

## Karyotype and genome size determination of *Jarilla chocola*, an additional sister clade of *Carica papaya*

Dessiré Patricia Zerpa-Catanho<sup>1</sup>, Tahira Jatt<sup>2</sup>, Ray Ming<sup>1\*</sup>

<sup>1</sup>Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

<sup>2</sup>Department of Botany, Shah Abdul Latif University, Khaipur, Sindh, 66020, Pakistan

\*Corresponding author: rayming@illinois.edu

### Abstract

*Jarilla chocola* is an herbaceous plant species that belongs to the *Jarilla* genus and the Caricaceae family. No information on chromosome number or genome size has been reported for *J. chocola* that confirms the occurrence of dysploidy events and explore the existence of heteromorphic sex chromosomes. Therefore, the total number of chromosomes of this species was determined by karyotyping and counting the number of chromosomes observed, and the genome size of female and male plants was estimated separately by flow cytometry. Results showed that *J. chocola* has eight pairs of chromosomes ( $2n = 2x = 16$ ), and its chromosomes are classified as metacentric for five pairs, submetacentric for two pairs and telocentric for one pair. The nuclear DNA content (1C-value) in picograms and diploid genome size was estimated separately from female and male plants using two species as the standards, *Phaseolus vulgaris* (1C = 0.60 pg) and *Carica papaya* (1C = 0.325 pg), to look for the possible existence of heteromorphic sex chromosomes. *C. papaya* proved to be a better standard for the determination of *J. chocola* DNA content and diploid genome size. No significant difference on the DNA content was observed between female (1C =  $1.02 \pm 0.003$  pg) and male (1C =  $1.02 \pm 0.008$  pg) plants. The estimated genome size of *J. chocola* per haploid genome in base pairs was calculated from the obtained C-values. Results showed an estimated genome size per haploid genome of  $1018.44 \pm 3.07$  Mb and  $1022.08 \pm 7.76$  Mb for female and male plants, respectively. Due to the observed chromosome number and genome size, only the occurrence of one of two previously reported dysploidy events in *Jarilla* could be confirmed for *J. chocola* and no evidence of heteromorphic sex chromosomes was found. These results provide fundamental information of the *J. chocola* genome and will expedite investigation of sex chromosomes and genome evolution in this species, the *Jarilla* genus and the Caricaceae family.

**Keywords:** Caricaceae, chromosome number, DNA content, flow cytometry, *Jarilla chocola*.

**Abbreviations:** 1C- 1C-value; 2C- 2C-value; AR-Arms ratio; RL-Relative length coefficient; TCL-Total chromosome length.

### Introduction

*Jarilla chocola* is an herbaceous plant species that belongs to the *Jarilla* genus of the Caricaceae family and it is classified as one of the closest relatives of *Carica papaya* (Carvalho and Renner, 2012). This genus was described for the first time in 1832 by Pablo de La Llave, as *Mocinna*, but later the name was corrected to *Jarilla* (Carvalho and Renner, 2013). This herbaceous species is distributed at low altitudes across the Pacific coast (from 100 to 1300 m) from the northern part of Mexico to Guatemala and El Salvador (Cardoza Ruiz et al., 2014). Besides *J. chocola*, the genus *Jarilla* includes two other herbaceous plant species, *J. caudata* and *J. heterophylla*, and a natural hybrid (*J. caudata*  $\times$  *J. heterophylla*), that are distributed in higher altitudes of Mexico (Carvalho and Renner, 2013; Cardoza Ruiz et al., 2014).

*J. chocola* plants are dioecious (male and female flowers on separate individuals), and, unlike papaya, their stem is green and very small (20-50 cm long, 1-8 mm diameter), they can be propagated asexually by tubers, they can be kept in the greenhouse and they have a relative short life cycle (Willingham and White, 1976; Cardoza Ruiz et al., 2014). Like papaya plants, these plants are not branched; their leaves are lobed (3 to 5) and produce a milky latex with a papain-like

enzyme that has proteinase activity (Tookey and Gentry, 1969; Willingham and White, 1976; Cardoza Ruiz et al., 2014). Male *J. chocola* plants produce staminate flowers arranged in an inflorescence (with 8-40 flowers each) with a long peduncle (7-21 cm long). Like papaya male flowers, they are white and present two groups of stamens (one superior and one inferior) and a small rudimentary ovary (3 mm) (Cardoza Ruiz et al., 2014). Female *J. chocola* plants produce solitary pistillate flowers and, like papaya female flowers, they are white in color with a short peduncle (Cardoza Ruiz et al., 2014). All these similarities make *J. chocola* an interesting resource for comparative research in papaya, as a dioecious herbaceous model for the study of sex chromosomes, whose existence has been reported in other dioecious Caricaceae species (Wu et al., 2010; Gschwend et al., 2013; Iovene et al., 2015).

Previous studies with *J. chocola* plants have been focused only on its potential as a crop for proteinase production and the characterization of a papain-like proteinase present in its latex (Tookey and Gentry, 1969; Willingham and White, 1976). A previous examination of the karyotype and genome size of different Caricaceae species has been done, including

*J. caudata* and *J. heterophylla* (Rockinger et al., 2016), but not *J. chocola*. According to this study, *Jarilla* species have doubled their genome size, but reduced their chromosome number ( $2n = 14$ ) compared to other Caricaceae species ( $2n = 18$ ), through two dysploidy events that occurred approximately 16.6 and 5.5 million years ago (Rockinger et al., 2016). There is no evidence of the occurrence of these events in *J. chocola* due to the lack of information on the total number of chromosomes for this species. Furthermore, there is no information regarding its genome size, nor any evidence supporting whether this dioecious species has sex chromosomes, creating a gap in genome evolution and sex chromosome evolution studies on Caricaceae.

To investigate the occurrence of the described dysploidy events in the *Jarilla* genus, to report *J. chocola* genome size and to examine the possible existence of heteromorphic sex chromosomes in *J. chocola*, a cytological study and a flow cytometry analysis were performed.

## Results

### **Chromosome number, length and karyotype analysis**

According to the observation of chromosome slide preparations of plant root tips, *J. chocola* is a diploid species with 8 pairs of chromosomes, for a total of 16 chromosomes (Fig. 1). The 16 chromosomes could be paired according to their morphology in eight pairs (Fig. 2). Six pairs of chromosomes were classified as metacentric chromosomes, one pair was classified as submetacentric and one pair was classified as telocentric chromosomes (Table 1, Fig. 3).

According to their relative length, only five metacentric chromosomes were classified as long (chromosome pair 1 to 5, Table 1), one metacentric and one submetacentric pair were classified as slightly short (chromosome pair 6 and 7, Table 1), while the telocentric chromosome was classified as short (chromosome 8, Table 1). It is important to mention that the chromosome pair number 8 (telocentric chromosome pair) showed a slight variation in size among plants, but this difference could correspond to slide preparation or chromosome condensation changes, and it is not enough proof of the existence of heteromorphic sex chromosomes in *J. chocola* (Fig. 1).

### **Nuclear DNA content and genome size**

Once the plants flowered, the nuclear DNA content (1C-value) of *J. chocola* female and male plants could be calculated using flow cytometry (Table 2). Compared to *P. vulgaris*, *J. chocola* had approximately double the content of nuclear DNA (Fig. 4). Compared to *C. papaya*, *J. chocola* had approximately three times the content of nuclear DNA (Fig. 4). An example of Gate selection for building flow cytometry histograms are shown in Supplementary Figures 1 and 2.

Using the mean position of G0/G1 detected peaks and common bean (*P. vulgaris*) as the standard, a total of  $1.13 \pm 0.01$  pg and  $1.12 \pm 0.01$  pg of DNA was calculated for female and male plants, respectively (Table 2). Therefore, the estimated genome size per haploid genome was  $1130.78 \pm 5.29$  Mb and  $1116.27 \pm 9.10$  Mb for female and male plants, using common bean (*P. vulgaris*) as the standard (Table 2). Using the mean position of G0/G1 detected peaks and papaya (*C. papaya*) as the standard, a total of  $1.02 \pm 0.003$  pg and  $1.02 \pm 0.01$  pg of DNA was calculated for female and male plants, respectively (Table 2). Therefore, the estimated genome size per haploid genome was  $1018.44 \pm 3.07$  Mb and  $1022.08 \pm 7.76$  Mb for female and male plants, using papaya

(*C. papaya*) as the standard (Table 2). No significant differences between female and male plants were detected regarding their DNA content or their genome size.

Variation coefficients were higher when using common bean as the standard (Fig. 4), probably due to the overlapping between G2/M nuclei peaks from *P. vulgaris* and G0/G1 *J. chocola*, which was evident when selecting the Gates to build the flow cytometry histograms (Supplementary Figure 1). *C. papaya* proved to be a better internal standard for the determination of *J. chocola* genome size than *P. vulgaris*, due to the lack of overlap among nuclei peaks from both species as shown by the selected Gates (Supplementary Figure 2).

## Discussion

### **Chromosome number, length and karyotype analysis**

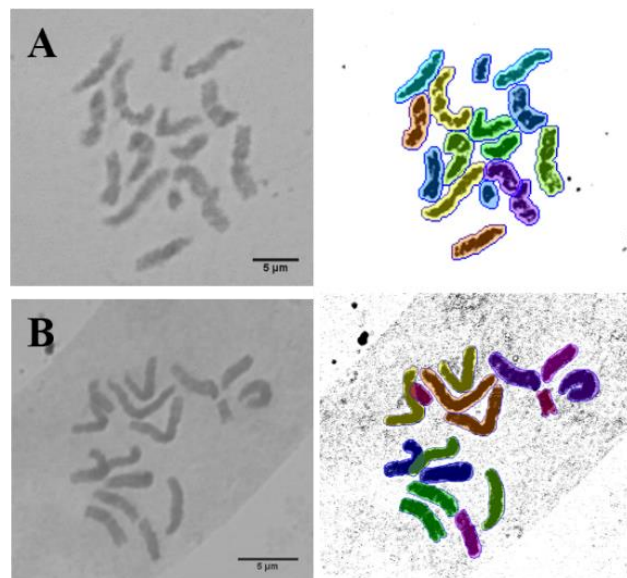
This study reports *J. chocola* karyotype for the first time, with a total of 8 pairs of chromosomes ( $2n = 2x = 16$ ) and their respective morphologies. According to a recent study, the ancestral chromosomal number of the Caricaceae plant family is  $2n = 18$ , but a reduction in the chromosome number and a doubling of the genome size occurred in the common ancestor of *Horowitzia* and *Jarilla* (closest papaya relatives), explaining the bigger genome size of the species in these two genera in comparison with the papaya genome size (Rockinger et al., 2016). The only *Horowitzia* species (*H. cnidoscolioides*) has a total of 8 pairs of chromosomes ( $2n = 2x = 16$ ) and since other species of *Jarilla* (*J. caudata* and *J. heterophylla*) have a total of 7 pairs of chromosomes ( $2n = 2x = 14$ ), *J. chocola* was expected to have 7 pairs of chromosomes (Rockinger et al., 2016). Our observations on the number of chromosomes of *J. chocola* ( $2n = 2x = 16$ ) differ from the expectation and they suggest that only one event of dysploidy or chromosome number reduction occurred in the common ancestor between *Horowitzia* and *Jarilla* approximately 16.6 million years ago, as reported previously (Rockinger et al., 2016). Our observations on the number of chromosomes of *J. chocola* ( $2n = 16$ ) also suggest that approximately 1.7 million years ago (instead of the proposed 5.5 million years ago), a second dysploidy event must have occurred in the common ancestor of *J. caudata* and *J. heterophylla* because the chromosome number is not conserved among the three *Jarilla* species (*J. caudata*, *J. heterophylla* and *J. chocola*).

According to our chromosome morphology analysis, *J. chocola* has five pairs of chromosomes that were classified as long metacentric chromosomes; two pairs that were classified as slightly short metacentric and submetacentric chromosomes; and one pair that was classified as short telocentric chromosomes. While the karyotype and size of each pair of chromosomes in other *Jarilla* species has not been reported, the karyotype of other three cultivated Caricaceae species (*C. papaya*, *V. monoica* and *V. cundinarmacensis*) has been previously reported (Corrêa Damasceno Junior et al., 2009; Rockinger et al., 2016). *Carica papaya*, *V. monoica* and *V. cundinarmacensis* all have nine pairs of chromosomes ( $2n = 2x = 18$ ) and the chromosomes in these three species have been described as symmetric (homomorphic chromosomes), metacentric and small (1.52 to 2.29  $\mu\text{m}$  in *C. papaya*, 1.35 to 2.49  $\mu\text{m}$  in *V. monoica* and 1.66 to 2.45  $\mu\text{m}$  in *V. cundinarmacensis*) (Corrêa Damasceno Junior et al., 2009). A recent study of the karyotype and genome size of *C. papaya* reported the same number of chromosomes in papaya ( $2n = 2x = 18$ ), with seven metacentric (1, 2, 3, 4, 5, 7 and 8) and two submetacentric

**Table 1.** Morphology of the chromosomes of *Jarilla chocola*.

No.	Length ( $\mu\text{m}$ )			AR <sup>a</sup>	RL <sup>b</sup>	Centromere position Type <sup>c</sup>	Length Type <sup>d</sup>
	Long arm $\pm$ S.E.	Short arm $\pm$ S.E.	Total Chromosome Length (TCL) $\pm$ S.E.				
1	5.07 $\pm$ 0.12	4.82 $\pm$ 0.08	9.89 $\pm$ 0.15	1.05	1.43	m	L
2	4.99 $\pm$ 0.08	4.57 $\pm$ 0.19	9.56 $\pm$ 0.25	1.09	1.39	m	L
3	4.46 $\pm$ 0.13	3.79 $\pm$ 0.06	8.25 $\pm$ 0.14	1.18	1.20	m	L
4	3.82 $\pm$ 0.09	3.32 $\pm$ 0.05	7.14 $\pm$ 0.09	1.15	1.03	m	L
5	3.69 $\pm$ 0.04	3.14 $\pm$ 0.09	6.83 $\pm$ 0.08	1.18	0.99	m	L
6	3.38 $\pm$ 0.14	2.39 $\pm$ 0.04	5.77 $\pm$ 0.14	1.41	0.84	m	M1
7	4.06 $\pm$ 0.15	1.36 $\pm$ 0.02	5.42 $\pm$ 0.15	2.99	0.79	sm	M1
8	2.21 $\pm$ 0.08	0.15 $\pm$ 0.01	2.36 $\pm$ 0.08	14.73	0.34	t	S
		$\Sigma$ TCL	55.22				
		ACL <sup>e</sup>	6.90				

<sup>a</sup>Arms ratio (AR) = long arm length/short arm length; <sup>b</sup>relative length coefficient (RL) = TCL/ACL; <sup>c</sup>m = metacentric, sm = submetacentric, t = telocentric; <sup>d</sup>L = long, M1 = slightly short, S = short; <sup>e</sup>Average chromosome length =  $\Sigma$  TCL/n, with n = 8.

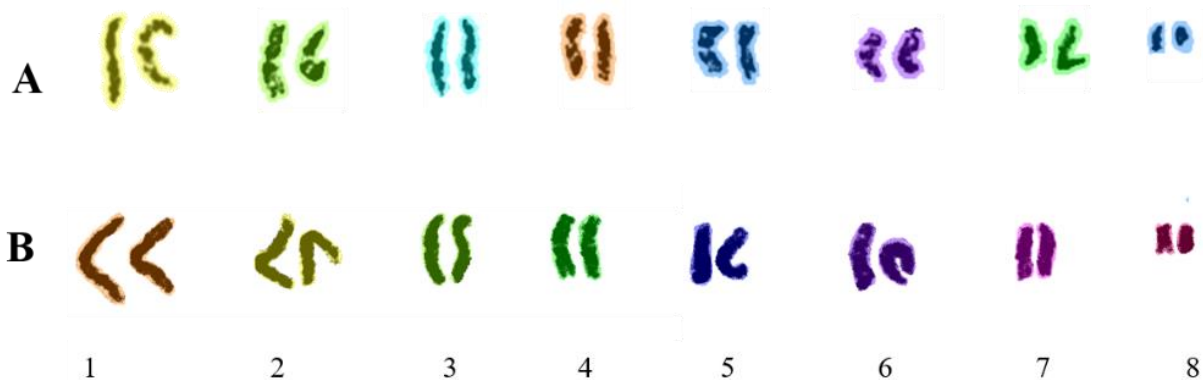


**Fig 1.** Chromosomes of two different *Jarilla chocola* plants during metaphase (X = 100). a. scale bar = 5  $\mu\text{m}$ . b. scale bar = 5  $\mu\text{m}$ . Colors represent chromosome pairs identified for two of five analyzed plants.

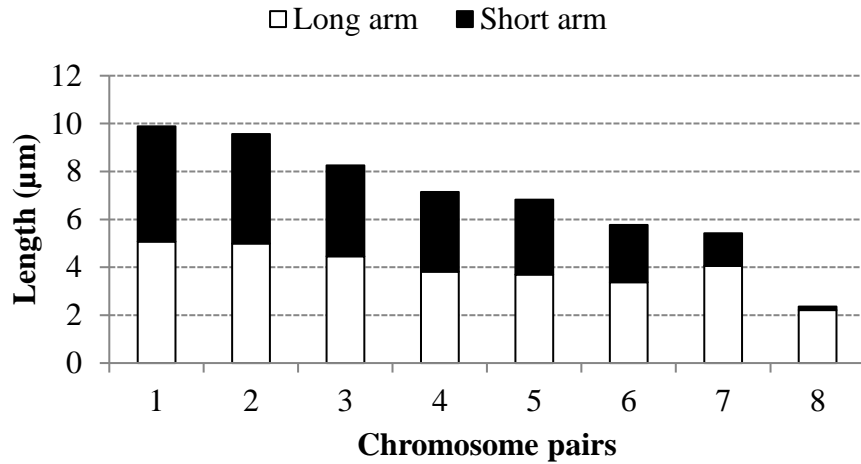
**Table 2.** DNA content and genome size of *Jarilla chocola* plants measured by flow cytometry.

Species used as the standard	<i>J. chocola</i> sample	n	2C-value* (pg)	1C-value* (pg)	Genome size (Mbp) per haploid genome*
<i>Phaseolus vulgaris</i>	Female	2	2.31 $\pm$ 0.01	1.13 $\pm$ 0.01	1130.78 $\pm$ 5.29
	Male	2	2.28 $\pm$ 0.02	1.12 $\pm$ 0.01	1116.27 $\pm$ 9.10
<i>Carica papaya</i>	Female	7	2.08 $\pm$ 0.01	1.02 $\pm$ 0.003	1018.44 $\pm$ 3.07
	Male	7	2.09 $\pm$ 0.02	1.02 $\pm$ 0.01	1022.08 $\pm$ 7.76

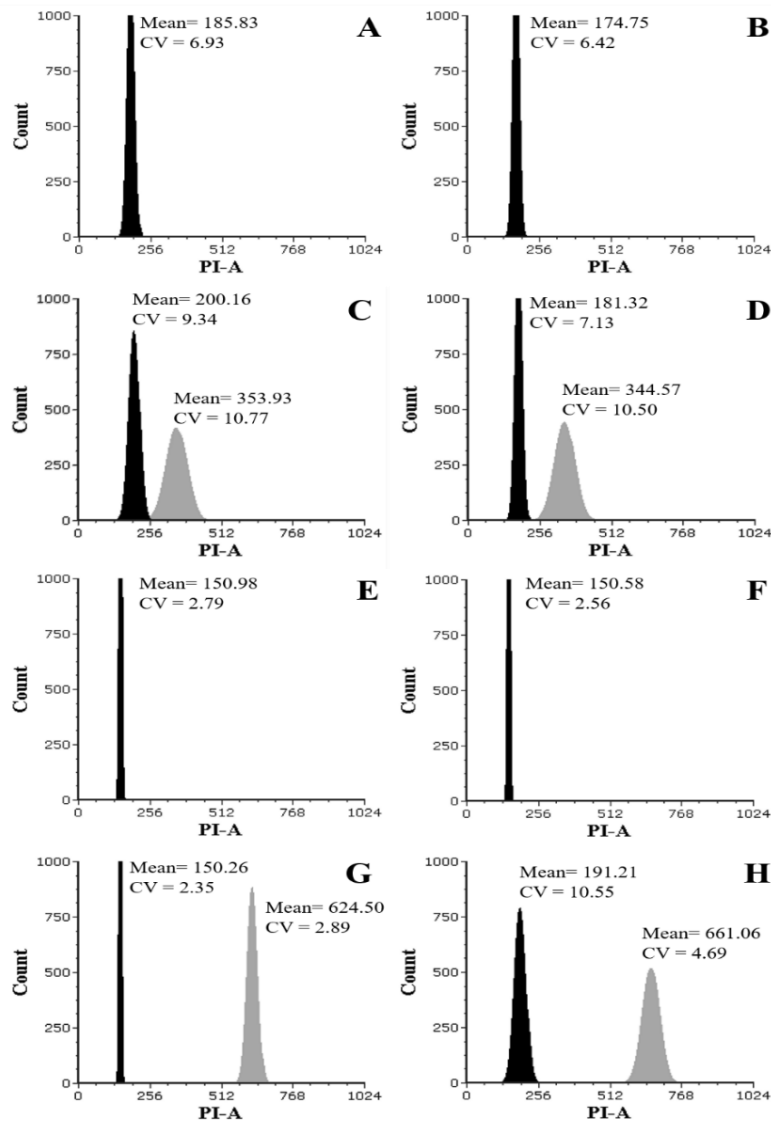
\*Mean  $\pm$  std. error. No significant difference was detected between male and female, p-value > 0.05 (Two tailed T-test).



**Fig. 2** Karyotype analysis of previous *Jarilla chocola* plants (a and b, Fig 1). Homoeologous chromosomes were paired based on their morphology (Table 1). Colors are the same as Fig. 1 and represent the chromosome pairs identified in each plant.



**Fig 3.** Ideogram of *Jarilla chocola* chromosomes. The figure shows the length of the long arm and the length of the short arm for each chromosome pair. Chromosome pairs one to six are metacentric, while pair seven is submetacentric and pair eight is telocentric.



**Fig 4.** Flow cytometry histogram. (A) *P. vulgaris* as the standard. (B) replicate of *P. vulgaris* as the standard. (C) *J. chocola* (female plant with *P. vulgaris* as the internal standard). (D) *J. chocola* (male plant with *P. vulgaris* as the internal standard). (E) *C. papaya* as the standard. (F) replicate of *C. papaya* as the standard. (G) *J. chocola* (female plant with *C. papaya* as the internal standard). (H) *J. chocola* (male plant with *C. papaya* as the internal standard).

(6 and 9) pairs of chromosomes (Rockinger et al., 2016). It is important to notice that no telocentric chromosomes and no heteromorphic chromosomes have been reported for any of the analyzed Caricaceae species. Therefore, *J. chocola* chromosomes are longer and are morphologically different than the chromosomes from three reported Caricaceae species (*C. papaya*, *V. monoica* and *V. cundinarmacensis*).

#### **Nuclear DNA content and genome size**

This study reports *J. chocola* DNA content in female ( $1C = 1.02 \pm 0.003$  pg =  $1018.44 \pm 3.07$  Mbp) and male ( $1C = 1.02 \pm 0.01$  pg =  $1022.08 \pm 7.76$  Mbp) plants. The duplication of the genome size in *J. chocola* appears to be supported by our nuclear DNA content analysis results, which is also consistent with a big genome size reported for other two *Jarilla* species (*J. caudata* = 924 Mbp, *J. heterophylla* = 836 Mbp) by Rockinger et al. (2016). The estimated genome size per haploid genome of *J. chocola* is even bigger than that of the other two species, making *J. chocola* the species with the biggest genome size in the *Jarilla* genus. A different number of duplication events, or a differential accumulation of repetitive elements in *J. chocola* after the first dysploidy event (Rockinger et al., 2016), could explain its bigger genome size compared to *J. caudata* and *J. heterophylla*. However, more studies are required to understand how these events affected the genome of *Jarilla chocola* compared to the other two *Jarilla* species.

According to our results, no evidence of heteromorphic sex chromosomes was found in *J. chocola* and the difference in genome size between female and male plants was not considered statistically significant. Therefore, we can conclude that no heteromorphic sex chromosomes exist in *J. chocola*. Still, these results do not imply that sex chromosomes do not exist in this species. Differences in genome size between sexes in other Caricaceae species with known sex chromosomes (like *C. papaya*, *Vasconcellea quercifolia* and *V. parviflora*) are considered to be too small to be detectable by flow cytometry (Wu et al., 2010; Gschwend et al., 2013; Iovene et al., 2015), which is consistent with our results. Furthermore, no heteromorphic sex chromosomes have been described for any of the Caricaceae species with sex chromosomes (Wu et al., 2010; Gschwend et al., 2013; Iovene et al., 2015). Thus, the existence of homomorphic sex chromosomes in *J. chocola* is still possible and it should be explored by other techniques.

In conclusion, our results provide basic information about the genome of *Jarilla chocola* and contribute to the future research on sex chromosomes and genome evolution of this species, the *Jarilla* genus and the Caricaceae family.

#### **Materials and methods**

##### **Plant material**

*Jarilla chocola* seeds were bought online from a commercial company: Trade Wind Fruits (Santa Rosa, CA, USA). Thirty-two seeds of *J. chocola* were incubated in an aqueous solution of  $KNO_3$  (1mol/L) at 28°C and 150 rpm for 48 hours. After the incubation,  $KNO_3$  solution was removed and seeds were washed with tap water three times to remove possible salt residues. Once washed, seeds were sown in a mixture of Sungro professional growing mix (SS#1-F1P RSi) and medium size Vermiculite (in a ratio 3:1) and watered at convenience (every two days). Two weeks later, the seeds started to germinate. Nevertheless, the germination percentage was low since they require a long period after ripening to

germinate. Seedlings were grown under greenhouse conditions, with a photoperiod of 12 hours of light/12 hours of darkness and a constant temperature of 27 °C. Two months after germination, once plants were considered robust, lateral root tip samples (0.6 cm) were collected between 9:00 and 11:00 am from five different plants (chosen randomly among 16 plants) to perform a cytological and a karyotype analysis. Approximately, four to six months later (once plants flowered), young leaves were collected from female and male plants, separately, to perform a nuclear DNA content analysis by flow cytometry.

##### **Root tip processing**

Excised roots tips were first washed with tap water to remove soil particles and later rinsed three times with distilled, deionized water. Washed root tips were incubated in an 8-hydroxyquinoline (2 mM) solution for two hours. After incubation, root tips were rinsed again with distilled deionized water and fixed in a solution composed of ethanol and glacial acetic acid (ratio 3:1) at 4°C for 24 hours. After fixation, root tips were incubated in deionized water for five minutes, followed by an incubation in HCl (5 N) for nine minutes at room temperature. Then, the hydrolyzed root tips were rinsed with chilled water and transferred into Feulgen stain reagent for 1 hour 30 minutes. Later, root tips were bleached with a solution of k-metabisulfite (0.05% v/v of 10% k-metabisulfite and 0.05% v/v of 1N HCl in water) and stored in a solution of acetocarmine (1% acetocarmine in 45% v/v acetic acid in water) for not more than 48 hours. Finally, root tips were put over a drop of 45% v/v acetic acid in a microscope slide, covered with a cover slip and carefully crushed against the slide. The slides were examined under a light microscope (Olympus BX/61) equipped with a Magnifire digital camera (ChipCoolers, Warwick, RI, USA) at 100X magnification and oil immersion. To record the results, photographs were taken from well-spread chromosome preparations.

##### **Karyotype analysis**

At least five well-spread microscope slides with a similar degree of chromatin condensation (metaphase) were used to count the total number of chromosomes per plant, determine the position of the centromere on the chromosomes, and measure their length. Measurements were done from the photographs using ImageJ (Schindelin et al., 2015), and SmartType software (Digital Scientific, Cambridge, UK) was used to karyotype. No less than 200 cells per plant were examined and different stages of cells divisions were observed (early metaphase and metaphase). Total chromosome length (TCL), average chromosome length (ACL), arms ratio (AR) and relative length coefficient (RL) were calculated from the data. Chromosomes were arranged in pairs according to their length, arms ratio and the position of their centromere. Chromosome pairs were classified according to the arms ratio (AR) as metacentric (if AR = 1.7), submetacentric (if AR between 1.7 to 3.0), subacrocentric (if AR between 3.0 and 7.0) or telocentric (if AR > 7.0). Chromosome pairs were classified according to their relative length coefficient (RL) as long (L, if RL > 1.26), median long (M2, if RL between 1.01 and 1.25), median (M1, if RL between 0.75 and 1.00) or short (S, if RL < or = to 0.75) (Kutarekar and Wanjari, 1983; Kakeda and Fukui, 1994).

### Flow cytometry analysis

To analyze nuclear DNA content, nuclei were extracted. The one-step protocol with Galbraith's Buffer (and modifications that were made to this buffer) proved to be unsuccessful for isolation of nuclei from *J. chocola* and *Carica papaya* leaves, but not for *Phaseolus vulgaris*, making it necessary to choose a two-step protocol for nuclei isolation. Therefore, a procedure to analyze nuclear DNA content from papaya leaves was modified, which worked with all three species (Araújo et al., 2010). Nuclei suspensions were prepared from fragments of young leaves (2 cm<sup>2</sup>) of the standards (*Phaseolus vulgaris*, cv. Valentino, 2n = 22, 1C-value = 0.60 pg and *Carica papaya*, cv. SunUp, 2n = 18, 1C-value = 0.325 pg) and the female or the male samples (*J. chocola*). Two biological replicates were analyzed using common bean (*P. vulgaris*) as the standard, and seven biological replicates were analyzed using papaya (*C. papaya*) as the standard. To obtain a nuclei suspension, the leaf tissue was placed in a sterile 15 x 90 mm plastic petri dish on ice and sliced into 0.25 mm to 1 mm segments with a new razor blade in 500 µl of Otto-I Buffer (0.1M citric acid monohydrate 0.1M, 0.5% tween 20, 2.0 mM dithiothreitol, 50 µg/ml RNase and 1X Halt™ Protease Inhibitor Cocktail). After the tissue was sliced, 500µl of Otto-I Buffer were added and the suspension was filtered through a 40 µm nylon filter and collected into a 1.5 ml microcentrifuge tube. To purify the nuclei, the tubes were centrifuged at 100 g for 10 min at 4°C and the supernatant was poured out until the 50 µl mark on the microcentrifuge tube. Another 100 µl of Otto-I Buffer were added (for a total of 150 µl) to the nuclei pellet and the pellet was carefully resuspended with a micropipette and incubated for 10 min on ice. After incubation, purified nuclei were stained for 30 min at 4°C with 150 µl Otto-I:Otto-II Buffer Solution (1:2 proportion, supplemented with 75 µM propidium iodide, 2.0 mM dithiothreitol, 50 µg/ml RNase and 1X Halt™ Protease Inhibitor Cocktail).

Stained nuclei were analyzed at the Flow Cytometry Facility of the Roy J. Carver Biotechnology Center (CBC) at the University of Illinois, with a BD FACS Diva software. Important flow cytometer parameters (like wavelength and voltage) were determined based on a preliminary and separate analysis of the standards. Not less than 10 000 events were measured for each sample. FCS files were used to build flow cytometry histograms using the software FCS Express 6 Flow Research Edition (Version 6.06.0021, De Novo Software, CA, USA).

### Nuclear DNA content and genome size estimation

Genome size estimates are described as "C" values, with 2C representing the value for a diploid genome estimated by flow cytometry, and 1C representing the value for a haploid genome (VanBuren et al., 2011). 2C-values were calculated with the mean position of the G0/G1 peaks from flow cytometry results (Araújo et al., 2010), using the following equation:

$$2C - \text{value (pg)} = \frac{\text{G0/G1 peak position of the sample} \times \text{Standard 2C DNA content (pg)}}{\text{G0/G1 peak position of the standard}}$$

The mean 2C-value was calculated using two biological replicates when using common bean (*P. vulgaris*) as the standard, and seven biological replicates when using papaya (*C. papaya*) as the standard (Table 2).

1C-values for nuclear DNA content in picograms were calculated from the 2C-values using the following equation:

$$1C - \text{value (pg)} = \frac{2C - \text{value (pg)}}{2}$$

The mean 1C-value was calculated using two biological replicates when using common bean (*P. vulgaris*) as the standard, and seven biological replicates when using papaya (*C. papaya*) as the standard (Table 2).

The 2C-value for each sample in picograms (pg) was used to calculate the diploid genome size (Mbp) using a conversion factor of 1 pg = 978 Mbp (Araújo et al., 2010; VanBuren et al., 2011). For sequencing purposes, the genome size per haploid genome in basepairs (Mbp) was estimated from the calculated 2C-values using the following equation:

$$\begin{aligned} \text{Genome size per haploid genome} \\ = 2C - \text{value (pg)} \times \frac{978 \text{ Mbp}}{1 \text{ pg}} \times \frac{1}{2} \end{aligned}$$

The mean genome size per haploid genome was calculated using two biological replicates when using common bean (*P. vulgaris*) as the standard, and seven biological replicates when using papaya (*C. papaya*) as the standard (Table 2).

### Statistical analysis

A two-tail T-test was performed to detect any significant difference between female and male plants regarding their nuclear DNA content (pg) and their genome size (bp).

### Acknowledgments

This work was supported by National Science Foundation (NSF) Plant Genome Research Program Award 1546890 to Dr. Ray R. Ming. The authors would like to thank Dr. A. Lane Rayburn for providing access and space to his laboratory at the University of Illinois, with the necessary equipment (microscope and camera) to record the results. Authors would like to thank Dr. Barbara Pilas for helping with instrument operation, training and protocol development for flow cytometry analysis at her laboratory in the Flow Cytometry Facility of the University of Illinois. The authors would like to thank Dr. Melina Mancini and Mr. Ernesto Machado for their valuable comments to the manuscript.

### References

- Araújo FS, Carvalho CR, Clarindo WR (2010) Genome size, base composition and karyotype of *Carica papaya* L. Nucleus. 53:25-31.
- Cardoza Ruiz FS, Ledis Linares J, Aguilar Grijalva AE (2014) *Jarilla chocola* (Caricaceae), un nuevo género y especie para la flora de El Salvador. J Bot Res Inst Tex. 8(2):595-601.
- Carvalho FA, Renner SS (2012) A dated phylogeny of the papaya family (Caricaceae) reveals the crop's closest relatives and the family's biogeographic history. Mol Phylogenet Evol. 65(1):46-53.
- Carvalho FA, Renner SS (2013) Correct names for some of the closest relatives of *Carica papaya*: a review of the mexican/guatemalan genera *Jarilla* and *Horovitzia*. PhytoKeys. 29:63-74.
- Corrêa Damasceno Junior P, Da Costa FR, Pereira TNS, Neto MF, Pereira MG (2009) Karyotype determination in three Caricaceae species emphasizing the cultivated form (*C. papaya* L.). Caryologia. 62(1):10-15.

- Gschwend AR, Wai CM, Zee F, Arumuganathan AK, Ming R (2013) Genome size variation among sex types in dioecious and trioecious *Caricaceae* species. *Euphytica*. 189:461-469.
- Iovene M, Yu Q, Ming R, Jiang J (2015) Evidence for emergence of sex-determining gene(s) in a centromeric region in *Vasconcellea parviflora*. *Genetics*. 199(2):413-421.
- Kakeda K, Fukui K (1994) Dynamic changes in the morphology of barley chromosomes during the mitotic metaphase stage. *Jpn J Genet*. 69(5):545-554.
- Kutarekar DR, Wanjari KB (1983) Karyomorphological studies in some of the varieties of bengal gram (*Cicer arietinum*, Linn.). *Cytologia*. 48(3):699-705.
- Rockinger A, Sousa A, Carvalho FA, Renner SS (2016) Chromosome number reduction in the sister clade of *Carica papaya* with concomitant genome size doubling. *Am J Bot*. 103(6):1082-1088.
- Schindelin J, Rueden CT, Hiner MC, Eliceiri KW (2015) The ImageJ ecosystem: an open platform for biomedical image analysis. *Mol Reprod Dev*. 82(7-8):518-529.
- Tookey HL, Gentry HS (1969) Proteinase of *Jarilla chocola*, a relative of papaya. *Phytochemistry*. 8(6):989-991.
- VanBuren R, Li J, Zee F, Zhu J, Liu C, Arumuganathan AK, Ming RR (2011) Longli is not a hybrid of longan and lychee as revealed by genome size analysis and trichome morphology. *Trop Plant Biol*. 4(3-4):228-236.
- Willingham BC, White GA (1976) Agronomic evaluation of prospective new crop species. v. *Jarilla chocola* - a proteinase source. *Econ Bot*. 30:189-192.
- Wu X, Wang J, Na J, Yu Q, Moore RC, Zee F, Huber SC, Ming R (2010) The origin of the non-recombining region of sex chromosomes in *Carica* and *Vasconcellea*. *Plant J*. 63(5):801-810.