Cytometry. 2023. https://doi.org/10.1002/cyto.a.24798

15524930, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.24798 by Cochrane Portugal, Wiley Online Library on [03/11/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License