Analysis of Genome Size of Sixteen Coffea arabica Cultivars Using Flow Cytometry

Josué Ortega-Ortega, Francisco Arturo Ramírez-Ortega, Roberto Ruiz-Medrano, and Beatriz Xoonostle-Cázares
Departamento de Biotecnología y Bioingeniería, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional Av. Instituto Politécnico Nacional, 2508 San Pedro Zacatenco, 07360 CDMX, México

Abstract. Coffee is an important crop worldwide, grown on about 10 million hectares in tropical regions including Latin America, Africa, and Asia. The genus Coffea includes more than 100 species; most are diploid, except for C. arabica, which is allotetraploid and autogamous. The genetic diversity of commercial coffee is low, likely due to it being self-pollinating, in addition, the widespread propagation of few selected cultivars, such as Caturra, Bourbon, and Typica. One approach is the analysis of genome size in these cultivars as a proxy to study its genetic variability. In the present work, genome size of 16 cultivars was assessed through high-resolution flow cytometry (FCM). Nuclear DNA was analyzed using a modified procedure that uses propidium iodide (PI) and 4’-6’-diamino-2-phenylindole dihydrochloride hydrate (DAPI) staining. The C. arabica cultivars investigated possessed a nuclear DNA content ranging from 2.56 ± 0.016 pg for Typica, to 3.16 ± 0.033 pg for ICATU, which had the largest genome size. All cultivars measured using both fluorochromes had greater estimates with DAPI than PI. The proportion of the genome composed of guanosine and cytosine (GC%) among the cultivars evaluated in this study ranged from 37.03% to 39.22%. There are few studies of genome size by FCM of distinct important C. arabica cultivars, e.g., hybrids and artificial crosses. Thus, this work could be valuable for coffee breeding programs. The data presented here are intended to expand the genomic understanding of C. arabica and could link nuclear DNA content with evolutionary relationships such as diversification, hybridization and polyploidy.

Coffee is an important commodity in terms of international trade (The Observatory of Economic Complexity, 2018) and is grown in Latin America, Africa, and Asia, covering a surface of 10 million ha (Mishra and Slater, 2012). Coffee cultivars belong to the genus Coffea (Rubiaceae family), which includes more than 100 mostly diploid species (2n = 2x = 22), except for C. arabica (2n = 4x = 44), which is autogamous and allotetraploid (Noirot et al., 2003a). Most world coffee production comes from C. arabica, of which the most common cultivars are Caturra, Catuai, Bourbon, and Typica (International Coffee Organization, 2018). However, genetic diversity of C. arabica is considered low, due likely to genetic bottlenecks caused by propagation of few individuals from selected cultivars for commercial purposes (Cubry et al., 2008; Lashmeshes et al., 2011). Thus, besides its economic and agricultural relevance, Coffea species and C. arabica cultivars can shed light on the domestication and evolution of this genus through analysis of genome sequence and nuclear genome size, among other approaches (Barre et al., 1996; Herrera et al., 2002). Genome size is congruent with the ploidy level and number of basic chromosome sets of an organism (Huang et al., 2013; Tatum et al., 2006). Nuclear DNA content (C value) is an important trait revealing the correlation between genome size and phenotype. For a haploid genome, the abbreviations of C-value are represented as 1C (Bennett and Leitch, 2005; Greilhuber and Doležel, 2009; Greilhuber et al., 2005a). More than 8500 C-values of plant species have been estimated to date (Bennett and Leitch, 2005; Greilhuber et al., 2005a). The DNA content in angiosperms varies a protocol for purification of nuclei was used to the identification of diverse groups. Moreover, a protocol for purification of nuclei was used to avoid spurious variations and erroneous genome size estimations. The data presented here are intended to expand the genomic understanding of C. arabica and correlate nuclear DNA content as well as base composition with evolutionary relationships such as diversification, hybridization, and polyploidy.

Materials and Methods

Plant material
Leaves were collected from 16 healthy C. arabica cultivars in consideration of its
commodity value, according to International Coffee Organization (2018) and the World Coffee Research (2018). Glycine max (soybean) and Coffea canephora leaves were used as standards with known genome content. Plants were grown in greenhouse at full irrigation and long-day photoperiod (16-/8-h) conditions.

**FCM analysis**

**Nuclei isolation.** This process was carried out essentially as described (Noirot et al., 2005). Basically, sections of young leaves (2 cm²) were macerated in a small plastic petri dish using a sharp razor blade in 400 µL of nuclei extraction buffer (CyStain ultraviolet Precise P, Partec, Göttingen, Germany); 2 mM dithiothreitol, which also helps maintain the integrity of chromatin (Dolezel and Bartos, 2005). Samples were incubated for 2 h in the dark, after which the mixture was filtered through a 30-µm nylon mesh disposable filter (CellTrics; Sysmex, Milton Keynes, UK) and transferred to a 2-mL microcentrifuge tube. To avoid artifactual variations and erroneous genome size estimations, the mixture was centrifuged at 200g for 15 min and the supernatant discarded. Finally, nuclei were treated with 50 µg/mL RNase A (Sigma-Aldrich) and stained. Two different compounds were used for staining, PI (50 µg/mL) (excitation/emission wavelengths: 480–575/550–740 nm) (Sigma-Aldrich) as well as a nuclei staining buffer containing DAPI (excitation/emission wavelengths: 320–385/415–520) (Shapiro 2005) (CyStain ultraviolet Precise P; Partec). Both staining treatments were supplemented with 2 mM dithiothreitol, which also helps maintain the integrity of chromatin (Dolezel and Bartos, 2005).

**Experimental design.** The experimental design was adopted following the recommendations of Clarindo et al. (2012) for nuclei extraction from multiple leaves, from two standards, soybean (2C = 2.30 pg and C. canephora (2C = 1.29 pg) (values obtained in the present work). FCM parameters, such as gain and channel, were determined for each sample in picograms (pg) by multiplying the mean ratio by the 2C value of the standard (Dolezel and Greilhuber, 2010).

**Base composition**

The base pair (bp) composition of 16 cultivars was evaluated. %bp was estimated according to the equation: adenine thymine percentage (AT%) = AT% for internal standard [(mean fluorescence standard DAPI / mean fluorescence sample DAPI) / (mean fluorescence standard PI / mean fluorescence sample PI)] (binding length) (Contreras and Shearer, 2018; Godelle et al., 1993), where AT% of primary standard is 63.6 and the binding length of DAPI is 3.5 (Meister and Barow, 2007). Values reported in Table 1 were calculated as GC% = 100 – AT%.

**Statistical analysis**

To assess for significant differences between C. arabica cultivars genome size values, Fisher’s least significant difference (LSD) test was employed, LSD₀.₀₅ = 0.13356 was calculated to determine the greater limit allowed between each treatment mean to consider whether they belong to the same population. Then, the real difference was calculated between the population average, relative to the LSD₀.₀₅ = 0.13356. A t test was used to compare genome size values for 16 cultivars calculated using both DAPI and PI to determine whether differences were significant using these two fluorochromes. Sigma Plot 14.0 software (Systat Software, San Jose, CA) was used to perform the statistical analysis.

**Results**

**Standardization of nuclear DNA content quantification.** To carry out the quantification of DNA content of the C. arabica cultivars, the G0/G1 peak value of soybean (primary standard) was tuned to fluorescence channel 163. The genome size mean values were obtained in picograms (2C = 23.0 ± 0.030 pg) (Fig. 1). C. canephora also was used as secondary standard (data shown in Supplemental Tables 1–3) to determine FCM parameters; the G0/G1 peak was tuned to channel 92 and the genome size was calculated to 2C = 1.29 ± 0.051 pg (Fig. 1). FCM parameters were determined for each DNA content measurement, based on FCM assessments of standards and samples.

**FCM histograms of C. arabica cultivars.** Mean fluorescence values from 16 C. arabica cultivars showed G0/G1 nuclei peaks in a fluorescence range from 181.7 to 216.91 with nuclei count between 3500 and 5000, four representative cytograms are shown in Fig. 1A–D. In terms of genome content, two groups were identified. In the first group are those with progenitors are the ‘Bourbon’ and ‘Typica,’ whereas the second group includes those that possess some genetic traits from another cultivar (artificial crosses) or species (hybrids), mainly C. canephora and also sometimes C. liberica as well as hybrids (Fig. 2) (World Coffee Research, 2018). The histograms of suspensions of isolated nuclei stained with PI exhibited cv values varying between 4.12% and 5.73%, which suggests that the resolutions of the histograms were appropriate for genome size analysis (Table 1).

**Assessment of genome size in C. arabica cultivars.** Considering the relative G0/G1 nuclei peak of PI fluorescence corresponding to the primary standard (soybean) and to the sample, the mean ratio of 2C values was calculated as a linear relationship between the ratio of 2C value peaks of the sample and standards. C. arabica cultivars possessed a nuclear DNA content ranging from 2.56 ± 0.016 to 3.16 ± 0.033 pg. ‘Typica’ had the smallest genome size, whereas ICATU had the largest genome size (Fig. 3). The nuclear DNA content values of ‘Bourbon’ (2C = 2.86 ± 0.010), ‘Maragogype’ (2C = 2.61 ± 0.049 pg), ‘Pluma Hidalgo’ (2C = 2.66 ± 0.007 pg), ‘Villa Sarchi’ (2C = 2.76 ± 0.065 pg), and ‘Caturra’ (2C = 2.88 ± 0.040 pg) (‘Typica’ genetic family) are congruent with the genealogy of the coffee cultivars analyzed in this study as shown in Fig. 4. The cultivars Catuai and Mundo Novo genome size values (2C = 2.88 ± 0.053 and 2.90 ± 0.031 pg, respectively) were similar to the Brazilian cultivars Catuai Vermelho IAC 15 (UFV 2237 cova 148 ELJ) as well as Mundo Novo IAC 376-4-32 (UFV 2150) Cova 39 (Bourbon Vermelho × Sumatra) reported by Clarindo and Carvalho (2009), Clarindo et al. (2013), and Fontes (2013). ‘Garnica’ (2C = 2.93 ± 0.021 pg) and ‘Garena’ (2C = 3.03 ± 0.035 pg) showed similar DNA content values, whereas ‘Colombia’ (2C = 3.05 ± 0.022), ‘IAPAR 59’ (2C = 3.05 ± 0.080 pg), ‘Costa Rica’ (2C = 3.12 ± 0.125), and ‘Oro Azteca’ (2C = 3.15 ± 0.021 pg) showed a greater DNA content value, possibly due to its hybrid nature (Fig. 4). The genome size values obtained with DAPI were always greater in all cultivars (Table 1). The fluorochromes comparison between DAPI and PI for all cultivars showed significant difference (P < 0.001) in 2C values (Table 1). Genome size differences using both fluorochromes varied in a range of 0.18 to 0.52 (Table 1).

The lowest difference in genome size using DAPI and PI was found in Costa Rica cultivar (0.18), whereas the greatest difference was found in the Maragogype cultivar (0.52). The AT% was determined in the cultivars evaluated in a range of 60.78% to 62.97%. The GC% contents ranged from 37.03% to 39.22%.
Moreover, the genome size with PI was reproducible using two standards (C. canephora and soybean); a linear correlation between the fluorescence and the DNA content with an $R^2 = 0.98$ was found, and a similar slope in all evaluated C. arabica cultivars. This reproducible parameter allowed us to estimate that the nuclear DNA content in picograms is very similar, regardless of which standard was used to compare the DNA content (Fig. 5).

**Discussion**

In the present work, 16 C. arabica cultivars were selected for analysis, according to their global commodity value in different countries. C. arabica lineages analyzed in this study were classified into two groups that agree with previously reported phylogenies obtained through other methods (Anthony et al., 2002; Machado et al., 2017; Steiger et al., 2002). The first group corresponds to the ‘Bourbon’ and ‘Typica’ genetic family; the most important cultivars of this group are ‘Pluma Hidalgo’ and ‘Maragogype’, which are natural mutations of ‘Typica’. ‘Villa Sarchi’ and ‘Caturra’ are natural mutations of ‘Bourbon’, besides hybrids such as ‘Mundo Novo’ (‘Typica’ × ‘Bourbon’) and ‘Catuai’ (‘Mundo Novo’ × ‘Caturra’). These cultivars are associated with high cup quality but are susceptible to the major coffee diseases, such as coffee rust. Due to the presence of this disease, cultivars that show some degree of resistance have been introduced. The second group includes artificial crosses between different C. arabica cultivars, for example, ‘Garnica’ (‘Mundo Novo’ × ‘Caturra’) and their natural mutation ‘Garena’. This also group includes hybrids, such as ‘IAPAR 59’ (‘Villa Sarchi’ × ‘Timor Hybrid 832/2’), ‘Colombia’ (‘Caturra’ × ‘Timor Hybrid 1343’), ‘Costa Rica’ (‘Timor Hybrid 832/1’ × ‘Caturra’), ‘Oro Azteca’ (‘Timor Hybrid 832/1’ × ‘Caturra’), and

### Table 1. Genome size and bp composition of 16 C. arabica cultivars stained with PI and DAPI using soybean (2C = 2.30 pg) as the standard.

| Cultivars      | Fluorescence mean | Nuclei number | CV (%) | 2C genome size (pg) | GC% | AT% |
|----------------|-------------------|---------------|--------|---------------------|-----|-----|
| Typica         | 90.83             | 181.7         | 2173   | 3549                | 5.85| 5.23|
| Maragogype     | 93.79             | 185.49        | 6118   | 5580                | 5.85| 5.65|
| Pluma Hidalgo  | 93.86             | 188.91        | 9193   | 2924                | 5.55| 5.43|
| Villa Sarchi   | 94.19             | 196.07        | 7570   | 3072                | 6.22| 4.21|
| Bourbon        | 94.35             | 202.97        | 5910   | 4504                | 5.88| 5.66|
| Catara         | 95.49             | 204.37        | 7357   | 4738                | 5.44| 6.13|
| Caturai        | 95.61             | 204.83        | 9121   | 5505                | 5.44| 4.44|
| Mundo Novo     | 95.72             | 205.67        | 7229   | 3213                | 5.48| 4.29|
| Garnica        | 95.95             | 208.11        | 4834   | 4474                | 5.65| 5.65|
| Pacamara       | 97.5              | 212.63        | 6267   | 3246                | 5.64| 6.32|
| Garena         | 97.77             | 214.95        | 5950   | 3606                | 5.55| 4.97|
| Colombia       | 98.15             | 216.6         | 6908   | 3354                | 4.87| 5.15|
| IAPAR 59       | 98.42             | 216.91        | 3628   | 2716                | 5.09| 4.96|
| Costa Rica     | 99.12             | 221.9         | 8102   | 3700                | 5.42| 5.73|
| Oro Azteca     | 104.26            | 224.04        | 5237   | 2833                | 5.4| 5.22|
| ICATU          | 104.96            | 224.35        | 6404   | 4896                | 5.09| 6.02|

*Significant difference ($P < 0.001$) based on t test comparing mean 2C genome size determined using DAPI with PI.

bp = base pair; PI = propidium iodide; DAPI = 4',6'-diamino-2-phenylindole dihydrochloride hydrate.

---

Table 1. Genome size and bp composition of 16 C. arabica cultivars stained with PI and DAPI using soybean (2C = 2.30 pg) as the standard.

GC% = 100 – {AT% for standard × [(mean fluorescence standard DAPI / mean fluorescence sample DAPI) / (mean fluorescence standard PI / mean fluorescence sample PI)] (1/binding length)} (Contreras and Shearer, 2018) (Godelle et al., 1993), where AT% of internal standard = 63.6 and binding length of DAPI = 3.5 (Meister and Barow 2007).

---

**Fig. 1.** Representative cytograms of fluorescence intensity by flow cytometry of G0/G1 nuclei from some C. arabica cultivars and standards. Representative cytograms showing G0/G1 peaks with cvs ranging between 3.76% and 4.58%, obtained from propidium iodide–stained nuclear suspensions prepared from leaves of comparison standards; 1) soybean (channel 163.6) (2C = 2.30 pg) and 2) C. canephora (channel 92.85) (2C = 1.29 pg). Four representative C. arabica cultivars are shown. (A) Catuai 3) channel fluorescence 204.83 (2C value = 2.88 pg); (B) ICATU 4) channel fluorescence 224.35 (2C value = 3.16 pg); (C) Oro Azteca 5) channel fluorescence 224.04 (2C value = 3.15 pg); and (D) Typica 6) channel fluorescence 181.7 (2C value = 2.56 pg).
are no cultivars harboring all the resistance genes identified to date (Anthony et al., 2001; Diola et al., 2013); therefore, techniques that increase genetic diversity will be very important for the control or mitigation of coffee rust, among other diseases of this fundamental crop.

The first step in FCM analysis was to define the 2C-values of the standards, a closer value (in this study, 2C = 2.30 ± 0.030 pg) to those reported by Arumuganathan and Earle (1991) (2C = 2.31 pg) and Vilhar et al. (2001) (2C = 2.34 pg) in soybean, which contains AT% = 63.6% (Abreu et al., 2011; Barow and Meister 2002; Meister and Barow, 2007). Dolezel and Greilhuber (2010) suggest that, to be useful as a primary standard, a plant must have, among others, similar, but not identical, genome size to the analyzed plant, and the G0/G1 peaks of the standard should not overlap to the peaks of the sample. In addition, the standard must be easy to use, genetically stable, nuclei must be obtained in enough amounts for analysis, and its genome size must be known with great precision as well as similar AT%. Soybean has all these characteristics and is one of the eight best primary standards according to Praca-Fontes et al. (2011). Besides, considering the occurrence of pseudo variations and consequent inaccurate genome size estimations, it is desirable to use more than one standard. In this work C. canephora (2C = 1.29 ± 0.051 pg) also was employed as secondary standard (data shown in Supplemental Tables 1–3) considering its relatedness to C. arabica. The genome content of C. canephora has been reported ranged from 1.20 to 1.40 pg (Cros et al., 1995; Noirot et al., 2003a, 2003b) with a AT% = 64.46 (Clarindo et al., 2012).

Genome size determination tests depend on different factors, due to its indirect nature, influenced by the GC%, as well as the potential presence of compounds that fluoresce in the same wavelength, thus over or underestimating the experimental values. Considering that the genome of soybean and C. canephora is already known, the closest values to the actual estimate are those obtained with the method using PI. In Fig. 5, genome size values is shown using two standards, in which the estimates of the cultivars based on the comparison with the soybean genome, are greater. As a speculative note, it can be considered that the staining of the cultivars using two different standards could be similar, because they are of the same species. However, soybean, being a different tissue from another species, may have a greater affinity for the dye, associated with inherent properties at the time of sample processing. Despite this, the variation is consistent, which allows us to rely on the estimates obtained.

FCM parameters of 16 C. arabica cultivars, such as G0/G1 nuclei peaks, nuclei count, and CV values are close to 5% (Table 1), were similar to the results obtained by Clarindo et al. (2013). In agreement with Dolezel and Bartos (2005), CV lower than 5% are considered better for
In the presence of inhibitors (Choudhury et al., 2003), a protocol was developed for nuclei isolation to avoid erroneous genome size estimations. It is important to mention that in all analyzed cultivars, the genome size determined by PI resulted in slightly different values compared with the estimated values obtained with DAPI (Table 1). The difference in genome size estimates varied by as much as 0.52 pg. This suggests that as bp composition of the cultivars deviates from the bp composition of the primary standard (soybean, AT% = 63.6%). Some differences in base composition can help to identify those cultivars with artificial hybridization. This highlights to determine with relative ease whether populations are truly different from the rest, more so when standard deviations and means are not useful parameters in this regard. The tool is particularly useful for the study of agronomically important traits. In the case of C. arabica it is important to determine the genetic variability between different cultivars (including genome size) since most of them are not biologically defined species, but rather the result of somatic mutations and conventional breeding. The LSD analysis shown here allowed us to distinguish variations of 0.133 pg in nuclear DNA content. The cultivar with lower DNA content was ‘Typica’ with 2.56 pg, in which 0.133 pg corresponds to 5% sensitivity, which is acceptable for a statistical

Fig. 4. Genealogy of the coffee cultivars analyzed in this study. Different lineages are distinguished in commercial cultivars giving rise to the extant cultivars: Pluma Hidalgo and Typica, Typica and Bourbon, Bourbon and Caturra, and Typica and Robusta. The cultivars indicated with a dashed arrow were obtained as natural mutations (‘Pluma Hidalgo’, ‘Pacamara’, ‘Maragogype’, ‘Bourbon’, ‘Villa Sarchi’, and ‘Caturra’). The Bourbon and Typica cultivars generated ‘Mundo Novo’, which with ‘Caturra’ generated ‘Catuai’, ‘Garnica’, and ‘Garena’, whereas ‘Villa Sarchi’ and ‘Timor’ hybrid 832/2 generated ‘JAPAR 59’. Another identified group resulted from the ‘Typica’ and ‘Robusta’ genetic cross, thus generating the ‘Timor’ hybrid 832/1, which with ‘Caturra’ produced ‘Oro Azteca’ and ‘Costa Rica’ cultivars. The ‘Timor’ hybrids 832/1 and 1343 generated the commercial cultivar Colombia. Cultivars showing partial resistance to coffee rust are highlighted in shaded boxes. International code for Timor cultivars accessions are indicated in parenthesis.

Fig. 5. Linear comparison of DNA nuclear content among C. arabica cultivars using two different standards. Similar slope (\(R^2 = 0.98\)) in all evaluated cultivars was found when nuclei was stained with propidium iodide.
test; however, among many cultivars, the detectable difference is lower than the statistical analysis allows. In contrast, even though some of these cultivars arose as the result of natural mutations and are thus related, cultivars that resulted from these mutations have a different DNA content relative to their precursors (such as the case of ‘Pacamara’ and ‘Maragogype’). Other cultivars have different genetic origin and yet possess similar DNA content, i.e., ‘Pacamara’ and ‘Garena’.

In general, cultivars resulting from spontaneous mutations have greater DNA content than cultivars from which they originated. This is the case of the Typica cultivar that gave rise to ‘Maragogype’, ‘Pacamara’, and ‘Pluma Hidalgo’. A similar situation is found in the case of the Catuaba cultivar with greater DNA content than its parent ‘Bourbon’. The estimation of genome with PI using primary standard shows two groups, which are significant different (\( \Delta S_{0.05} = 0.13356 \)). The first represents Typica-related cultivars, with lower genome content (‘Maragogype’, ‘Pluma Hidalgo’, ‘Villa Sarchi’, ‘Bourbon’, ‘Catuara’, ‘Catuai’, and ‘Mundo Novo’), whereas the second includes the interspecific coffee species from Africa (Baack et al., 2005; Giles et al., 2019). Intraspecific variation, compared with the parental hybrid populations exhibit similar intraspecific DNA content variation, which would have under-
given different genome size changes as compared with F1s or are simply a mixture of hybrid genotypes, not necessarily from late hybrid generations.

In conclusion, FCM was used for the estimation of genome size and GC% content in 16 C. arabica cultivars, which allowed the identification of different groups that agree with the previously grouping described, obtained through phylogenetic and amplified fragment length polymorphism analysis (Anthony et al., 2002; Giles et al., 2019; Machado et al., 2017; Steiger et al., 2002). This information could be helpful for C. arabica breeding programs, given the paucity of genome size studies with FCM in different important C. arabica cultivars (such as the group corresponding to the ‘Bourbon’ and ‘Typica’ genetic family hybrids and artificial crosses). Additionally, these results may be relevant for the wine quality as well as for a better understanding of C. arabica evolutional-ity, relationships, diversification, hybridization, and polyploidy.

Literature Cited

Abreu, I.S., C.R. Carvalho, C.R. Carvalho, G.M.A. Carvalho, and S.Y. Motoike. 2011. First kar- yotype, DNA C-value and AT/GC base com- position of macaw palm (Aeroamia aculeata, Areaceae) a promising plant for biodiesel production. Austral. J. Bot. 59:149–155.

Anthony, F., B. Bertrand, O. Quiros, A. Wilches, J. Bertrand, and A. Charrier. 2001. Genetic di- vertic of wild coffee (Coffea arabica L.) using molecular markers. Euphytica 118:53–65.

Anthony, F., M.C. Combes, C. Astorga, B. Bertrand, G. Graziosi, and P. Lashermes. 2002. The origin of cultivated Coffea Arabica L. cultivars revealed by AFLP and SSR markers. Theor. Appl. Genet. 104:894–900.

Arunugamanathan, K. and E.D. Earle. 1991. Nuclear DNA content of some important plant species. Plant Mol. Biol. Rpt. 9:208–218.

Baack, E.J., K.D. Whitney, and L.H. Rieseberg. 2005. Hybridization and genome size evolu- tion: Timing and magnitude of nuclear DNA content increases in Helianthus homoploid hybrid species. New Phytol. 167:623–630.

Barow, M. and A. Meister. 2002. Lack of correla- tion between AT frequency and genome size in higher plants and the effect of nonrandomness of base sequences on dye binding. Cytometry A 47:1–7.

Barre, P., M. Noiriot, J. Louarn, C. Duperray, and S. Hamon. 1996. Reliable flow cytometric esti- mation of nuclear DNA content in coffee trees. Cytometry 24:32–38.

Bennett, M.D. and J.I. Leitch. 2005. Nuclear DNA amounts in angiosperms: Progress, problems, and prospects. Ann. Bot. 95:45–90.

Bennett, M.D. and I.J. Leitch. 2005. Plant DNA flow cytometry and estimation of nuclear genome size. Annu. Bot. 95:99–110.

Dolezél, J. and J. Rous. 2005. Plant DNA flow cytometry and estimation of nuclear genome size. Annu. Bot. 95:99–110.

Dolezél, J. and J. Greilhuber. 2010. Nuclear ge- nome size: Are we getting closer? Cytometry A 77:635–642.

Dolezél, J., J. Greilhuber, S. Lucrettii, A. Meister, M.A. Lysatk, and L. Lardii. 1998. Plant Genome Size Estimation by Flow Cytometry Inter-laboratory Comparison. Ann. Bot. 82: 17–26.

Dolezél, J., J. Greilhuber, and J. Suda. 2007. Estimation of nuclear DNA content in plants using flow cytometry. Nat. Protoc. 2:2233–2244.

Dolezél, J., S. Sgorbati, and S. Lucrettii. 2007. Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. Physiol. Plant. 85:625–631.

Fontes, B.P.D. 2003. Citogenética, citometria de fluxo e citometria de imagem em Coffea spp. PhD Thesis, Univ. of Víncosa, Brazil.

Giles, J.A.D., A.D. Ferreira, L.F. Partelli, E.M. Aoyama, J.C. Ramalho, A. Ferreira, and A.R. Falqueato. 2019. Divergence and genetic param- eters between Coffea sp. species based in foliar morpho-anatonic traits. Scientia Hortic. 245:231–236.

Giles, J.A.D., L.F. Partelli, A. Ferreira, J.P. Rodrigues, G. Oliosi, and F.H. Silva. 2018. Genetic diversity of promising ‘conilon’ coffee clones based on morpho-agronomic variables. Ann. Bras. Cienc. 90:2437–2446.

Godelle, B., D. Cartier, D. Marie, S.C. Brown, and S. Siljak-Yakovlev. 1993. Heterochromatin
study demonstrating the non-linearity of fluorometry useful for calculating genomic base composition. Cytometry A 14:618–626.

Greilhuber, J. 2005. Intraspecific variation in genome size in angiosperms: Identifying its existence. Bot. Rev. 71:145–198.

Greilhuber, J. and J. Dolezl. 2009. 2C or not 2C: A closer look at cell nuclei and their DNA content. Chromosoma 118:391–400.

Greilhuber, J., J. Dolezl, M.A. Lysak, and M.D. Bennett. 2005. The origin, evolution and proposed stabilization of the terms ‘genome size’ and ‘C-value’ to describe nuclear DNA content. Ann. Bot. 86:859–909.

Grover, C.E. and J.F. Wendel. 2010. Recent insights into mechanisms of genome size change in plants. J. Bot. 2010:1–8.

Grover, C.E., Y. Yu, R.A. Wing, A.H. Paterson, and J.F. Wendel. 2008. A phylogenetic analysis of indel dynamic in the cotton genus. Mol. Biol. Evol. 7:1415–1418.

Herrera, J.C., M.C. Combos, H. Cortina, G. Alvarado, and P. Lashermes. 2002. Gene introgression into Coffea arabica by way of triploid hybrids (C. arabica × C. canephora). Heredity 89:488–494.

Huang, H., Y. Tong, Q.J. Zhang, and L.Z. Gao. 2013. Genome size variation among and within Camellia species by using flow cytometric analysis. PLOS One. 8:1–5.

International Coffee Organization. 3 Jan. 2018. <http://www.icco.org/trade_statistics.asp>.

Lashermes, P., M.C. Combos, C. Ansaldi, E. Gichuru, and S. Noir. 2011. Analysis of alien introgression in coffee tree (Coffea arabica L.). Mol. Bot. 27:27–232.

Leong-Skornickova, J., O. Sida, V. Jarolimová, M. Sabu, T. Fér, P. Trávníček, and J. Suda. 2007. Chromosome numbers and genome size variation in Indian species of Carcina (Zingiberaceae). Ann. Bot. 100:505–526.

Machado, C.M.S., N.S. Pimentel, A. Golynsk, A. Ferreira, H.D. Vieira, and F.L. Partelli. 2017. Genetic diversity among 16 genotypes of Coffea Arabica in the Brazilian cerrado. Genet. Mol. Res. 16:1–13.

Marques, I., G. Nieto-Feliner, M. A. Martins-Loucao, and J. Fuertes-Aguilar. 2011. Genome size and base composition variation in natural and experimental Narcissus (Amaryllidaceae) hybrids. Ann. Bot. 190:257–264.

Martins, M.Q., A.S. Fortunato, W.P. Rodrigues, F.L. Partelli, E. Campanostrini, F.J.C. Lidon, F.M. DaMattá, J.C. Ramalho, and A.I. Ribeiro-Barros. 2017. Selection and validation of reference genes for accurate RT-qPCR data normalization in Coffea spp. under a climate change context of interacting elevated [CO2] and temperature. Front. Plant Sci. 8:307.

Martins, M.Q., W.P. Rodrigues, A.S. Fortunato, A.E. Leitão, A.P. Rodrigues, I.P. Pais, L.D. Martins, M.J. Silva, F.H. Reboredo, F.L. Partelli, E. Campanostrini, M.A. Tomaz, P. Scotti-Campos, A.I. Ribeiro-Barros, F.J.C. Lidon, F.M. DaMattá, and J.C. Ramalho. 2016. Protective response mechanisms to heat stress in interaction with high [CO2] conditions in Coffea spp. Front. Plant Sci. 7:947.

Meister, A. and M. Barow. 2007. Nuclear DNA content measurement. p. 67–101. In: J. Dolezl, J. Greilhuber, and J. Suda (eds). Flow cytometry with plant cells: Analysis of genes, chromosomes and genomes. Wiley-VCH, Weinheim, Germany.

Mishra, M.K. and A. Slater. 2012. Recent advances in the genetic transformation of coffee. Biotechnol. Res. Intl. 2012:1–17.

Noirot, M., P. Barre, D. Christin, S. Hamon, and A. De Kochko. 2005. Investigation on the causes of stoichiometric error in genome size estimation using heat experiments: Consequences on data interpretation. Ann. Bot. 95:111–118.

Noirot, M., P. Barre, C. Duperray, J. Louarn, and S. Hamon. 2003b. Effects of caffeine and chlorogenic acid on propidium iodide accessibility to DNA: Consequences on genome size evaluation in coffee tree. Ann. Bot. 92:259–264.

Noirot, M., P. Barre, J. Louarn, C. Duperray, and S. Hamon. 2000. Nucleus–cytosol interactions—a source of stoichiometric error in flow cytometric estimation of nuclear DNA content in plants. Ann. Bot. 86:309–316.

Noirot, M., P. Barre, J. Louarn, C. Duperray, and S. Hamon. 2002. Consequences of stoichiometric error on nuclear DNA content evaluation in *Coffea liberica* var. dewevrei using DAPI and propidium iodide. Ann. Bot. 89:385–389.

Noirot, M., V. Poncelet, P. Barre, P. Hamon, S. Hamon, and A. De Kochko. 2003a. Genome size variations in diploid African *Coffea* species. Ann. Bot. 92:709–714.

The Observatory of Economic Complexity. 2018. 12 Oct. 2018. <https://atlas.media.mit.edu/en/>.

Parris, J.K., T.G. Ranney, H.T. Knap, and W.V. Baird. 2010. Ploidy levels, relative genome sizes, and base pair composition in Magnolia. J. Amer. Soc. Hort. Sci. 135:533–547.

Praga-Fonts, M.M., C.R. Carvalho, W.R. Clarindo, and C.D. Cruz. 2011. Revisiting the DNA C-values of the genome size-standards used in plant flow cytometry to choose the “best primary standards.” Plant Cell Rep. 30:1183–1191.

Price, H.J. 1988. DNA content variation among species of *Erianthus arundinaceus* (Poaceae). Cell Res. 27:209–215.

Rodrigues, W.P., M.Q. Martins, A.S. Fortunato, A.P. Rodrigues, N.J. Semedo, M.C. Simões-Costa, I.P. Pais, A.E. Leitão, F. Colwell, L. Goulao, C. Mágua, R. Maia, F.L. Partelli, E. Campanostrini, P. Scotti-Campos, A.I. Ribeiro-Barros, F.C. Lidon, F.M. DaMattá, and J. Ramalho. 2016. Long-term elevated air [CO2] strengthens photosynthetic functioning and mitigates the impact of supra-optimal temperatures in tropical *Coffea arabica* and *C. canephora* species. Glob. Change Biol. 22:415–431.

Rothleutner, J.J., M.W. Friddle, and R.N. Contreras. 2016. Ploidy levels, relative genome sizes, and base pair composition in *Cotonuaster* el. J. Amer. Soc. Hort. Sci. 141:457–466.

Sanglard, N.A., P.M. Amaral-Silva, M.C. Sattler, S.C. Cristina de Oliveira, A.C.P. Nunes, T.C.B. Soares, C.R. Carvalho, and W.R. Clarindo. 2017. From chromosome doubling to DNA sequence changes: Outcomes of an improved in vitro procedure developed for allotriploid "Hibrido de Timor" (Coffea arabica L. × *Coffea canephora* Pierre ex A. Froehner). Plant Cell Tissue Organ Cult. 131:223–231.

Sattler, M.C., C.R. Carvalho, and W.R. Clarindo. 2016. Regeneration of allotriploid *Coffea* plants from tissue culture: Resolving the propagation problems promoted by irregular meiosis. Cytologia 81:125–132.

Shapiro, H.M. 2005. Practical flow cytometry. 4th ed. Wiley, Hoboken, NJ.

Sisko, M., A. Ivanic, and B. Bohaneck. 2003. Genome size analysis in the genus *Cucurbita* and its use for determination of interspecific hybrids obtained using embryo-rescue technique. Plant Sci. 165:663–669.

Steiger, D.L., C. Nagai, P.H. Moore, C.W. Morden, R.V. Osgood, and R. Ming. 2002. AFLP analysis of genetic diversity within and among *Coffea Arabica* cultivars. Theor. Appl. Genet. 105:209–215.

Tatum, T.C., L. Nunez, M.M. Kushad, and A.L. Rayburn. 2006. Genome size variation in pumpkin (*Cucurbita sp*). Ann. Appl. Biol. 149:145–151.

Vilhar, B., J. Greilhuber, J.D. Koce, E.M. Temsch, and M. Dermastia. 2001. Plant genome size measurement with DNA image cytometry. Ann. Bot. 87:719–728.

Williams, R.R., S. Broad, D. Sheer, and J. Ragoussis. 2002. Subchromosomal positioning of the epidermal differentiation complex (EDC) in keratinocyte and lymphoblast interfase nuclei. Exp. Cell Res. 272:163–175.

World Coffee Research. 2018. 30 Sept. 2018. <https://worldcofferesearch.org/>.

Yan, J., J. Zhang, K. Sun, D. Chang, S. Bai, Y. Shen, L. Huang, J. Zhang, Y. Zhang, and Y. Dong. 2016. Ploidy level and DNA content of *Eriathus arundinaceus* as determined by flow cytometry and the association with biological characteristics. PLOS One 11:1–14.
Supplemental Table 1. Three independent repetitions of nuclear DNA amount of *C. arabica* cultivars using soybean (2C = 2.30 pg) as primary standard and *C. canephora* (2C = 1.29 pg) as secondary standard.

| Cultivars Replication | Mean of fluorescence |
|-----------------------|----------------------|
| Soybean 163.39 163.36 163.35 163.36 ± 0.03 |
| Coffea canephora var. Robusta 91.31 91.36 91.33 91.33 ± 0.03 |
| Coffea arabica var. Typica 181.54 181.86 181.70 181.70 ± 0.16 |
| Coffea arabica var. Maragogype 185.42 185.49 185.56 185.49 ± 0.05 |
| Coffea arabica var. Pluma Hidalgo 188.91 188.90 188.92 188.91 ± 0.01 |
| Coffea arabica var. Villa Sarchi 196.01 196.07 196.14 196.07 ± 0.07 |
| Coffea arabica var. Bourbon 202.97 202.96 202.98 202.97 ± 0.01 |
| Coffea arabica var. Caturra 204.41 204.37 204.33 204.37 ± 0.04 |
| Coffea arabica var. Catuaí 204.79 204.83 204.87 204.83 ± 0.04 |
| Coffea arabica var. Mundo Novo 205.66 205.70 205.64 205.67 ± 0.03 |
| Coffea arabica var. Gourmetica 208.09 208.10 208.13 208.11 ± 0.02 |
| Coffea arabica var. Pacamara 212.65 212.63 212.60 212.63 ± 0.03 |
| Coffea arabica var. Pluma H 214.91 214.95 214.98 214.95 ± 0.04 |
| Coffea arabica var. Villa Sarchi 216.60 216.65 216.62 216.62 ± 0.03 |
| Coffea arabica var. IAPAR 59 216.83 216.99 216.91 216.91 ± 0.08 |
| Coffea arabica var. Costa Rica 221.76 221.99 221.96 221.90 ± 0.13 |
| Coffea arabica var. Oro Azteca 224.03 224.02 224.06 224.04 ± 0.02 |
| Coffea arabica var. ICATU 224.37 224.36 224.33 224.35 ± 0.02 |

Supplemental Table 2. Nuclear DNA amount of *C. arabica* varieties stained with PI using *C. canephora* as standard.

| Cellular cycle (G0/G1) | IP AU12 |
|-----------------------|--------|
| Typica 181.7 3549 5.23 2.52 ± 0.026 | 2.52 ± 0.026 |
| Marago 185.49 5580 5.65 2.62 ± 0.019 | 2.58 ± 0.019 |
| Pluma H 188.91 2924 5.43 2.58 ± 0.034 | 2.62 ± 0.034 |
| Villa sarchi 196.07 3072 4.21 2.62 ± 0.004 | 2.72 ± 0.004 |
| Bourbon 202.97 4504 5.66 2.82 ± 0.029 | 2.84 ± 0.016 |
| Caturra 204.37 4738 6.13 2.84 ± 0.016 | 2.85 ± 0.055 |
| Catuaí 204.83 5505 4.44 2.85 ± 0.055 | 2.86 ± 0.064 |
| Mundo N 205.67 3213 4.29 2.86 ± 0.042 | 2.89 ± 0.042 |
| Gourmetica 208.11 4474 5.65 3.2 ± 0.028 | 2.95 ± 0.028 |
| Pacamara 212.63 3246 6.32 3.2 ± 0.010 | 2.99 ± 0.010 |
| Garena 214.95 3606 4.97 3.01 ± 0.031 | 3.01 ± 0.086 |
| Colombia 216.6 3354 5.15 3.01 ± 0.031 | 3.01 ± 0.031 |
| IAPAR 59 216.91 2716 4.96 3.00 ± 0.086 | 3.00 ± 0.086 |
| Costa rica 221.9 3700 5.73 3.08 ± 0.050 | 3.08 ± 0.050 |
| Oro azteca 224.04 2833 5.22 3.11 ± 0.062 | 3.11 ± 0.062 |
| ICATU 224.35 4896 6.02 3.12 ± 0.037 | 3.12 ± 0.037 |

PI = propidium iodide.

Supplemental Table 3. Nuclear DNA amount of *C. arabica* cultivars stained with DAPI using soybean and *C. canephora* as standards.

| Cellular cycle (G0/G1) | DAPI 4',6'-diamino-2-phenylindole dihydrochloride hydrate. |
|-----------------------|--------------------------------------------------|
| Typica 90.83 2173 3.85 3.03 ± 0.042 | 2.85 ± 0.031 |
| Marago 93.79 6118 5.85 3.13 ± 0.049 | 2.94 ± 0.012 |
| Pluma H 93.86 9193 5.55 3.13 ± 0.050 | 2.94 ± 0.056 |
| Villa sarchi 94.19 7570 6.22 3.14 ± 0.033 | 2.95 ± 0.073 |
| Bourbon 94.35 5910 5.88 3.15 ± 0.027 | 2.96 ± 0.040 |
| Caturra 95.49 7357 5.44 3.18 ± 0.031 | 2.99 ± 0.066 |
| Catuaí 95.61 9121 5.44 3.19 ± 0.029 | 3.00 ± 0.025 |
| Mundo N 95.72 7229 5.48 3.19 ± 0.032 | 3.00 ± 0.015 |
| Garena 95.95 4834 5.65 3.20 ± 0.027 | 3.01 ± 0.084 |
| Pacamara 97.50 6267 5.64 3.25 ± 0.025 | 3.06 ± 0.035 |
| Garena 97.77 5950 5.55 3.26 ± 0.026 | 3.06 ± 0.041 |
| Colombia 98.15 6908 4.87 3.27 ± 0.019 | 3.08 ± 0.097 |
| IAPAR 59 98.42 3628 5.09 3.28 ± 0.022 | 3.08 ± 0.026 |
| Costa rica 99.12 8102 5.42 3.30 ± 0.015 | 3.11 ± 0.038 |
| Oro azteca 104.26 5237 5.4 3.48 ± 0.033 | 3.27 ± 0.075 |
| ICATU 104.96 6404 5.09 3.50 ± 0.030 | 3.45 ± 0.043 |