

Morphological index and flow cytometry as tools for discriminating Musa L. germplasms from Northeast India for estimation of ploidy and genome composition

Rudra Prasad Banerjee

CSIR-National Botanical Research Institute

Padma Tamang

CSIR-National Botanical Research Institute

Anath Bandhu Das

Utkal University

Rabindra Kumar Sinha

Tripura University

Satya Narayan Jena (✉ satyanarayanjena2005@yahoo.co.in)

CSIR-National Botanical Research Institute <https://orcid.org/0000-0003-3472-4018>

Research Article

Keywords: 2C DNA, genome constitution, North-east India, Musa, morphological scoring

Posted Date: July 19th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-3168973/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Genome characterization is an essential aspect for understanding the origin and physiological adaptation of any species. The size of the genome, its constitution, and ploidy level provide the valuable information on the taxonomic relationship, which is further utilized in genetic crop improvement program concentrating on traits of interest. *Musa* sp., also known as the banana/plantain, is a fruit and vegetable that is grown all over the world. Because of the inter- and intra-specific hybridization with their wild diploids, the majority of banana cultivars are triploid and seedless in nature. Today's banana cultivars have different genome sizes and ploidies as a consequence, which contributes to their unique shape and climate adaptation. However, limited information is available on the genomic resources of wild and cultivated bananas commonly found in the northeastern region of India. In this study, we performed the morphological characterization, genome size estimation, and ploidy analysis of 16 *Musa* germplasm collected from northeast India. Morphological characterization and flow cytometry-based (FCM) 2C DNA content revealed significant variation in the ploidy level among the collected *Musa* germplasms. Further, the present result revealed that the morphological scoring based on ploidy and genome composition was not always consistent with the flowcytometry DNA content data. The neighbor-joining (NJ) tree based on morphology scoring of *Musa* sp. var. *fushrey jhapari* (NBRI-0020) clustered into the diploid group (Group II), however, it was not a diploid, which confirmed in flow cytometry data-based NJ tree (Group I). It is further revealed that the *Musa* sp. var. *fishery jhapari* (NBRI-0020) had a total morphological score of 54, which did not fall under any genome composition category, however, it showed relatively closer to diploid group in the morphology-based NJ Tree. In the contrary, *Musa balbisiana* var. *jungli kol* (NBRI-0027) and *Musa* sp. var. *jungli kol* (NBRI-0028) were found to be diploid in flow cytometry data, while those on morphology scoring were categorized as triploid, thereby flow cytometric data must be supplemented with morphology-based data for an accurate genome composition categorization. Our findings in this present study provides valuable insight into the composition of wild and cultivated *Musa* genome. In most of the cases (~ 68.75%; 11 out of 16) of Northeast India *Musa* genotypes, the morphology scoring-based genome composition did not match with FCM-based genome composition, indicating the morphology-based genome composition of *Musa* genotypes from North-East India is not straightforward and accurate. This present data on *Musa* genotypes of Northeast India will supplement the morphological traits-driven breeding program for banana genetic improvement. Furthermore, our findings shed light on the genome evolution of bananas in north-east India which is still remained unexplored because of its complex evolutionary processes.

Introduction

Bananas and plantains (*Musa* sp.), commonly regarded as herbaceous perennial members of the Musaceae family, are commercially significant fruit crops and widely farmed in tropical and subtropical countries. (Lysák et al. 1999; Asif et al. 2001). It is the fourth-most widely grown crop in the world and is worldwide staple food. The cultivated species are generally grown by small-holder farmers out of which ~ 85% is consumed domestically. In India, the cultivation of banana is commonly practiced throughout the country that contribute a massive 22.15% of the global production (<https://www.promusa.org/India>).

However, the distribution of *Musa* is restricted in the North-East India, Western Ghats, and Andaman and Nicobar Island (Prasad et al. 2013; Hareesh et al. 2017) which harbour genetically diverse banana genotypes of cultivated, wild, seeded and non-pulpy varieties (Uma et al. 2006). The majority of the edible bananas are seedless with parthenocarpic fruits, that have been originated through the inter- and intra-specific hybridization between wild and seeded banana. The wild and cultivated *Musa* varieties also differ in ploidy level and genome size, and morphological traits. The wild *M. acuminata* and *M. balbisiana* are diploid and seeded in nature with a genome composition of AA and BB respectively. However, varieties originated through cross-breeding between these two wild cultivars are polyploids, seedless having distinct genome composition of A- and B-genome (Uma et al. 2006). Such natural variations in the ploidy and genome composition resulted in the origin of genetically diverse *Musa* groups. The cultivated varieties have economic importance as food and fibre, but a limited genetic information is available on their wild relatives. Therefore, it is crucial to expand valuable insights on

taxonomic and genomic composition of wild and cultivated *Musa* varieties to implement a successful breeding strategy for developing elite *Musa* varieties for disease resistance and greater productivity.

For several decades, the morpho-taxonomical characterization of banana using scoring index was considered as a reliable approach for estimation of ploidy and genome composition (Simmonds and Shepherd 1955). Besides, other physiological traits such as stomatal size, density, pollen size and chromosome numbers have also been considered for determining ploidy in *Musa* (Hamill et al. 1992). However, as a result of inconsistency in data reproducibility and reliability due to genotypic influences, these approaches have failed to produce the genomic information in an efficient way (Vandenhout et al. 1995). The morphological scoring method involves a set of 15 characters including the vegetative and reproductive phases of banana, which has further been assigned a scoring index between 0 to 5 (Simmonds and Weatherup 1990). The cumulative value of all the taxonomical traits result in the characterization of an unknown *Musa* genotypes into different genomic groups of different ploidy. Though this approach is effective in determining the ploidy and genome composition of many cultivars, it is still obscure and requires expertise in taxonomy. Furthermore, the complexity in the origin and diversification of diploid cultivars such as *M. acuminata* could not be completely resolved using only morphological data. Therefore, it is utmost important to focus on other alternative approaches that can further be supplemented with morpho-taxonomical characterisation to strengthen the conceptual framework on ploidy and genome composition in banana.

In recent times, the advancement of Flow cytometry (FCM) has provided an opportunity to work with the nuclear DNA contents in plants (Doležel and Bartoš 2005). This facilitates the appropriate measurement of nuclear DNA content and genome size of any unknown sample in a time efficient manner (Doležel et al. 2007; Galbraith 2012). Besides, the ease of sample preparation, processing and data analysis has made this a convenient approach for its application in fundamental research like ploidy analysis, nuclear DNA content (2C DNA) estimation, and genome size very accurately in cellular level (Doležel et al. 2007). In general, FCM involves the use of two common intercalating dyes namely Propidium Iodide (PI) and 4,6-diamidino-2-phenylindol (DAPI); PI intercalates randomly in the genome without any base-dependent bias, while DAPI binds specifically with the AT rich regions (Doležel and Bartoš 2005; Doležel et al. 2007). Such unique criteria provide an advantage of using PI over DAPI for estimating 2C DNA content and genome size with greater coverage.

In this study, we have combined the morphological index and flow cytometry to determine the ploidy level and estimation of genome size in 16 *Musa* germplasms collected from the North-Eastern India to investigate the genomic DNA content variation among cultivated and wild *Musa* varieties to understand their origin and phylogenetic relationships. We further tried to re-examine the morphological scoring data with flow cytometry data of 16 studied *Musa* germplasms to determine the ploidy level and genome composition.

Materials and methods

Plant material used

A total number of 16 *Musa* germplasms were collected from different states of North-East India and grown at the experimental field of CSIR-NBRI, Lucknow (Table 1). The seeds of the internal reference standard banana (FHIA 03) obtained from National Research Centre of Banana (NRCB), Tamil Nadu, India and were also grown in CSIR-NBRI experimental field. The external reference standards namely *Solanum lycopersicum* and *Glycine max* used in flow cytometry were grown from the seeds provided by Dr. Jaroslav Doležal from the Experimental Institute of Botany, Czech Republic. Fresh leaf samples of *Musa* germplasms and external and internal germplasms were collected from the field grown samples and subsequently processed for sample preparation and analysis in flow cytometer.

Table 1

Summary of ploidy and genome composition of 16 *Musa* germplasms using the morpho taxonomical scoring index proposed by Simmonds and Shephard (1955)

Sl No.	Species name/ Local name	Area of collection	Coordinates	Accession number	Accession type	Morpho-taxonomical Scoring	Ploidy level	Genome composition
1	<i>Musa balbisiana</i>	Majhitar, Sikkim	27°11'27.98"N 88°33'10.27"E	NBRI-002	Wild	71	Diploid	BB
2	<i>Musa</i> sp. var. <i>jungli kol</i>	Rangli, Sikkim boarder	27°10'58.13"N 88°32'18.41"E	NBRI-003	Wild	55	Diploid	AB
3	<i>Musa balbisiana</i>	Baghmara, West Bengal	26°49'12.36"N 88°20'37.41"E	NBRI-006	Wild	71	Diploid	BB
4	<i>Musa acuminata</i> var. <i>malbhog</i>	Lulagaon, West Bengal	27°02'53.11"N 88°52'27.30"E	NBRI-0012	Cultivated	33	Triploid	AAB
5	<i>Musa balbisiana</i>	Lulagaon, West Bengal	27°02'44.33"N 88°50'50.88"E	NBRI-0013	Wild	75	Diploid	BB
6	<i>Musa</i> sp. var. <i>fushrey jhapari</i>	Rangli, Sikkim boarder	27°10'26.71"N 88°31'49.47"E	NBRI-0020	Cultivated	54	Triploid	ABB
7	<i>Musa acuminata</i> var. <i>chini champa</i>	Lebong, West Bengal	27°05'27.79"N 88°27'24.47"E	NBRI-0026	Cultivated	36	Triploid	AAB
8	<i>Musa balbisiana</i> var. <i>jungli kol</i>	Namrup, Assam	27°21'11.34"N 95°31'52.90"E	NBRI-0027	Wild	38	Triploid	AAB
9	<i>Musa</i> sp. var. <i>Jungli Kol</i>	Narem pothar, Assam	27°18'31.97"N 95°28'35.06"E	NBRI-0028	Wild	38	Triploid	AAB
10	<i>Musa acuminata</i> var. <i>kachkol</i> (NBRI-0033)	Bagmara Thaichimog para, Tripura	23°55'26.62"N 91°50'42.67"E	NBRI-0033	Cultivated	42	Triploid	AAB

Sl No.	Species name/ Local name	Area of collection	Coordinates	Accession number	Accession type	Morpho-taxonomical Scoring	Ploidy level	Genome composition
11	<i>Musa</i> sp. var. <i>jungli kol</i>	Dharampur, Tripura	23°46 23.79''N 91°43 54.72''E	NBRI-0035	Wild	49	Diploid	AB
12	<i>Musa acuminata</i> var. <i>gopi</i>	Madhupur, Tripura	23°45 43.82''N 91°15 37.32''E	NBRI-0036	Cultivated	60	Triploid	ABB
13	<i>Musa paradisiaca</i> var. <i>kachkol</i>	Madhupur, Tripura	23°45 52.29''N 91°15 29.89''E	NBRI-0037	Cultivated	59	Triploid	ABB
14	<i>Musa</i> sp. var. <i>jungli kol</i>	Pandabpur, Tripura	23°45 14.67''N 91°15 11.97''E	NBRI-0038	Wild	41	Triploid	AAB
15	<i>Musa acuminata</i> var. <i>sobri</i>	Khowai, Tripura	24°03 59.70''N 91°36 21.09''E	NBRI-0039	Cultivated	43	Triploid	AAB
16	<i>Musa acuminata</i> var. <i>champa</i>	Suryamani nagar, Tripura	23°45 29.55''N 91°16 43.58''E	NBRI-0041	Cultivated	38	Triploid	AAB

Morpho-taxonomical scoring

Morphological characterization was carried out by collecting the vegetative and reproductive parts of the plant from the field grown samples. Cross sections were made from each of the selected plant portion and detailed observations were carried out using a simple microscope. Subsequently, detailed photographs of all the studied plant parts were done for preparing morphological plates. The morphological characterization of 16 *Musa* germplasms was carried out following the method of Simmonds and Shepherd (1955). A total of 15 morphological parameters from the vegetative and reproductive stages were chosen for ploidy and genome characterization. These characters were selected on the basis of their discrimination ability between *M. acuminata* and *M. balbisiana* genome. The scoring system was defined as 1 and 5, where providing a score of 1 to a particular trait corresponded to its closeness with *M. acuminata* (AA) genome while a score of 5 was assigned with *M. balbisiana* (BB) genome. In addition, any score falling within the range of 2–44 was also considered to be closely associated to *M. balbisiana* based on the intensity of a particular character. Lastly, the individual scores of all the 15 characters were combined and grouped into different genome compositions following Simmonds and Shepherd's (1955) scoring index, and modified by Singh and Uma (1996) (Simmonds and Shepherd 1955; Atom et al. 2015) scoring index. The combined score for all the 15 germplasms ranged between 15 and 75 for *M. acuminata* and *M. balbisiana* respectively.

Preparation of samples for FCM analysis

Approximately 50 mg of fresh leaf samples were finely chopped with a sharp blade in a sterile petri plate using 500 µl nuclei extraction buffer with minor modifications (Doležel et al. 2007). The chopping process was performed by keeping the

petri plates on a bucket filled with ice to reduce any metabolic activity. Samples were carefully chopped for 5–10 sec and the entire cell suspension was filtered through a 30-micron mesh filter (Partec) and the filtrate was collected in a sterile collection tube. The remaining tissue debris accumulated on the filter was discarded. Then, 1.5 ml of suspension buffer containing propidium iodide (PI) was added to the filtrate and the entire solution was gently mixed through pipetting. The homogenized solution was incubated for 1h at dark condition and which was subjected to analysis using Sysmex CyFlow Cube 8 Partec II Flow Cytometer. Similar methodology was followed to prepare the reference standards where both the sample and reference were chopped together and processed subsequently (Doležel and Bartoš 2005; Doležel et al. 2007).

The ploidy level was determined by comparing the peaks of the samples with that of an internal reference standard viz., FHIA 03 (genome composition = AABB) while the estimation of genome size was achieved by calculating the 2C DNA content of the samples and comparing it with the external standards viz., *Solanum lycopersicum* L. cv. Stupické polní rane (2C = 1.96 pg) and *Glycine max* Merr. Cv. Polanka (2C = 2.50 pg). The threshold and voltage for FL2 filter was adjusted to 400 while the same for FL3 was fixed to 400 and 480 respectively, to obtain better G₀/G₁ peak position and low percentage of Coefficient of Variation (CV%) (Ulrich and Ulrich 1991). An average number of 5,000 nuclei were counted per sample and the total number of events were fixed to 4100. The flow rate was set to 0.1 µl/ min and absorbance of PI was recorded at 488 nm. All the samples and standards were run in triplicates and the mean value of peak median was calculated for genome size estimation. The data analysis was performed in FCS express software provided by the manufacturer. Ploidy level was determined by using the following formula.

$$\text{Sample ploidy} = \frac{\text{Mean of sample peak}}{\text{Mean of standard peak}} \times \text{standard ploidy}$$

The 2C DNA content of the sample was estimated by the following formula;

$$\text{Sample 2C DNA (pg)} = \frac{\text{Mean of sample 2C peak}}{\text{Mean of standard 2C peak}} \times \text{standard 2C value}$$

The genome composition and size of all the 16 *Musa* germplasms were determined by comparing the 2C DNA content of the studied samples with the internal reference standards such as *M. lateriata* (AB), Jurmany (BB), and FHIA 03 (AABB).

Statistical analysis

Statistical analysis was performed using 'R' version 3.6.2 (64-bit). Taxonomic grouping of 16 *Musa* germplasms was done by Hierarchical clustering method. The complete linkage method was chosen for the analysis following the Euclidean distance as dissimilarity coefficient. Besides, one-way ANOVA was performed to compare the variation in nuclear DNA content between the sample and standard groups.

Results

Morpho-taxonomical scoring and genome composition

The vegetative and reproductive characters of each accession of 16 *Musa* germplasms collected from various states of North-East India was further examined after dissection (Table 1, Fig. 1a-p). Scoring index ranged between 33–75 across all the 16 *Musa* germplasms. The lowest scoring of 33 was achieved by *M. acuminata* var. *malbhog* (NBRI-0012), while the highest of 75 was noticed for *M. balbisiana* (NBRI-0013) (Table 1). Among the 16 germplasms, 11 were triploid with a genome composition of either AAB or ABB, whereas 5 germplasms exhibited diploid with either AB or BB genome. Three germplasms namely *M. balbisiana* (NBRI-002), *M. balbisiana* (NBRI-006) and *M. balbisiana* (NBRI-0013) were classified

into BB genome category with a score of 71, 71 and 75, respectively. Similarly, two germplasms viz., *Musa* sp. var. *jungli kol* (NBRI-003) and *Musa* sp. var. *jungli kol* (NBRI-0035) were categorised as AB having a score of 55 and 49, respectively. Eight germplasms (*M. acuminata* var. *malbhog* (NBRI-0012), *M. acuminata* var. *chini champa* (NBRI-0026), *M. balbisiana* var. *jungli kol* (NBRI-0027), *M. sp.* var. *jungli kol* (NBRI-0028), *M. acuminata* var. *kachkol* (NBRI-0033), *Musa* sp. var. *jungli kol* (NBRI-0038), *M. acuminata* var. *sobri* (NBRI-0039) and *M. acuminata* var. *champa* (NBRI-0041) were represented by AAB genome group with a score ranging between 33 to 42. The genome group ABB was observed in *Musa* sp. var. *fushrey jhapari* (NBRI-0020), *M. paradisiaca* var. *kachkol* (NBRI-0037), and *M. acuminata* var. *gopi* (NBRI-0036) with a score of 54, and, 59 and 60, respectively.

The dendrogram based on the morphological scoring index of 16 germplasms revealed the presence of two distinct groups i.e., Group-I (Mostly triploids), and Group-II (mostly diploids). The group I (Mostly triploids) was further sub-divided into two sub-groups based on similarity coefficient on the basis of morphological scoring and genome composition (Fig. 2). The first subgroup of Group I comprised of 5 *Musa* germplasms such as *M. acuminata* var. *chini champa* (NBRI-0026), *Musa* sp. var. *jungli kol* (NBRI-0027), *M. acuminata* var. *malbhog* (NBRI-0012), *Musa acuminata* var. *champa* (NBRI-0041), and *Musa* sp. var. *jungli kol* (NBRI-0038) while the second subgroup of group I contained *Musa* sp. var. *jungli kol* (NBRI-0028), *M. acuminata* var. *sobri* (NBRI-0039), and *M. acuminata* var. *kachkol* (NBRI-0033). The Group-II contained mostly diploid including 3 triploid germplasms. The germplasms such as *M. balbisiana* (NBRI-006), *M. balbisiana* (NBRI-0013), and *M. balbisiana* (NBRI-002), with BB genome composition formed one clade. The two triploids such as *M. acuminata* var. *gopi* (NBRI-0036), and *M. paradisiaca* var. *kachkol* (NBRI-0037) clustered in this group II with closer affinity to germplasms with BB-genome and other triploid (*Musa* sp. var. *fushrey jhapari* (NBRI-0020) clustered in this group II in between AB genome germplasm (*Musa* sp. var. *jungli kol* (NBRI-003) and *Musa* sp. var. *jungli kol* (NBRI-0035).

2C DNA content, ploidy and genome size estimation

Estimation of nuclear DNA of 16 *Musa* germplasms using *Glycine max* Merr. Polanka as reference standard revealed variation in the 2C DNA content ranging from 1.09-1.74 pg (Table 2). Subsequently, determination of ploidy using as *Musa* internal reference [FHIA-03 (AABB), *M. laterita* (AB) and Jurmany (BB)] resulted in the identification of 7 diploids and 9 triploid germplasms. The amount of 2C DNA in seven diploid *Musa* varieties ranged between 1.09 pg to 1.14 pg while the same varied between 1.54pg to 1.74 pg in the remaining nine triploids (Fig. 3, Table 2). Unlike to the morphology-based data, the flow cytometry data revealed seven germplasm as diploid while nine were triploids. *Musa* sp. var. *jungli kol* (NBRI-27) and *Musa* sp. var. *jungli kol* (NBRI-28) which were classified as triploid based on morphological scoring were found to be diploid in FCM data. Further, these two germplasms were also found to contained lowest DNA content (1.09pg) among all studied germplasm. The coefficient of variation (CV%) ranged between 2.42–4.81 across all the studied germplasms. Among the diploids, the highest DNA content (1.14pg) as recorded in *Musa* sp. var. *jungli kol* (NBRI-003) followed by 1.12pg in *M. balbisiana* (NBRI-006) and *Musa* sp. var. *jungli kol* (NBRI-0035), 1.11pg in *M. balbisiana* (NBRI-002) and *M. balbisiana* (NBRI-0013).

Table 2
Estimation of genome size, ploidy and 2C DNA content in 16 *Musa* germplasms using *Glycine max* Merr. Polanka as external reference standard

Sl no.	Species/local name	<i>Musa</i> accessions	2C DNA content (pg)		Ploidy level		Genome size 1C (Mb)
			Median	<i>Glycine</i> reference	CV% Standard	Internal <i>Musa</i>	<i>Glycine</i> reference
1	<i>Musa balbisiana</i>	NBRI-002	115397.5±0.00474	1.11	4.00	Diploid	542.79
2	<i>Musa</i> sp. var. <i>jungli kol</i>	NBRI-003	118287.67±0.00421	1.14	2.42	Diploid	557.46
3	<i>Musa balbisiana</i>	NBRI-006	116525.17±0.00474	1.12	4.02	Diploid	547.68
4	<i>Musa acuminata</i> var. <i>malbhog</i>	NBRI-0012	169958.16±0.0073	1.64	4.30	Triploid	801.96
5	<i>Musa balbisiana</i>	NBRI-0013	114969±0.00489	1.11	4.26	Diploid	542.79
6	<i>Musa</i> sp. var. <i>fushrey jhapari</i>	NBRI-0020	174498.5±0.00687	1.68	3.94	Triploid	821.52
7	<i>Musa acuminata</i> var. <i>chini champa</i>	NBRI-0026	160008±0.00768	1.54	4.81	Triploid	753.06
8	<i>Musa balbisiana</i> var. <i>jungli kol</i>	NBRI-0027	113302.17±0.00458	1.09	4.05	Diploid	533.01
9	<i>Musa</i> sp. var. <i>Jungli Kol</i>	NBRI-0028	113225.67±0.00449	1.09	3.97	Diploid	533.01
10	<i>Musa acuminata</i> var. <i>kachkol</i>	NBRI-0033	177832±0.00592	1.71	3.33	Triploid	836.19
11	<i>Musa</i> sp. var. <i>jungli kol</i>	NBRI-0035	116237.33±0.00405	1.12	3.48	Diploid	547.68
12	<i>Musa acuminata</i> var. <i>gopi</i>	NBRI-0036	180295.5±0.0069	1.74	3.83	Triploid	850.86
13	<i>Musa paradisiaca</i> var. <i>kachkol</i>	NBRI-0037	170076.83±0.00643	1.64	3.77	Triploid	801.96
14	<i>Musa</i> sp. var. <i>jungli kol</i>	NBRI-0038	181102.33±0.00727	1.74	4.02	Triploid	850.86
15	<i>Musa acuminata</i> var. <i>sobri</i>	NBRI-0039	165452±0.00751	1.59	4.55	Triploid	777.51

Sl no.	Species/local name	<i>Musa</i> accessions	2C DNA content (pg)		Ploidy level		Genome size 1C (Mb)
16	<i>Musa acuminata</i> var. <i>champa</i>	NBRI-0041	171007±0.00749	1.65	4.39	Triploid	806.85
17	<i>Musa balbisiana</i>	Reference standard	125737.33±0.00549	1.21	4.38	Diploid	591.69
18	<i>Musa</i> sp. var. <i>jungli kol</i>	Reference standard	115728.16±0.00495	1.11	4.28	Diploid	542.79
19	<i>Musa balbisiana</i>	Reference standard	232157.33±0.00984	2.24	4.24	Tetraploid	1095.36
20	<i>Musa acuminata</i> var. <i>malbhog</i>	Reference standard	258848.83±0.00814	2.50	3.15	Diploid	1222.5

Cluster analysis using the Euclidean distance as dissimilarity coefficient of FCM data revealed that all the seven diploids [*M. balbisiana* (NBRI-002), *Musa* sp. var. *jungli kol* (NBRI-003), *M. balbisiana* (NBRI-006), *M. balbisiana* (NBRI-0013), *Musa* sp. var. *jungli kol* (NBRI-0027), *Musa* sp. var. *jungli kol* (NBRI-0028) and *Musa* sp. var. *jungli kol* (NBRI-0035)] were clustered as group-II while nine triploids were clustered in group-I (Fig. 3). In our flow cytometry data based diploid cluster of Neighbour Joining (NJ) tree, all the diploid was of BB genome type while the same was little differentiated in our NJ tree of morphology index. However, unlike the diploids, genome composition in the nine triploids of Group-I NJ tree of Flow cytometry data resulted in two categories of genome composition viz., AAB and ABB.

Majority of the germplasms viz., *M. acuminata* var. *malbhog* (NBRI-0012), *Musa* sp. var. *fushrey jhapari* (NBRI-0020), *M. acuminata* var. *kachkol* (NBRI-0033), *M. acuminata* var. *gopi* (NBRI-0036), *M. paradisiaca* var. *kachkol* (NBRI-0037), *Musa* sp. var. *jungli kol* (NBRI-0038) and *M. acuminata* var. *champa* (NBRI-0041) belonged to AAB category while only two germplasms namely *M. acuminata* var. *chini champa* (NBRI-0026) and *M. acuminata* var. *sobri* (NBRI-0039) were represented in AAB group. Statistical analysis performed using one-way ANOVA test indicated significant variation in 2C DNA content within and among the *Musa* germplasms ($P < 0.001$). Subsequently, the estimated genome size of 16 *Musa* germplasms ranged between 533Mb in, genotype name with *M. balbisiana* var. *jungli kol* (NBRI-0027), and *M. balbisiana* var. *jungli kol* (NBRI-0028) and to 850 Mb – the largest genome size in *M. acuminata* var. *gopi* (NBRI-0036) and *Musa* sp. var. *jungli kol* (NBRI-0038). The genome size of the diploid cultivars ranged between 533–557 Mb while the same varied between 753–850 Mb in the triploid varieties. Comparing the results obtained through morphometric and FCM based analysis on ploidy and genome composition revealed differences in genome composition in majority of the studied *Musa* germplasms (Table 3). The G_0/G_1 peaks have shown a positional shift based on their ploidy and 2C DNA content when compared with the reference standard (Fig. 4, 5).

Table 3
Comparison of ploidy and genome composition of 16 *Musa* germplasms using morpho taxonomical score and flow cytometric analysis

SI No.	Species/local name	<i>Musa</i> accessions	Ploidy level and genome composition			
			Morpho taxonomy		FCM analysis	
			Ploidy	Genome group	Ploidy	Genome group
1	<i>Musa balbisiana</i>	NBRI-002	Diploid	BB	Diploid	AB
2	<i>Musa</i> sp. var. <i>jungli kol</i>	NBRI-003	Diploid	AB	Diploid	AB
3	<i>Musa balbisiana</i>	NBRI-006	Diploid	BB	Diploid	AB
4	<i>Musa acuminata</i> var. <i>malbhog</i>	NBRI-0012	Triploid	AAB	Triploid	AAB
5	<i>Musa balbisiana</i>	NBRI-0013	Diploid	BB	Diploid	AB
6	<i>Musa</i> sp. var. <i>fushrey jhapari</i>	NBRI-0020	Triploid	ABB	Triploid	AAB
7	<i>Musa acuminata</i> var. <i>chini champa</i>	NBRI-0026	Triploid	AAB	Triploid	ABB
8	<i>Musa balbisiana</i> var. <i>jungli kol</i>	NBRI-0027	Triploid	AAB	Diploid	BB
9	<i>Musa</i> sp. var. <i>Jungli Kol</i>	NBRI-0028	Triploid	AAB	Diploid	BB
10	<i>Musa acuminata</i> var. <i>kachkol</i> (NBRI-0033)	NBRI-0033	Triploid	AAB	Triploid	AAB
11	<i>Musa</i> sp. var. <i>jungli kol</i>	NBRI-0035	Diploid	AB	Diploid	AB
12	<i>Musa acuminata</i> var. <i>gopi</i>	NBRI-0036	Triploid	ABB	Triploid	AAB
13	<i>Musa paradisiaca</i> var. <i>kachkol</i>	NBRI-0037	Triploid	ABB	Triploid	AAB
14	<i>Musa</i> sp. var. <i>jungli kol</i>	NBRI-0038	Triploid	AAB	Triploid	AAB
15	<i>Musa acuminata</i> var. <i>sobri</i>	NBRI-0039	Triploid	AAB	Triploid	ABB
16	<i>Musa acuminata</i> var. <i>champa</i>	NBRI-0041	Triploid	AAB	Triploid	AAB
17	<i>Musa laterieta</i>	Internal <i>Musa</i> Ref. I	Diploid	AB	Diploid	AB
18	Jurmany	Internal <i>Musa</i> Ref. II	Diploid	BB	Diploid	BB
19	FHIA 03	Internal <i>Musa</i> Ref. III	Tetraploid	AABB	Tetraploid	AABB

All these seven diploid germplasms were found to be composed of BB genome on the basis of their DNA content with comparison to internal *Musa* reference standards (Fig. 6). However, with reference to morphological scoring index, two diploid such as, *Musa* sp. var. *jungli kol* (NBRI-003) and *Musa* sp. var. *jungli kol* (NBRI-0035) showed more affinity towards mixed characters of A and B genome. The FCM data of these two germplasms also showed that the DNA content lied in between the range of BB (*Musa* ref Standard. I: *Musa balbisiana* var *jurmany*) and AB (*Musa* ref standard II: *Musa laterieta*) genome (Fig. 6). The germplasms with the above genomes were not clearly differentiated in our flow cytometric data. However. The DNA content value did not perfectly match with that of AB-genome. We have kept these germplasms in BB-category as the value were closure to the BB-genome with comparison to AB-genome. Similarly, the triploid varieties exhibited least amount of DNA content in *M. acuminata* var. *chini champa* (NBRI-0026) (1.54 pg) and greater amount in *M. acuminata* var. *gopi* (NBRI-0036) and *Musa* sp. var. *jungli kol* (NBRI-0038) (1.74 pg). Comparing the DNA contents of the

internal *Musa* reference standard, seven germplasms [i.e. *M. acuminata* var. *malbhog* (NBRI-0012), *Musa* sp. var. *fushrey jhapari* (NBRI-0020), *M. acuminata* var. *kachkol* (NBRI-0033), *M. acuminata* var. *gopi* (NBRI-0036), *M. paradisiaca* var. *kachkol* (NBRI-0037),] *Musa* sp. var. *jungli kol* (NBRI-0038) and *M. acuminata* var. *champa* (NBRI-0041)] were found to composed of AAB and two (*M. acuminata* var. *chini champa* (NBRI-0026) and *M. acuminata* Colla var. *sobri* (NBRI-0039) were of ABB type among nine triploids.

Discussion

In this study, we have performed the morphological characterisation and genome size estimation of 16 *Musa* accessions collected from the North-east India. Despite being considered North-East, India as one of the centres of origins of *Musa*, a limited number of studies have focused on the wild and cultivated banana species of this region. Furthermore, few studies have also shown that banana cultivars grown in the north-east region harbours an unique gene pool for which 27 wild species from this region exhibits greater genetic diversity (Uma et al. 2006). However, information available in this aspect was inadequate to support this hypothesis. Thus, understanding the morphology, ploidy level and genome composition of wild and cultivated *Musa* germplasms is crucial for genetic improvement of *Musa*.

For several decades, morpho-taxonomical scoring has been used as an effective approach for determination of ploidy and genome composition in banana (Simmonds and Weatherup 1990; Singh et al. 2014; Atom et al. 2015; Deepthi 2016; Nath et al. 2021). The morphological characters optimised by Simmonds and Shepherd (1955) exhibits a perspicuous differentiation across *M. acuminata* (AA) and *M. balbisiana* (BB) and therefore can be used for genome characterization. Providing a score to an unknown accession using the taxonomic scoring index results in the determination of ploidy level and genome composition by comparing with the known standards of *M. acuminata* or *M. balbisiana* genome. Using the taxonomic scoring of Simmonds and Shepherd (1955) and Singh and Uma (1996), our results confirmed that majority of the diploid genomes belonged to BB genome group suggesting their close relatedness with *M. balbisiana*.

This result also suggests that the banana germplasms concentrated in Northeast India are primarily comprised of B-genome. However, among the triploids, a greater number of accessions belonged to AAB genome group indicating a high coverage of *M. acuminata* genome than *M. balbisiana*. This finding strongly supports the event of inter- and intra-specific hybridisation between *M. acuminata* and *M. balbisiana* genome which is a common phenomenon among the wild banana varieties (Lysák et al. 1999; Pillay et al. 2006). Besides, the intermediate height of the pseudostem, wider leaf lamina and development of seedless fruits also indicates the hybrid origin of these varieties (Hapsari et al. 2015). However, lack of any karyotypic data in our study can be considered as one of the limitations which could help in strengthening our findings in addition to the morpho-taxonomical study (Nath et al. 2021).

Flow cytometry is an effective analytical technique with several implications in medical and fundamental research (Doležel et al. 2007). The principle used in this method provides accurate and reliable results in short time without any technical difficulties unlike other sophisticated analytical techniques. FCM based genome size estimation is always relative where a reference standard with known genome size is used to estimate the genome size of the unknown sample (Praça-Fontes et al. 2011; Temsch et al. 2021). This method allows measuring the sample and standard in two independent set of reactions. Subsequently, the fluorescent intensity of the isolated nuclei is usually measured in the G1 phase of cell cycle. The comparative analysis of the peak positions of standard and sample further determines the nuclear 2C DNA content (Bennett and Leitch 2011). However, in certain circumstances, use of external standards is non-advisable as both sample and standard are not treated under identical conditions. In such situations, simultaneous mixing of the sample and standard is not recommended. Although, in case of unknown samples, the standard and samples are recommended to be chopped and stained simultaneously to obtain reliable result.

In this study, we have used flow cytometry method to determine the ploidy level and estimate the 2C DNA content and genome size of 16 *Musa* germplasms. Using *Glycine max* Merr. Polanka as reference standard, our findings revealed a

significant variation in the DNA content with a range of 1.09–1.74 pg across the germplasms. The 2C DNA content among the diploid cultivars with BB genome content ranged from 1.09 to 1.14 pg thus suggesting low level of intraspecific variation within the *M. balbisiana* genome. Such small amount of variation was also reflected in the genome size of these cultivars that varied from 533 Mb to 557 Mb. Likewise, among the triploids with ABB and AAB genome composition, the 2C DNA content and genome size ranged between 1.54pg to 1.74 pg and 753–850 Mb respectively suggesting a variation of approximately 100 Mb in the *M. acuminata* genome. This finding was consistent with the previous results which showed that 2C DNA content of triploid *Musa* cultivars ranged between 1.61pg to 2.23 pg (Kamaté et al. 2001). Variation in the 2C DNA content among the triploids was evident due to the difference in the number of A and B genome constituents as suggested in other studies (Lysák et al. 1999). Furthermore, in comparison with the internal standard *M. acuminata* cv. FHIA 03 (AABB = 2.24 pg), the amount of 2C DNA contents in the diploid and triploid were expected to be 1.12 pg and 1.68 pg respectively. According to our findings, the estimated range of 1.09 pg to 1.14 pg in diploids and 1.54 pg to 1.74 pg DNA in the triploids corroborates with the expected values of 1.12 pg and 1.68 pg without any significant variation. Though the intraspecific variation was low within the diploid and triploid cultivars, it was statistically significant ($p < 0.005$) and may arise due to geographical isolation and/or vegetative propagation of distinct *Musa* cultivars (Lysák et al. 1999). However, despite low level of intraspecific variation within the cultivars, there was an increase of ~ 1.5 fold in the 2C DNA content (1.09 to 1.14 pg vs. 1.54 to 1.74 pg) and genome size (533 to 557 Mb vs. 753–850 Mb) among the *M. balbisiana* (BB) and *M. acuminata* (ABB/AAB) genomes. This finding was consistent with the previous reports on FCM analysis of *Musa* (Doležel et al. 1994; Lysák et al. 1999; Kamaté et al. 2001).

Our findings differ from some of the earlier reports which showed greater intra-specific variations in the 2C DNA content and genome size in other crops (Cavallini and Natali 1991; Cavallini et al. 1993). Nonetheless, whether such variations existed because of plasticity in the genome or due to error in the methodology needs further investigation. In our study, the variation observed in the DNA content and genome size of ABB or AAB cultivars could be due to the change in dose of “A” and “B” present in their genome due to hybridization without change in the chromosome number (Martel et al. 1997; Le Thierry D’Ennequin et al. 1998). Besides, presence of retrotransposons as non-coding or heterochromatic region in the genome may also influence genome size (Bennetzen and Kellogg 1997). Further investigation at genome scale using next-generation sequencing approach is required to address such assumptions. Angiosperms are known to exhibit a significant amount of variation in their genome size ranging from 60 Mb in *Genlisea aurea* to approximately 150 Gb in *Paris japonica* (Lee and Kim 2014). However, despite such variation in the genome size across the tree of life, the total genome content (1C DNA) vs. the level of complexity in organisms remained poorly understood. Such type of paradox was further represented as ‘C-value paradox’ which suggests a negative correlation between the DNA content and genome size. In other words, a larger genome may comprise of fewer number of genetic elements in comparison with a smaller genome with greater genetic elements. While several studies have shown a positive correlation between the genome size and gene content in prokaryotes and primitive eukaryotes, such patterns are missing in higher eukaryotes (Lynch and Conery 2003; Konstantinidis and Tiedje 2004; Gregory 2005). Consequently, these findings have also speculated a disproportionate relationship between genome size and number of genetic elements present on it. According to our findings, the largest genome was noticed in *M. acuminata* var. *gopi* (NBRI-0036) and *Musa* sp. var. *jungli kol* (NBRI 0038) accessions viz., ~ 850Mb with a 2C DNA content of 1.74 pg. Conversely, the smallest genome was recorded in *M. balbisiana* var. *jungli kol* (NBRI-0027) and *Musa* sp. var. *jungli kol* (NBRI 0028) with a genome size and 2C DNA content of ~ 534 Mb and 1.09 pg respectively. These findings clearly indicate a positive correlation between the 2C DNA content and genome size in the studied *Musa* germplasms. However, further investigations are required to confirm the existence of such discrepancies in other plant groups.

Despite using the morphological and FCM based approaches for estimation of ploidy and genome composition, our findings revealed an inconsistency in these two methods. Although majority of the *Musa* varieties in the present study showed a similar ploidy nature irrespective of the two approaches used, the genome composition differed significantly while choosing one method over the other. While the quality of flow cytometric data is reliable, precise and accurate due to

the use of reference standards (Doležel et al. 1994, 2007; Doležel 1997; Doležel and Bartoš 2005), the morpho-taxonomical approach is obsolete. Use of arbitrary scoring index requires expertise in the field of taxonomy. In addition, there could be a high chance of variability in the individual scores against each character depending upon the experimenter. This variation may further influence the total achieved score, thus providing an error in determining the genome composition. We propose herewith an improved taxonomical scoring method that can be robust, reliable, and uniform across any given condition. Besides, complex events such as hybridisation and polyploidisation can also affect the ploidy and genome size in banana (Lysák et al. 1999; Pillay et al. 2006). Thus, considerable attention needs to be given to these factors while characterising the genome using taxonomical keys.

Dendrograms obtained from the cluster analysis using morpho-taxonomical and FCM data revealed distinct grouping of the studied banana cultivars. While the taxonomical scoring approach resulted in two separate groups with genome composition of BB, AB, AAB and ABB, the FCM based data revealed two distinct clusters of diploids and triploids such as BB genome in one group and, ABB and AAB in another group. Though the formation of each cluster is based on the similarity in morphological trait and 2C DNA content among the germplasms, their grouping is expected to be similar irrespective of the method used in our analysis. However, our study did not support the morphological classification with FCM based genome characterisation which could be due to inappropriate taxonomical scoring method, and extensive natural inter- and intra-specific hybridisation resulting in the formation of polyploids. According to our present knowledge, flow cytometric method can efficiently estimate the genome size and ploidy level of bananas and plantains. However, this approach sometimes can also fail to evaluate the genome composition. As in our study itself the germplasm (NBRI-003 and NBRI-0035) that show a typical “A” genome character was also found to comprise 2C genome content similar to that of germplasm with “BB” genome. Although there is an availability of A and B genome specific molecular markers in the database, their reliability is questionable and thus may not provide enough information on genome composition. In other words, despite their abilities to determine the genome composition, these markers does not provide much details on the independent proportions of A and B genomes (Mabonga and Pillay 2017; Singh et al. 2020). In recent years, the advancement of next generation sequencing has provided an opportunity to decipher the genome architecture in a cost-effective manner. Therefore, future studies will focus on incorporating such advancements to unravel the genome size and composition in wild and cultivated banana which can further accelerate its breeding programme for genetic improvement.

Conclusion

Our study highlighted the morphological characterization, ploidy level and genome size estimation of the wild and cultivated banana cultivars of north-east India. Using the taxonomical scoring index and flow cytometry-based analysis, we found significant variation in 2C DNA content, genome size and composition among the diploid and triploid *Musa* cultivars. However, we observed variations in the results of taxonomical and flow cytometry-based methods as these two approaches have shown separate clustering of the studied banana accessions. Therefore, careful evaluation shall be given to all the traits while performing taxonomic scoring for classification purpose. In addition, selection of appropriate reference standards during FCM analysis is desirable. Nonetheless, our findings will provide a foundation for initiating the successful genetic improvement programmes of wild and cultivated banana through molecular breeding program in future.

Declarations

Acknowledgements

The authors are indebted to The Director, CSIR-National Botanical Research Institute, Lucknow for providing the facilities to conduct the research work. We acknowledge Dr. Himanshu Tillu, application specialist, Sysmex, India, for his valuable suggestions during the FCM analysis.

Funding

Funding from the Department of Biotechnology (DBT) under the project number GAP 3435 is duly acknowledged. [Ref file No. DBT-NER/AGRI/33/2016(Group-1)]

Institutional Review Committee

The manuscript number allotted by the ethical committee is **CSIR-NBRI_MS/2023/07/04**

Conflict of interest

The authors declare no conflict of interest

Author contributions

SNJ conceptualized and supervised the work. PT and RPB designed and carried out all the experiments and analysed the data. RPB prepared the first draft of the manuscript. PT and SNJ edited the draft MS. ABD and RKS has finalized the edited version the MS for the final form. All the authors have read the final version of the manuscript.

Data availability

Data generated during the study can be accessed on request to the corresponding author.

References

1. Asif MJ, Mak C, Othman RY (2001) Characterization of indigenous *Musa* species based on flow cytometric analysis of ploidy and nuclear DNA content. *Caryologia* 54:161–168. <https://doi.org/10.1080/00087114.2001.10589223>
2. Atom AD, Lalrinfela P, Thangjam R (2015) Genome classification of banana genetic resources of Manipur using morphological characters. *Sci Vis* 15:189–195.
3. Bennett MD, Leitch IJ (2011) Nuclear DNA amounts in angiosperms: Targets, trends and tomorrow. *Ann Bot* 107:467–590. <https://doi.org/10.1093/aob/mcq258>
4. Bennetzen JL, Kellogg EA (1997) Do Plants Have a One-Way Ticket to Genomic Obesity? *Plant Cell* 9:1509. <https://doi.org/10.2307/3870439>
5. Cavallini A, Natali L (1991) Intraspecific variation of nuclear DNA content in plant species. *Caryologia* 44:93–107. <https://doi.org/10.1080/00087114.1991.10797023>
6. Cavallini A, Natali L, Cionini G, Gennai D (1993) Nuclear dna variability within *pisum sativum* (Leguminosae): Nucleotypic effects on plant growth. *Heredity* (Edinb) 70:561–565. <https://doi.org/10.1038/hdy.1993.82>
7. Deepthi V (2016) Taxonomic scoring and genomic grouping in bananas. *Flora and Fauna* 22:151–158
8. Dolezel J (1997) Application of flow cytometry for the study of plant genomes. *J Appl Genet* 38: 285–302
9. Doležel J, Bartoš J (2005) Plant DNA flow cytometry and estimation of nuclear genome size. *Ann Bot* 95:99–110. <https://doi.org/10.1093/aob/mci005>
10. Doležel J, Doleželová M, Novák FJ (1994) Flow cytometric estimation of nuclear DNA amount in diploid bananas (*Musa acuminata* and *M. balbisiana*). *Biol Plant* 36:351–357. <https://doi.org/10.1007/BF02920930>
11. Doležel J, Greilhuber J, Suda J (2007) Estimation of nuclear DNA content in plants using flow cytometry. *Nat Protoc* 2:2233–2244. <https://doi.org/10.1038/nprot.2007.310>
12. Galbraith DW (2012) Flow cytometry and fluorescence-activated cell sorting in plants: the past, present, and future. *Biomédica* 30:65. <https://doi.org/10.7705/biomedica.v30i0.824>
13. Gregory TR (2005) Synergy between sequence and size in large-scale genomics. *Nat Rev Genet* 6:699–708. <https://doi.org/10.1038/nrg1674>

14. Hamill SD, Smith MK, Dodd WA (1992) In vitro induction of banana autotetraploids by colchicine treatment of micropropagated diploids. *Aust J Bot* 40:887–896. <https://doi.org/10.1071/BT9920887>
15. Hapsari L, Wahyudi D, Azrianingsih R, Arumingtyas EL (2015) Genome identification of bananas (*Musa* L.) from East Java Indonesia assessed with PCR-RFLP of the internal transcribed spacers nuclear ribosomal DNA. *Int J Biosci* 7:42–52. <https://doi.org/10.12692/ijb/7.3.42-52>
16. Hareesh VS, Joe A, Alappatt JP, Sabu M (2017) Musaceae of Andaman and Nicobar Islands with two new synonyms and one distributional record. *Rheedea* 27:71–78. <https://doi.org/10.22244/rheedea.2017.27.2.12>
17. Kamaté K, Brown S, Durand P, et al (2001) Nuclear DNA content and base composition in 28 taxa of *Musa*. *Genome* 44:622–627. <https://doi.org/10.1139/gen-44-4-622>
18. Konstantinidis KT, Tiedje JM (2004) Trends between gene content and genome size in prokaryotic species with larger genomes. *Proc Natl Acad Sci U S A* 101:3160–3165. <https://doi.org/10.1073/pnas.0308653100>
19. Le Thierry D'Ennequin M, Panaud O, Brown S, et al (1998) First evaluation of nuclear DNA content in *Setaria* genus by flow cytometry. *J Hered* 89:556–559. <https://doi.org/10.1093/jhered/89.6.556>
20. Lee SI, Kim NS (2014) Transposable elements and genome size variations in plants. *Genomics Inform* 12:87–97. <https://doi.org/10.5808/gi.2014.12.3.87>
21. Lynch M, Conery JS (2003) The origins of genome complexity. *Science* 302:1401–1404. <https://doi.org/10.1126/science.1089370>
22. Lysák MA, Doleželová M, Horry JP, et al (1999) Flow cytometric analysis of nuclear DNA content in *Musa*. *Theor Appl Genet* 98:1344–1350. <https://doi.org/10.1007/s001220051201>
23. Mabonga L, Pillay M (2017) SCAR marker for the A genome of bananas (*Musa* spp. L.) supports lack of differentiation between the A and B genomes. *J Agric Sci* 9:64. <https://doi.org/10.5539/jas.v9n6p64>
24. Martel E, De Nay D, Siljak-Yakovlev S, et al (1997) Genome size variation and basic chromosome number in pearl millet and fourteen related *Pennisetum* species. *J Hered* 88:139–143. <https://doi.org/10.1093/oxfordjournals.jhered.a023072>
25. Nath B, Saha PR, Das AB, et al (2021) Morpho-taxonomical and karyological study of Gopi banana: an indigenous palatable cultivated variety of *Musa* × *paradisiaca*. *Vegetos*. <https://doi.org/10.1007/s42535-021-00267-1>
26. Pillay M, Ogundiwin E, Tenkouano A, Dolezel J (2006) Ploidy and genome composition of *Musa* germplasm at the International Institute of Tropical Agriculture (IITA). *African J Biotechnol* 5:1224–1232. <https://doi.org/10.4314/ajb.v5i13.43086>
27. Praça-Fontes MM, Carvalho CR, Clarindo WR, Cruz CD (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. <https://doi.org/10.1007/s00299-011-1026-x>
28. Prasad K, Joe A, Bheemalingappa M, Prasad BR (2013) *Musa sabuana* (Musaceae): A new species from Andaman and Nicobar islands, India. *Indian J For* 36:151–153
29. Simmonds NW, Shepherd K (1955) The taxonomy and origins of the cultivated bananas. *J Linn Soc London, Bot* 55:302–312. <https://doi.org/10.1111/j.1095-8339.1955.tb00015.x>
30. Simmonds NW, Weatherup STC (1990) Numerical taxonomy of the wild bananas (*Musa*). *New Phytol* 115:567–571. <https://doi.org/10.1111/j.1469-8137.1990.tb00485.x>
31. Singh SS, Singh WR, Sagolshemcha R, Shrivastava K (2020) Genome classification of banana cultivars found in Manipur, India using IRAP and RAPD. *Indian J Agric Res* 54:388–393. <https://doi.org/10.18805/IJARE.A-5347>
32. Singh WR, Singh SS, Shrivastava K (2014) Analysis of banana genome groups of wild and cultivated cultivars of Manipur , India using sScore card method. *Adv Appl Sci Res* 5:35–38
33. Temsch EM, Koutecký P, Urfus T, et al (2021) Reference standards for flow cytometric estimation of absolute nuclear DNA content in plants. *Cytom Part A cyto.a.24495*. <https://doi.org/10.1002/cyto.a.24495>

34. Ulrich I, Ulrich W (1991) High-resolution flow cytometry of nuclear DNA in higher plants. *Protoplasma* 165:212–215. <https://doi.org/10.1007/BF01322292>
35. Uma S, Siva SA, Saraswathi MS, et al (2006) Variation and intraspecific relationships in Indian wild *Musa balbisiana* (BB) population as evidenced by random amplified polymorphic DNA. *Genet Resour Crop Evol* 53:349–355. <https://doi.org/10.1007/s10722-004-0576-y>
36. Vandenhout H, Ortiz R, Vuylsteke D, et al (1995) Effect of ploidy on stomatal and other quantitative traits in plantain and banana hybrids. *Euphytica* 83:117–122. <https://doi.org/10.1007/BF01678038>

Figures



Figure 1

Morphological plate showing different vegetative and reproductive parts of (a) *Musa balbisiana* (NBRI-002): A. plant habit with inflorescence, B. cross-section of the petiole showing closed petiole canal, C. peduncle, D. bract of the male flower, E. male flower, F. male flower petal, G. free tepal of the male flower, H. rudimentary pistil with stamens, I. rudimentary pistil, J. stamens, K. mature fruit, L-M. longitudinal and cross-section of the fruit-bearing seeds. Scale bars in A=3.5 m, in B= 0.5 m, in C= 0.5m, in D= 22cm, in E= 7cm, in F= 5.5cm, in G= 3.0 cm, H= 6.8cm, I= 6.8cm, J= 7cm, K-L= 13cm, M= 3.4cm; (b) *Musa* sp. var. *jungli kol* (NBRI-003): A. plant habit with inflorescence, B. young fruits with male inflorescence, C-D. leaf base, E-F. leaf apex, G. male bract, H. male flower, I. male flower petal and tepal, J. rudimentary pistil with stamens, rudimentary pistil,

and stamens, K. immature fruit and its cross section showing ovules, L. longitudinal of the immature fruit. Scale bars in A=3.9 m, in B= 1 m, in C-F = 2.5m, in G= 35cm, H= 8cm, I= 7.6cm, 4.1cm, J= 8cm, 8cm, 7.1cm, K-L= length (11cm), diameter (1cm); (c) *Musa balbisiana* (NBRI-006): A. plant habit with inflorescence, B-C. adaxial and abaxial surface of the entire leaf D. cross-section of the petiole showing open petiole canal, E. male bract, F. peduncle, G. male flower, H. petal, and tepal, I. rudimentary pistil with stamens, J. rudimentary pistil, K. mature fruit, L. stamens, M-N. cross and longitudinal section of the fruit-bearing seeds. Scale bars in A=4.5 m, in B-C= 2.5 m, in D=0.5m, in E= 15cm, in F = 70cm, in G= 8.2cm, H= 7cm, 3cm, I= 8.2cm, J= 8.2cm, K=14cm, L= 7cm, M= 2.8cm, N= 7cm; (d) *Musa acuminata* var. malbhog. (NBRI-0012): A. plant habit with inflorescence, B. young fruits with male inflorescence, C. male flower petal, D-E. Cross section and entire leaf with petiole, F-G. male bract abaxial and adaxial surfaces, H. reproductive parts of the inflorescence and fruit, I. rudimentary pistil, J= tepal, K= immature fruit, L-M= longitudinal and cross-section of the fruit, N= stamens. Scale bars in A= 2.6m, in B= 0.7m, C= 6.3cm, in D-E= 1.7m, in F-G= 15cm, in I= 7.1cm, J= 1.3cm, K=11cm, L= 9cm, M= 1cm, N= 5.6cm; (e) *Musa balbisiana* (NBRI-0013): A. plant habit with inflorescence, B. young fruits with male inflorescence, C-E. adaxial and abaxial surface of the entire leaf, F-G. male bract abaxial and adaxial surfaces, H. male flower, I. tepal, J. male petals, K. rudimentary pistil with stamens, L. stamens, M. rudimentary pistil, N. immature fruit, O-P. cross and longitudinal section of the fruit. Scale bars in A= 3.4m, in B= 0.8m, C-E= 2m, in F-G= 15.5cm, in H= 6.6cm, I= 1.5cm, J= 5.6cm, K=6.7cm, L= 6cm, M= 6.7cm, N-P= 10cm, O= 1cm; (f) *Musa* sp. var. fushrey jhapari (NBRI-0020): A. plant habit with inflorescence, B. young fruits with male inflorescence, C-D. adaxial and abaxial surface of the entire leaf, E-F. male bract abaxial and adaxial surfaces, G. male flower, H. male petal, and tepal, I. rudimentary pistil with stamens, J. stamens, K. rudimentary pistil, L. immature fruit, and its longitudinal and cross-sectional view. Scale bars in A= 4.5m, in B= 1m, C-D= 2.3m, in E-F= 17cm, in G= 7.2cm, H= 5.8cm, 2.8cm, I= 7.3cm, J=5.2cm, K= 7.2cm, L= length (12.7cm), diameter (1.4cm); (g) *Musa acuminata* var. chini champa (NBRI-0026): A. plant habit with inflorescence, B. cross-section of the petiole showing open petiole canal, C. peduncle, D. male bract abaxial and adaxial surfaces, E. male flower, F. male petal, G. male tepal, H. stamens, I. rudimentary pistil, J. rudimentary pistil with stamens, K. immature fruit, L. longitudinal section of the fruit, L. cross-sectional view of the fruit. Scale bars in A= 3.5m, in B= 40m, C= 1.6m, D= 15cm, E=6.4cm, F= 5.4cm, G= 2.3 cm, H= 5.3cm, I= 6.2cm, J=6.2cm, K= 13cm, L= length (13cm), diameter (3cm); (h) *Musa balbisiana* var. jungli kol (NBRI-0027): A. plant habit with inflorescence, B-C. adaxial and abaxial surface of the entire leaf, D. young fruits with male inflorescence, E-F. male bract with male flowers, G-H. male flower, I. male petal, and tepal, J. rudimentary pistil with stamens, K. stamens, L. rudimentary pistil, M-N. immature and longitudinal section of the fruit, O. peduncle. Scale bars in A= 2.2m, in B-C= 1.7m, D= 0.6m, E-F=11.2cm, G-H= 4.7cm, I= 4cm, 1.3cm, J=4.7cm, K= 4cm, L= 4.7cm, M-N= 7.4cm O= 30cm; (i) *Musa balbisiana* var. jungli kol (NBRI-0028): A. plant habit with inflorescence, B. young fruits with male inflorescence, C. immature fruit, D-E. longitudinal section of the fruit, F. male tepal, G. leaf apex, H. leaf base, I. reproductive floral parts, J. adaxial surface of the leaf, K-L. male bract, M. male flower and petal, N. close view of young fruits with male inflorescence, O. rudimentary pistil with stamens, P. rudimentary pistil and stamens. Scale bars in A= 2.2m, in B= 0.7m, C= 10cm, D-E= 10cm, F=1.1cm, G-H, J= 2m, K-L= 11.2cm, M= 4.6cm, 3.9cm, N= 0.7cm, O= 4.6cm, P=4.6cm (pistil), 3.8cm (stamens); (j) *Musa acuminata* var. kachkol (NBRI-0033): A. plant habit with inflorescence, B-C. adaxial and abaxial surface of the entire leaf, D-E. leaf petiole showing closed petiole canal, F. peduncle, G. male inflorescence, H. male flower, I. male petal, and tepal, J. rudimentary pistil with stamens, K. immature fruit, L. male bract, M-N. longitudinal and cross-section of the fruit, O= rudimentary pistil and stamens, Scale bars in A= 4.5m, in B-C= 3m, D-E= 40cm, F=50cm, G= 1m, H= 8cm, I= 7cm, 3.6cm, J=8cm, K= 15cm, L= 18cm, M-N= 18cm, 3cm, O= 7.4cm (pistil), 6.3cm (stamens); (k) *Musa* sp. var. Jungli kol (NBRI-0035): A. plant habit with inflorescence, B. leaf petiole, C. male bract, D. peduncle, E. immature fruit, F-G. longitudinal and cross-section of the fruit, H. male flower, I. male petal, J. tepal, K. stamens, L. rudimentary pistil with stamens and rudimentary pistil. Scale bars in A= 4m, in B=50cm, C=11.2 cm, D=50cm, E=12cm, F= 12cm, G= 2.8cm, H= 6.7cm, I= 4.7cm, J= 2cm, K= 3.4cm, L= 6.7cm; (l) *Musa acuminata* var. gopi (NBRI-0036): A. plant habit with inflorescence, B-C. adaxial and abaxial surface of the entire leaf, D. leaf petiole showing open petiole canal, E. peduncle, F. male inflorescence, G. male flower, H. tepal, I. rudimentary pistil with stamens, J. stamens, K. rudimentary pistil, L. male petal, M. mature fruit, N-O. longitudinal and cross-section of the fruit. Scale bars in A= 4.5m, in B-C= 3m, D=50cm, E= 70cm, F=19cm, G= 7.9cm, H= 2.6cm, I=7.9 cm, J=6.7 cm, K=7.9cm, L= 6.9cm, M-N= 13cm, O= 2.7cm; (m) *Musa paradisiaca* var. kachkol (NBRI-0037): A. plant habit with inflorescence, B. leaf petiole showing

closed petiole canal, C. peduncle, D. male bract, E. male flower, F. petal, G. tepals, H. rudimentary pistil with stamens, I. stamens, J. rudimentary pistil, K. mature fruit, L-M. longitudinal and cross-section of the fruit. Scale bars in A= 3.7m, in B=60cm, C=65cm, D=21cm, E= 8.5cm, F=7.2cm, G= 2.7cm, H= 8.3cm, I=7cm, J=8.3cm, K-L=18cm, M=4cm; (n) *Musa* sp. var. *jungli kol* (NBRI-0038): A. plant habit with inflorescence, B-C. adaxial and abaxial surface of the entire leaf, D-E. leaf petiole showing closed petiole canal, F. peduncle, G-H. male bract adaxial and abaxial surfaces, I. hand with 12 fingers, J. male flowers in the aggregate, K. a single male flower, L. rudimentary pistil with stamens, M. rudimentary pistil, N. stamens, O-P. mature and longitudinal section of the fruit, Q. cross-section of the fruit, R. male petal, S. tepal. Scale bars in A= 4.5m, in B-C= 3m, D-E= 60cm, F= 40cm, G-H= 22cm, I= 15cm, J-K= 8cm, L=7.8 cm, M=7.8cm, N=5.6cm, O-P= 15cm, Q= 2.8cm, R= 6cm, S=3.1cm; (o) *Musa acuminata* var. *sobri* (NBRI-0039): A. plant habit with inflorescence, B-C. adaxial and abaxial surface of the entire leaf, D. leaf petiole showing closed petiole canal, E. male bract adaxial and abaxial surfaces, F. peduncle, G. male flowers, H. male petal, I. tepal, J. rudimentary pistil with stamens, K. rudimentary pistil, L. stamens, M. immature fruit, N. longitudinal section of the fruit, O. cross-section of the fruit. Scale bars in A= 2.2m, in B-C=1.6m, D=40cm, E=18cm, F= 30cm, G=7.2cm, H= 6cm, I= 2.8cm, J= 7.2cm, K=7.2cm, L=5.9cm, M-N=11cm, O= 1.8cm and (p) *Musa acuminata* var. *champa* (NBRI-0041): A. plant habit with inflorescence, B. young fruits with male inflorescence, C-D. section of the leaf, E. close view of the inflorescence, F. male flower, G. petals and tepals, H. rudimentary pistil with stamens, I. male bract, J. rudimentary pistil and stamens, K. immature fruit, L. longitudinal and cross-section of the fruit. Scale bars in A= 3m, in B= 1.2m, C-D= 2cm, E= 1m, F=7.7cm, G= 7cm, 2.1cm, H=7.4cm, I=20cm, J= 7.4cm (pistil), 5.4cm (stamens), K=12cm, L= 12cm (length), 2.5cm (diameter).

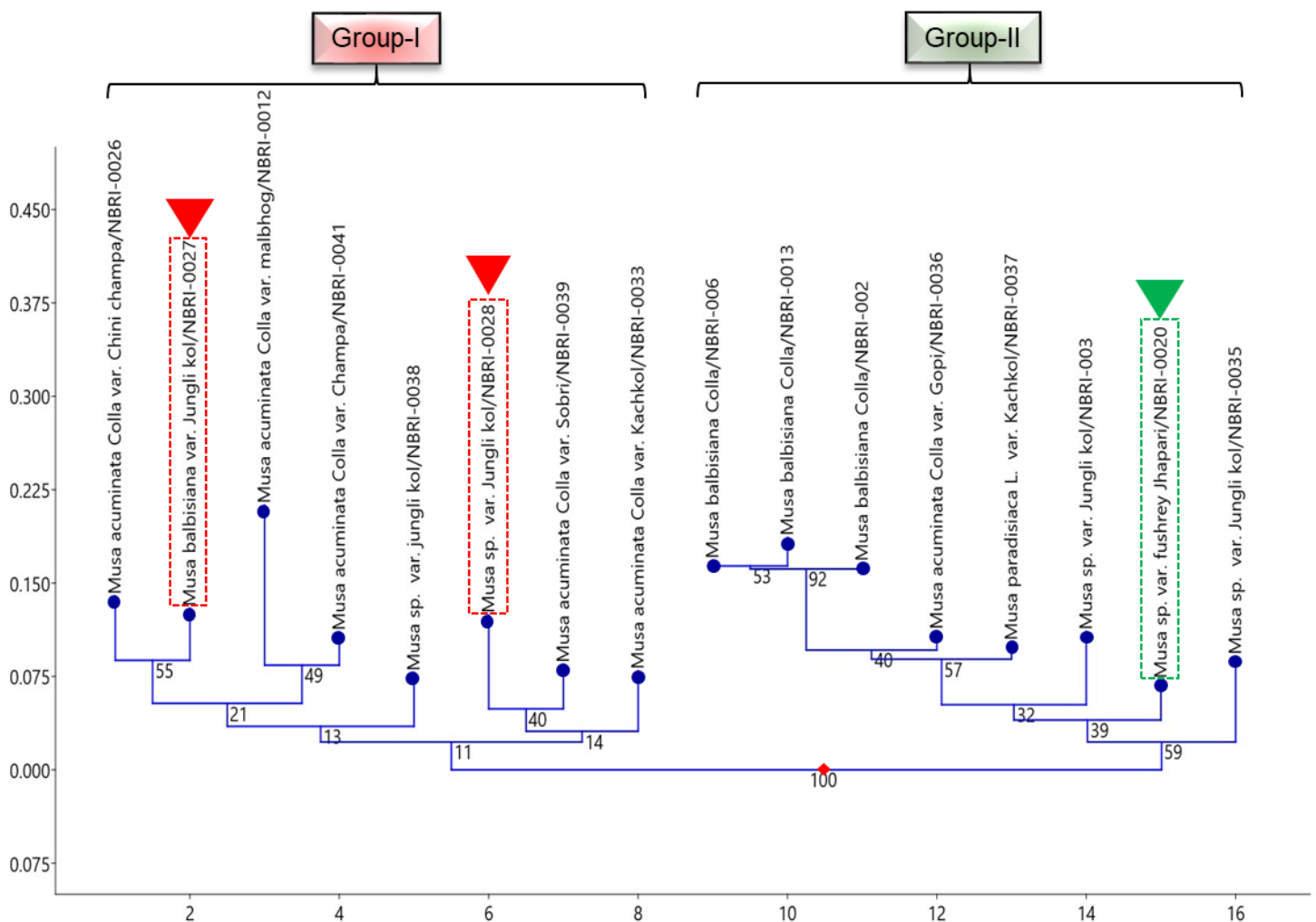


Figure 2

Dendrogram analysis showing the clustering of 16 *Musa* germplasms based on their morpho-taxonomical scoring

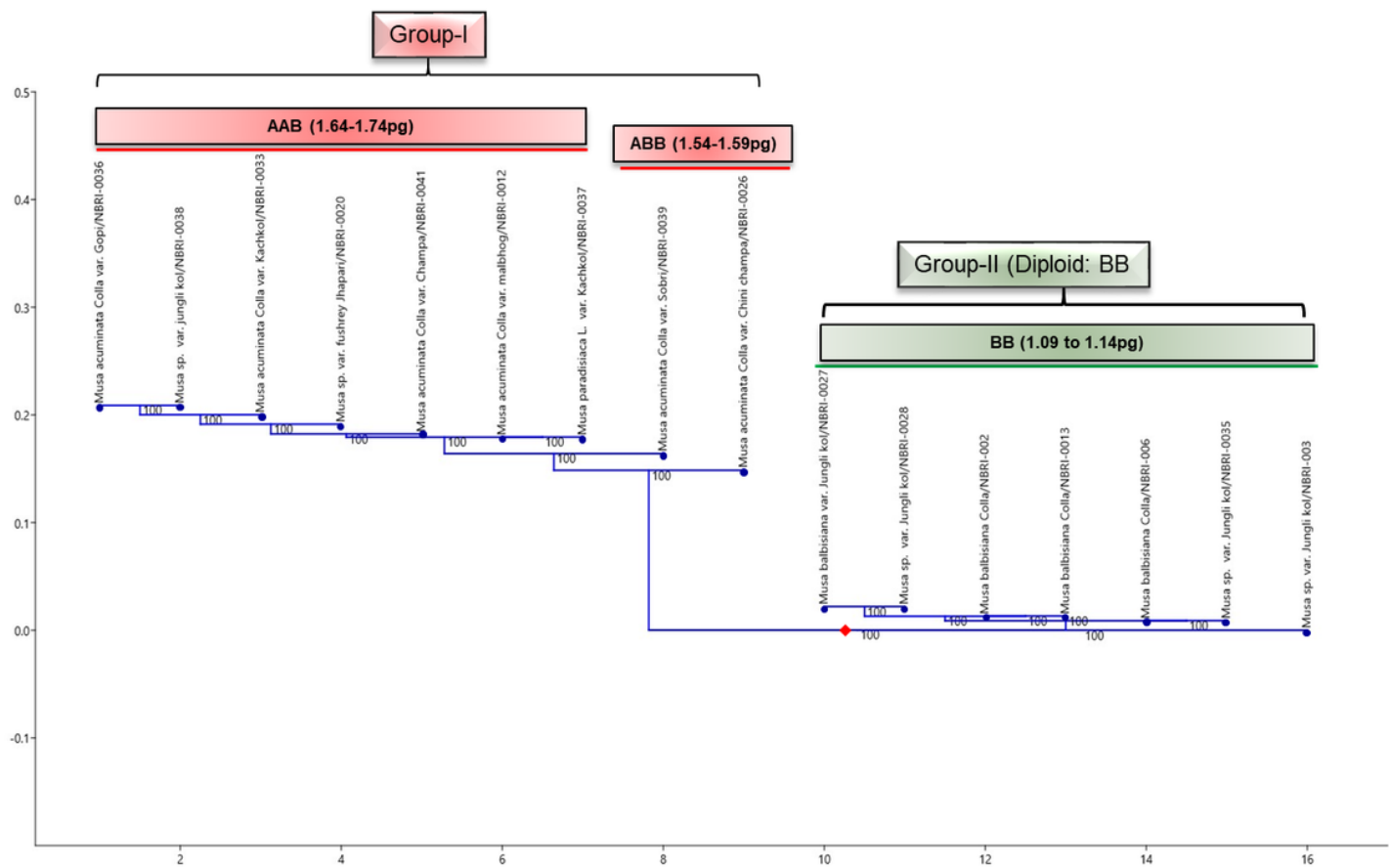


Figure 3

Dendrogram analysis showing the clustering of 16 *Musa* germplasms based on their 2C DNA content and genome constituent estimated using flow cytometry

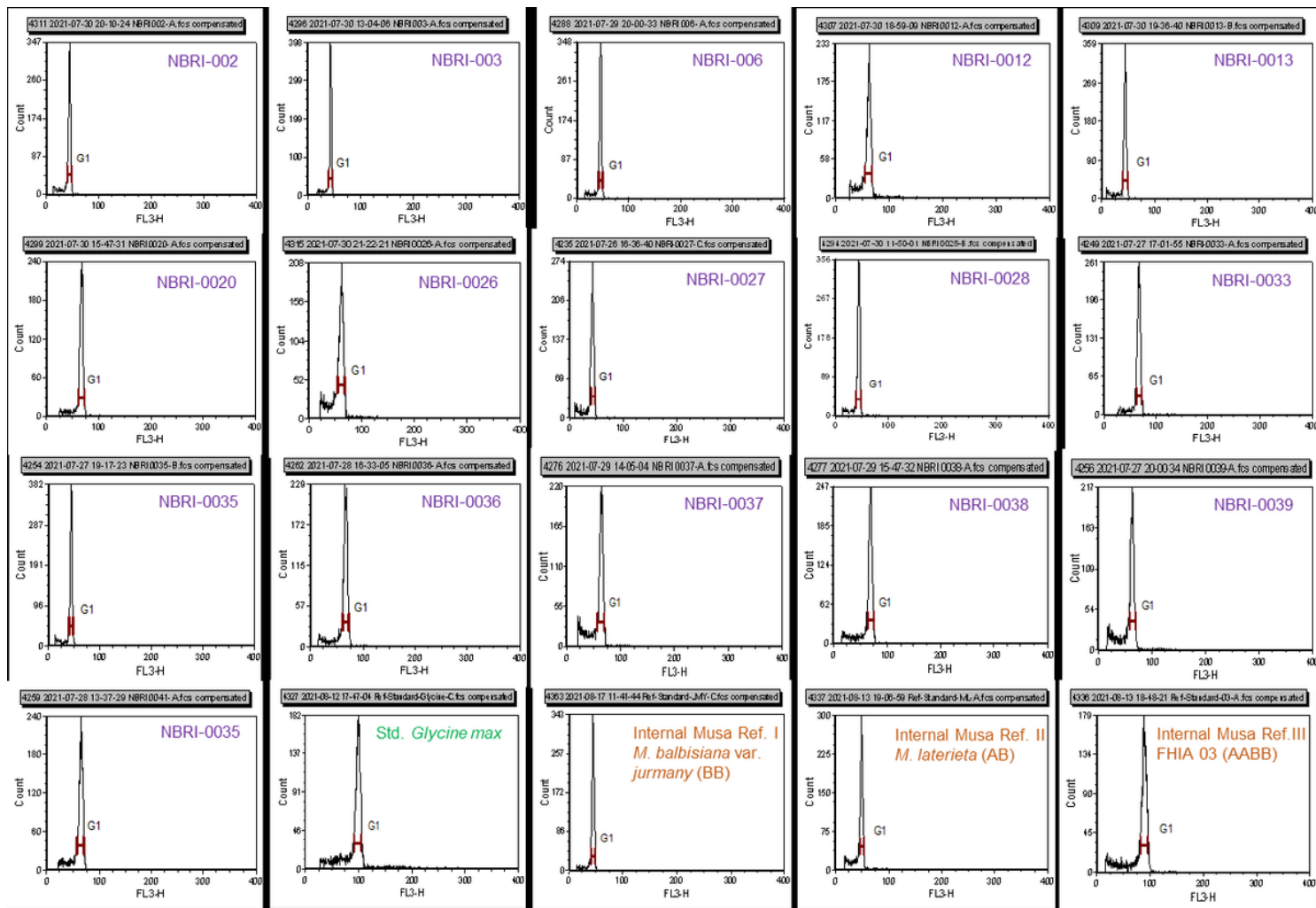


Figure 4

Histograms plots showing the G0/G1 peak position of 16 *Musa* germplasms along with the reference standard of *Solanum lycopersicum* L.'Stupicke' polni rane and *Glycine max* Merr. Polanka for estimating 2C DNA content

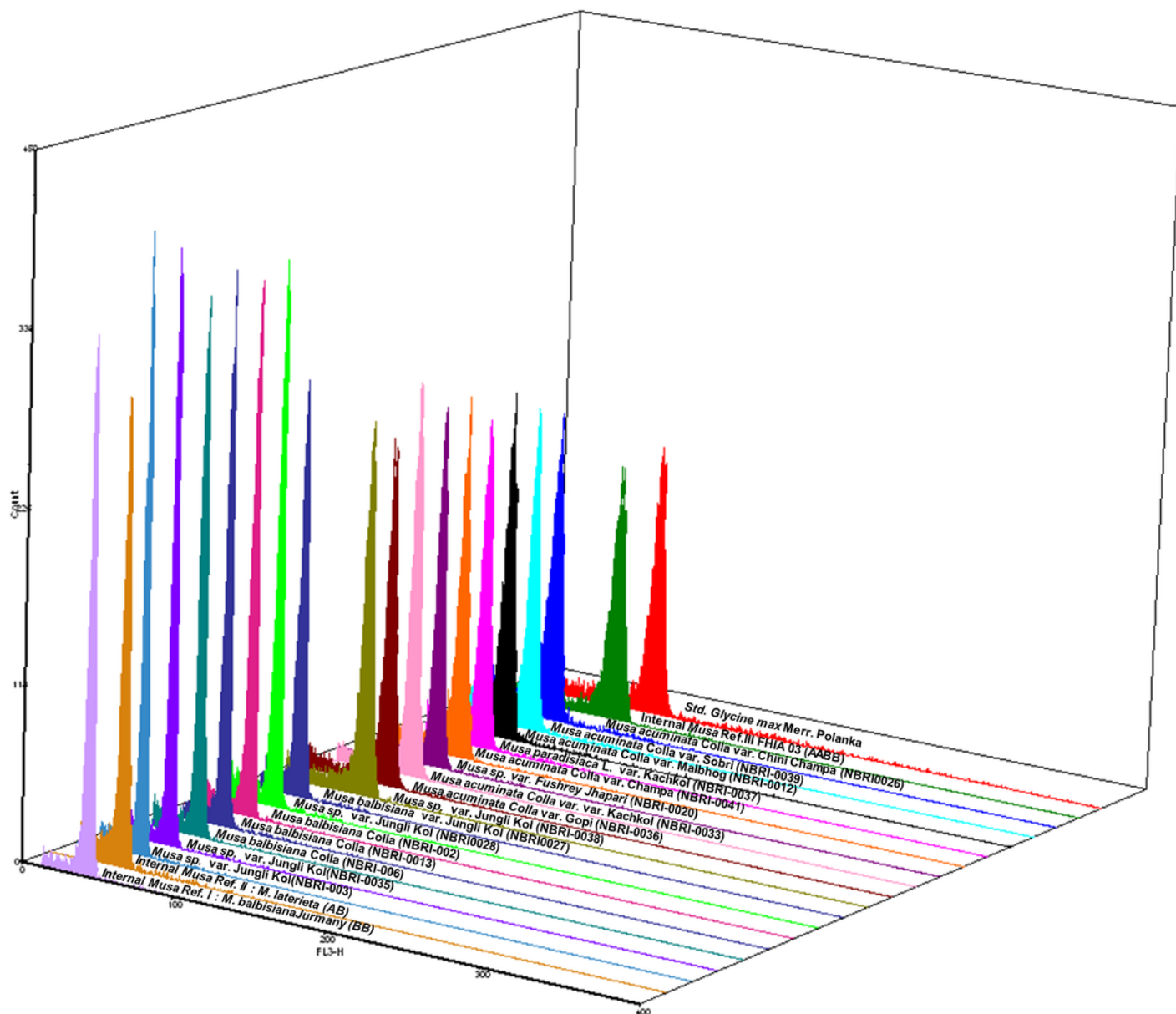


Figure 5

A three dimensional plot representing the relative G0/G1 peak position of 16 *Musa* germplasms using *Solanum lycopersicum* L.'Stupicke' polni rane and *Glycine max* Merr. Polanka as external reference standards

