Journal of Tropical Forest Science 35(1): 93–107 (2023)
https://doi.org/10.26525/jtfs2023.35.1.93
ISSN: 0128-1283, eISSN: 2521-9847

IDENTIFICATION OF GENETIC RELATIONSHIPS BETWEEN INDONESIAN BANANA CULTIVARS AND THE WILD RELATIVES USING DNA BARCODES

Ardiyani M[1,] *, Ermawar RA[2], Yulita KS[3], Dewi CLH[2], Ahmad F[4], Sulistyaningsih LD[1], Sari FP[2] & Fatriasari W[2]

[1]Research Center for Biosystematics and Evolution, National Research and Innovation Agency (BRIN), Jl. Raya Jakarta-Bogor, Cibinong, Bogor 16911, Indonesia
[2]Research Center for Biomass and Bioproducts, National Research and Innovation Agency (BRIN), Jl. Raya Jakarta-Bogor, Cibinong, Bogor 16911, Indonesia
[3]Research Centre for Ecology and Ethnobiology, National Research and Innovation Agency (BRIN), Jl. Raya Jakarta-Bogor, Cibinong, Bogor 16911, Indonesia
[4]Research Center for Genetic Engineering, National Research and Innovation Agency (BRIN), Jl. Raya Jakarta-Bogor, Cibinong, Bogor 16911, Indonesia

marlina.ardiyani@brin.go.id

Submitted April 2022; accepted August 2022

There are several studies on the DNA barcoding of banana worldwide, however studies on the DNA barcoding of Indonesian bananas are limited. This study aims to develop a DNA reference library of the Indonesian banana species, varieties and cultivars to reconstruct phylogenetic trees, and to identify and understand genetic relationships between varieties and cultivars using three barcoding markers (*rbc*L, *mat*K and *trn*L-F). A total of 26 accessions of banana were amplified and sequenced using these primers. Maximum likelihood (ML) and maximum parsimony (MP) were conducted on these sequences with addition of 9 accessions (*rbc*L), 13 accessions (*mat*K) and 14 accessions (*trn*L-F) from the National Center for Biotechnology Information (NCBI) GenBank. The aligned sequences were measured for genetic distances. The results showed that all markers were significant to differentiate among the A and B genomes. All markers resulted in a tree that contained two distinctive clades supported by high bootstrap value (> 50%). The *rbc*L marker was the most conserved, followed by *mat*K and *trn*L-F. The *rbc*L marker can only be used to distinguish generic level, while *mat*K and *trn*L-F can be used at species level. All markers cannot be used to identify subspecific, variety and cultivar levels. However, the markers are able to suggest the genome of the maternal parent.

Keywords: DNA barcoding, *mat*K, *Musa, Musaceae*, *rbc*L, *trn*L-F

INTRODUCTION

The bananas and plantains of the genus *Musa* are native to the paleotropics of Asia, Africa and Australia (Kress 1990, Kress et al. 2005, Kress & Specht 2006, De Langhe et al. 2009, Li et al. 2010). The *Musa* consists of 45 wild species and almost 2,000 cultivars among which 12 and more than 60 are respectively found in Indonesia (Nasution & Yamada 2001, Valmayor et al. 2000, Ruas et al. 2017). There are 8 wild subspecies and 7 varieties of *Musa acuminata* in Java (Nasution & Yamada 2001, Sulistyaningsih et al. 2014, Sulistyaningsih 2016). The edible bananas' diversity originated from inter- and intra-specific hybridisation of the wild relatives, namely *M. acuminata* (A genome) and *M. balbisiana* (B genome), grown in Southeast Asia (Cheesman 1948, Simmonds & Shepherd 1955, Simmonds 1962). Furthermore, the cultivars are diploid (AA and AB), triploid (AAA, AAB and ABB) or tetraploid (AAAB and AABB) (Simmonds 1962, Simmonds & Shepherd 1955).

Taxonomy of the wild *Musa* is complex (Häkkinen & Väre 2008). The traditional classification of the genus *Musa* has been questioned based on DNA analyses. The genus *Musa* consists of two sections, Callimusa and Musa. Callimusa section is a combination of three former sections, i.e. Australimusa, Callimusa

and Ingentimusa. Whereas, Musa section is formerly Eumusa and Rhodochlamys sections (Häkkinen 2013). Elucidating the taxonomy and phylogeny of wild *Musa* is important as it can provide valuable information for the collection and characterisation for further improvement. Although being a part of the center of banana origin and diversity, there are limited number of systematic studies that have been conducted to reveal the taxonomy and phylogeny of wild *Musa* in Indonesia. A taxonomic study of *M. acuminata* and its intraspecific taxa in Indonesia was conducted by Nasution (1991) and later revised by Ahmad et al. (2021) for wild *M. acuminata* in Sumatra. Large variation, numerous names and synonyms in banana and plantain cultivars led to difficulties in cultivar identification (Valmayor et al. 2000). Distinguishing cultivars based on morphological characters could be difficult because there is lack of information provided by morphological characters, particularly in determining morphologically overlapping cultivars (Sefc et al. 2001).

DNA barcoding is still employed in species identification besides using full genome information technology. This method involves utilising short molecular markers and needs to be conservative to simplify amplification and alignment among distant species (Hebert et al. 2003, Kress et al. 2005). The ITS2, *rbc*L and *mat*K are effective markers to barcode plants at species and generic taxonomic levels (Kress et al. 2005, Kress & Erickson 2007). Molecular studies on *Musa* species and its cultivars have been conducted using chloroplast markers of *atp*B-*rbc*L, *rps*16 and *trn*L-F, and nuclear markers of the Internal Transcribed Spacers to construct the phylogeny in family and infrageneric level (Li et al. 2010). Moreover, nuclear genes, ADH, G3PDH, IDH, CAT, GBSS and ARF are used to determine the origin of edible bananas (Volkaert 2009). The markers, Waxy and ADH1, and chloroplast markers of *mat*K and *trn*L-F cluster are employed for species phylogenetic study and estimation of cultivars' evolution (Li et al. 2013). The ITS2 is also used for phylogenetic study and barcoding of cultivars (Dhivya et al. 2020).

The Indonesian cultivars' DNA barcode has been studied in Bali bananas using ITS2, as well as in East Java bananas using *rbc*L and ITS, but only at a limited level considering the species and cultivars' high diversity (Hapsari et al. 2018, Ainiyah et al. 2020, Dwivany et al. 2020). Therefore, it is important to develop a DNA reference for identification purposes and understanding the evolutionary relationship between the species and cultivars. This study aims to develop a DNA reference library of the Indonesian banana species, varieties and cultivars, to reconstruct phylogenetic trees, and to identify and understand genetic relationships between varieties and cultivars using three barcoding markers (*rbc*L, *mat*K and *trn*L-F).

MATERIALS AND METHODS

Materials

In total, 26 accessions of species, subspecies, varieties and cultivars were collected from the banana germplasm information system of the Research Center for Biology and Purwodadi Botanic Garden, Indonesian Institute of Sciences (Table 2). For the ingroup, a total of 26 samples which were obtained by direct sequencing were used with an addition of 5 sequences of *rbc*L, 9 sequences of *mat*K and 10 sequences of *trn*L-F from the GenBank. The outgroups were all taken from the GenBank, i.e. 4 sequences of *rbc*L, 4 sequences of *mat*K and 4 sequences for *trn*L-F.

The leaf samples preservation using teabag method was in accordance to Wilkie *et al.* (2013). Healthy leaves were torn into small pieces and put into a teabag, which was then stored in an airtight container and submerged in silica gel. After being dried, the tissue became ready for DNA extraction.

DNA isolation

Total genomic DNAs were extracted using DNeasy Plant Mini Kit (QIAGEN) according to protocol from the manufacturer. Before this process, the tissue was homogenised into fine powder using a lyseror mixer mill (400 mm).

DNA amplification

The *rbc*L, *mat*K and *trn*L-F regions were amplified using a PCR kit in a thermo cycler with the following temperature profiles: initial denaturation for 3 min at 95 °C, then 35–40 cycles of denaturation for 1 min at 94 °C. Afterward,

there was annealing at 52-55 °C for *rbc*L, *mat*K and *trn*L-F primers for 30 s each, 1 min extension at 72 °C, and a final 5 min extension at 72 °C. The PCR cocktail of a 12.5 µL volume contains 2.5 µL of 5 x buffer Go taqPromega, 0.07 µL taq DNA polymerase, 0.25 µL dNTPs, 1 µL $MgCl_2$, 0.25 µL of 10 mmol each of forward and reverse primers, 7.18 µL of nuclease-free water, and 1 uL DNA template. The primer pairs are listed in Table 1.

Sequencing

The PCR products were purified and sequenced at 1st BASE using ABI PRISM 3730 × l genetic analyser.

Sequence alignment

Forward and reverse DNA sequences were made into contig and checked using the Geneious program and manual checking, based on chromatograms. The consensus sequences were exported to FastA format for subsequent analysis. Their alignment was performed using Geneious and MEGA X, then checked manually with eyes. Next, the aligned sequences were exported to NeXus format for phylogenetic analysis using PAUP ver.4.

Phylogenetic analysis and basic local alignment search tool (BLAST)

Phylogenetic analyses were performed using PAUP ver. 4 (Swofford 1998), MEGA X and Geneious with maximum parsimony (MP) and maximum likelihood (ML). Two closely related species of *Musa* (*Ensete ventricosum*, *Ensete glaucum* and *Musella lasiocarpa*) were selected as outgroups. All characters were considered to have equal weight in the parsimony optimality criteria. They were unordered and detected gaps were treated as missing data. Trees were generated using the MP optimality criteria. Heuristic searches were conducted through random sequence addition, with 500 replications, using various branch swapping and length optimisation options including one tree being held at each step. From the most parsimonious trees generated, a strict consensus tree was constructed. The consistency and homoplasy indices were calculated for all trees using distance (weighted least squares with power +2) as one of the criteria. Also, clade support was estimated using bootstrap values which were considered as having strong support when more than 50%. The BLAST was done in the NCBI GenBank to verify the closest entity to our DNA sequences.

RESULTS

Phylogenetics of *Musa* spp. based on *rbc*L gene

A total of 35 taxa representing 21 cultivars and 14 wild species (including two subspecies and five varieties) were used to analyse the phylogenetic relationship of *Musa* cultivars with their wild species by using the *rbc*L gene as a marker, based on ML and MP algorithms. A total of 26 samples used as ingroup were sequenced directly, while 5 sequences of *rbc*L, were taken from the GenBank. Meanwhile, 4 sequences including three sequences of *Ensete glaucum* and one of *Musella lasiocarpa*, retrieved from GenBank, were treated as outgroups.

The result of the phylogenetic tree constructed based on ML is similar to the one based on MP algorithm (Figures 1 and 2). The outgroup *Ensete* has a personal clade supported by a high bootstrap (BS) value (95%), and *Musella* is the

Table 1 List of primers for *rbc*L, *mat*K and *trn*L-F markers

Primer	Primer motifs	Size (bp)	References
*rbc*L 1F	5'- ATG TCA CCA CAA ACA GAA AC -3'	20	Kress et al. (2005)
*rbc*L 724R	5'- TCG CAT GTA CCT GCA GTA GC -3'	20	Kress et al. (2005)
*mat*K-1RKIM-f	5'- ACC CAG TCC ATC TGG AAA TCT TGG TTC -3'	25	Ki-Joong Kim (unpublished)
*mat*K-3FKIM-r	5'- CGT ACA GTA CTT TTG TGT TTA CGA G -3'	27	Ki-Joong Kim (unpublished)
*trn*L-c	5'- CGA AAT CGG TAG ACG CTA CG -3'	20	Taberlet et al. (1991)
*trn*L-f	5'- ATT TGA ACT GGT GAC ACG AG-3'	20	Taberlet et al. (1991)

sister to this clade with *Musa borneensis*. Two well-supported clades were found i.e. Clade I (BS 61% based on ML, BS 66% based on MP) which contains *Musa balbisiana,* its cultivars and hybrids such as "Pisang Klutuk", "Pisang Klutuk Wulung", "Pisang Raja", "Pisang Siem" and "Pisang Kepok". Clade I contains species and cultivars of the Genome B, while Clade II (BS 99% based on ML and MP), consists of Genome A species and its cultivars, and *M. acuminata,* its cultivars and hybrids such as "Pisang Ambon Kuning", "Cavendish", "Pisang Tanduk", "Raja Sereh" and "Pisang Rejang".

The MP algorithm of *rbc*L dataset produced one shortest tree (Figure 2) with 622 constant characters, while one variable character was parsimony-uninformative and 12 were parsimony-informative. Consistency index (CI) = 0.867 is a number that shows whether the tree is consistent in terms of quality with

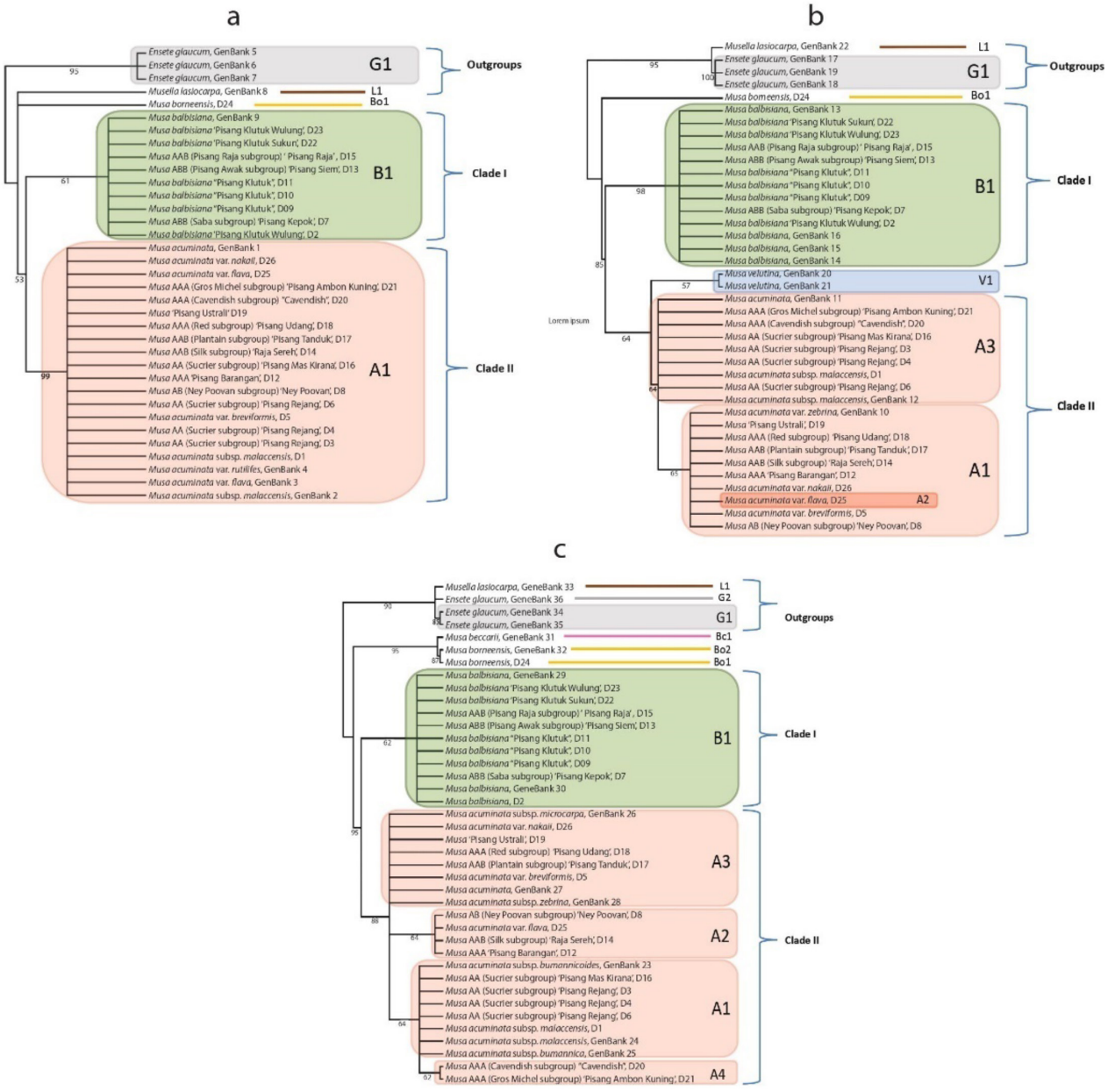

Figure 1 Phylogenetic tree constructed based on maximum likelihood analysis tree using rbcL (a), *mat*K (b), and *trn*L-F (c); according to the sequences of accessions, accessions with identical sequences are clustered in one group; letter in the right side of box represents clusters and chlorotypes names; bootstrap support values of 500 replicates (%) are shown at the nodes

the maximum value, one, and retention index (RI) = 0.977 explains the tree's quality by using different statistics from CI (Kitching et al. 1998). The value of CI, as well as RI, shows the high consistency and resolution of the phylogenetic tree (Swofford 1998). Based on CI and RI values produced by MP, the strict consensus tree had high quality and resolution, which was quite consistent. Meanwhile, homoplasy index (HI) = 0.133 explains the level of homoplasy of a character. This was produced from one substracted by CI value (Kitching et al. 1998). The low value of HI showed that the characters were used to construct a phylogenetic tree at a high level of homoplasy. However, the separation support of the ingroups and outgroups is low [BS = 53% (ML) and BS = 59% (MP)], meaning the topology trees have a high possibility of branching rearrangement (Kress et al. 2002).

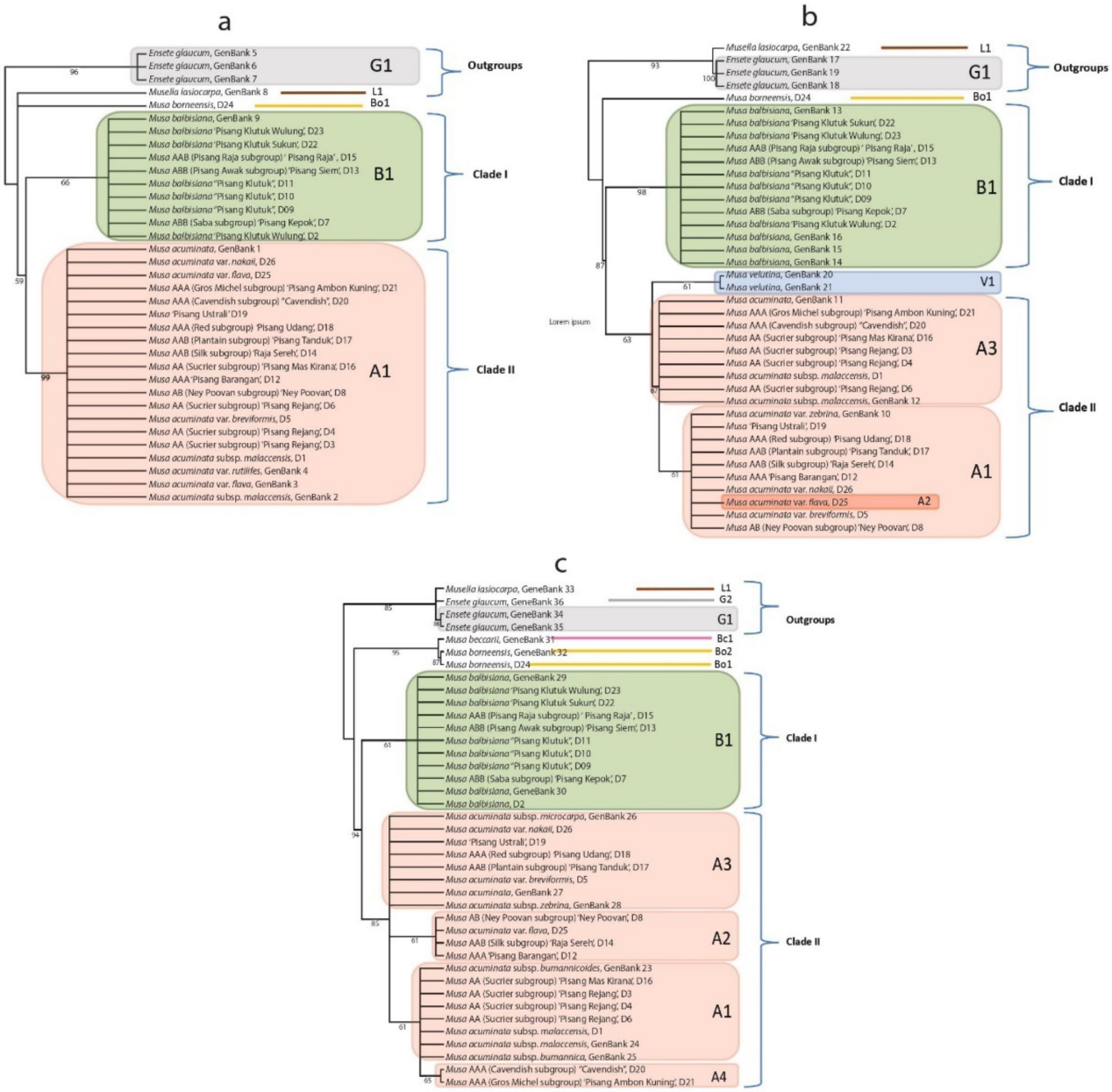

Figure 2 Phylogenetic tree constructed based on maximum parsimony (MP) analysis tree using *rbc*L (a), *mat*K (b) and *trn*L-F (c); according to the sequences of accessions, accessions with identical sequences are clustered in one group; letter in the right side of box represents clusters and chlorotypes names; bootstrap (BS) support values of 1,000 replicates (%) are shown at the nodes

Phylogenetics of *Musa* spp. based on *mat*K gene

Another analysis of the phylogenetic relationship of *Musa* cultivars with their wild species was conducted based on *mat*K gene using ML and MP algorithms. A total of 39 taxa representing 21 cultivars and 14 wild species (including two subspecies and four varieties) were used. Among 35 sequences treated as ingroups, 9 were obtained from the NCBI GenBank. The 3 sequences of *Ensete glaucum* and 1 of *Musella lasiocarpa* retrieved from GenBank were treated as outgroups. The results of phylogenetic trees, constructed based on ML and MP, were similar (Figures 1 and 2). The outgroups had their clade supported by a high BS value (95%). Moreover, *Musella* is a sister to *Ensette* separated from the ingroups.

The tree is similar to that constructed based on *rbc*L data where two well-supported clades were found i.e. Clade I (BS 98% based on ML and MP) which contained the Genome species and cultivars, as well as *Musa balbisiana,* its cultivars and hybrids. Clade II (BS 64% based on ML, 63% based on MP), consists of the Genome A species and cultivars, as well as *M. acuminata,* its cultivars and hybrids. The difference between the trees constructed based on *rbc*L and *mat*K data was in the Clade II. There were three well-supported clades, namely clade IIa containing *M. velutina*, IIb containing the AAA species plus IIc which also contains AB, AAA, AAB species, and *Musa* varieties (var. *zebrina*, var. *nakaii*, var. *flava* and var. *breviformis*).

The MP algorithm of the *mat*K dataset produced one shortest tree (Figure 2) with 706 constant characters, while 25 variable characters were parsimony-uninformative and 22 were parsimony-informative; CI = 0.940, HI = 0.060 and RI = 0.978. The value of CI, as well as RI, showed the high consistency and resolution of the phylogenetic tree. Hence, the strict consensus tree had a high quality and resolution.

Phylogenetic trees based on the *trn*L-F region

The aligned sequences of the *trn*L-F region obtained from 48 samples were 922 bps. Out of these, 889 were constant characters, while 5 were parsimony-uninformative, hence only 18 nucleotides were parsimony-informative characters.

The phylogenetic trees constructed from the ML and MP algorithms showed an almost similar pattern. The topology result from the analysis conducted based on ML showed the ingroups (Figure 1) consisted of *M. beccari, M. borneensis, M. acuminate, M. balbisiana* and a crossing between both species. The ingroup which contained *M. acuminata* and *M. balbisiana* formed a monophyletic unit with strong support from BS 95%, while their further groupings showed three groups of *M. balbisiana* (BS 62%) and *M. acuminata* (88%). The acuminata was further divided into two groups of *M. acuminata* and *M. acuminata* x *M. balbisiana.*

The phylogenetic trees constructed from MP algorithms showed the ingroups (Figure 2), also consisting of *Musa beccari, M. borneensis, M. acuminata, M. balbisiana,* and a crossing between both species. *Musa acuminata* and *M. balbisiana* formed a monophyletic unit with strong support from the bootstrap (94%). This clade was further grouped into *M. balbisiana* (BS 61%) and *M. acuminata* (85%). The acuminata clade had two clades that consist of *M. acuminata* and *M. acuminata* x *M. balbisiana.*

Clade I had polytomy topologies with low bootstrap support [BS = 62% (ML) and BS = 61% (MP)], while Clade II had strong bootstrap support [BS = 88% (ML) and BS = 85% (MP)]. Polytomy topology is the separation of branches that is incapable to distinguish or separate one species from another. It is probably caused by simultaneous evolution, hence causing uncertainty in this process (Kuhn et al. 2011).

Chlorotypes and BLAST results

All groups of identical sequences i.e. chlorotypes, at the species level according to *rbc*L, *mat*K and *trn*L-F, corresponded to the grouping in phylogenetic trees (Table 2, Figure 1). Based on these markers, *M. balbisiana* was defined in one chlorotype. Contrarily, *M. acuminata* was identifed in one chlorotype based on *rbc*L, three based on *matK* and five based on *trn*L-F (Table 2). The BLAST results showed that the ID of *Musa* accessions used in this study matched the species, subspecies and varieties deposited in the NCBI GenBank (Table 4).

DISCUSSION

The chloroplast marker (*rbc*L, *mat*K and *trn*L-F)

The topology of phylogenetic tree constructed by ML and MP algorithms based on *rbc*L, *mat*K and *trn*L-F data had two clades. Clade I consisted of BB ("Pisang Klutuk", "Pisang Klutuk Wulung" and "Pisang Klutuk Sukun") and ABB genome groups ("Pisang Raja", "Pisang Siem" and "Pisang Kepok"), as well as wild species (*Musa balbisiana*). Clade II consisted of AA ("Pisang Mas Kirana" and "Pisang Rejang"), AAA ("Pisang Ambon Kuning", "Cavendish", "Pisang Udang", "Pisang Barangan" and "Pisang Rejang"), AAAA ("Pisang Tarali"), AB ("Ney Poovan") and AAB genome groups ("Pisang Tanduk" and "Raja Sereh"). This also included wild banana species, viz. *Musa acuminata*, *M. acuminata* subsp. *malaccensis*, *M. acuminata* var. *breviformis*, *M. acuminata* var. *flava*, *M. acuminata* var. *nakaii* and *M. acuminata* var. *rutilifes*. The clades showed that the wild species, *M. balbisiana* was grouped with cultivated banana, from the B genome. Meanwhile, *M. acuminata* and its infraspecific taxa were grouped with cultivated banana from A genome. Therefore, the phylogenetic tree was concluded to be separate, based on the genome characters.

The majority of edible cultivated bananas originated from intraspecific or interspecific hybridisation between wild diploid *M. acuminata* (A-genome) and *M. balbisiana* (B-genome) species. Combinations of these A- and B-genomes have led to various genotypes of the fruit, including diploid (AA, BB and AB), triploid (AAA, AAB and ABB), and tetraploid (AAAB, AABB and ABBB) variants (Simmonds & Shepherd 1955). Results also showed that three groupings were in accordance with the genome content, acuminata (AA), balbisiana (BB) and its hybrid (AB).

The balbisiana group, even though weakly and moderately supported by the BS analysis, showed a strong tendency to form an inclusive group. Close examination on *trn*L-F marker nucleotide sequence identified a long repeat within the *trn*L intron region at positions 305–319 (Table 3). This long insertion was only recorded in 11 samples of *M. balbisiana*, but not in the hybrid, and it tends to have a relationship with the B genome. A long insertion/deletion event in the *trn*L intron has also been known to occur in Annonaceae, Fabaceae and Dipterocarpaceae (Pirie et al. 2007, Yulita 2013, D'yachenko et al. 2015). However, a recent genomic study on *M. balbisiana* (Wang et al. 2019) reported 55.75% of the B-genome assembly composed of repetitive sequences, which was higher than the 41.85% of the A-genome assembly.

Genetic diversity based on chloroplast genes

All markers used provided two main clades which were associated with A and B genomes. Chloroplast is transmitted maternally, hence, based on the present study, the cultivars or hybrid bananas' female parent genome can be estimated. Notably, the chloroplast of cultivated *M. acuminata*, AA and AAA genomes, is transmited from this species. In hybrid bananas, such as AB, AAB or ABB genome, chloroplast genes are useful to trace its donor. For instance, the chloroplast donor of "Pisang Kepok" (D7) was suggested from *M. balbisiana* indicating that this banana's maternal ancestor is from *M. balbisiana*.

Understanding the genome transmission is important for breeding programs as bananas' genetic is complex due to inter- and intraspecific hybridisation of two species *M. acuminata* and *M. balbisiana* (Heslop-Harrison & Schwarzacher 2007). The information obtained can help breeders to select potential donors for hybridisation.

CONCLUSION

This study showed that the *rbc*L gene relatively had low discrimination rates for the banana species and it lacked sufficient discriminatory power due to the species level. The large subunit gene of ribulose-1,5-bisphosphate carboxylase is part of the sequence located in the chloroplast DNA (cpDNA), a highly conserved region often used in barcoding for plants. The sequence has low mutation compared with other barcodes in cpDNA because it has a high level of similarity between different species. The *mat*K gene also had low rates but it is used to distinguish at the species level, while *trn*L-F was more variable compared to the other genes but it is only used to distinguish at the species level. Discrimination at subspecies, variety and cultivar levels was not performed with the three chloroplast genes.

Table 2 Banana species, varieties and cultivars with their living accession, chlorotypes and GenBank accession numbers

Sample code	Scientific name	Local name	Genome composition	Ploidy level	Accession no.	Collection	Chlorotype			GenBank accession no.		
							*rbc*L	*mat*K	*trn*L-F	*rbc*L	*mat*K	*trn*L-F
D1	*Musa acuminata* ssp. *malaccensis*	Pisang Hutan	AA	2x	LIPI-010	RCB-BRIN	A1	A3	A1	MZ395222	MZ318102	MZ318128
D2	*Musa balbisiana* Colla	Pisang Klutuk Wulung	BB	2x	LIPI-064	RCB-BRIN	B1	B2	B1	MZ395223	MZ318103	MZ318129
D3	*Musa* AA (Sucrier subgroup)	Pisang Rejang	AA	2x	LIPI-082	RCB-BRIN	A1	A3	A1	MZ395224	MZ318104	MZ318130
D4	*Musa acuminata*	Pisang Rejang	AA	2x	LIPI-007	RCB-BRIN	A1	A3	A1	MZ395225	MZ318105	MZ318131
D5	*Musa acuminata* var. *breviformis*	Cau Kole	AA	2x	LIPI-218	RCB-BRIN	A1	A1	A3	MZ395226	MZ318106	MZ318132
D6	*Musa acuminata*	Pisang Rejang	AAA	3x	LIPI-143	RCB-BRIN	A1	A3	A1	MZ395227	MZ318107	MZ318133
D7	*Musa* ABB (Saba subgroup)	Pisang Kepok	ABB	3x	LIPI-060	RCB-BRIN	B1	B1	B1	MZ395228	MZ318108	MZ318134
D8	*Musa* AB (Ney Poovan subgroup)	Ney Poovan	AB	2x	LIPI-035	RCB-BRIN	A1	A1	A2	MZ395229	MZ318109	MZ318135
D9	*Musa balbisiana* Colla	Pisang Klutuk	BB	2x	LIPI-054	RCB-BRIN	B1	B1	B1	MZ395230	MZ318110	MZ318136
D10	*Musa balbisiana* Colla	Pisang Klutuk Sukun	BB	2x	LIPI-062	RCB-BRIN	B1	B1	B1	MZ395231	MZ318111	MZ318137
D11	*Musa balbisiana* Colla	Pisang Klutuk Sukun	BB	3x	LIPI-061	RCB-BRIN	B1	B2	B1	MZ395232	MZ318112	MZ318138
D12	*Musa* AAA	Pisang Barangan	AAA	3x	LIPI-561	RCB-BRIN	A1	A1	A2	MZ395233	MZ318113	MZ318139
D13	*Musa* ABB (Pisang Awak subgroup)	Pisang Siem	ABB	3x	LIPI-315	RCB-BRIN	B1	B1	B1	MZ395234	MZ318114	MZ318140
D14	*Musa* AAB (Silk subgroup)	Pisang Raja Sereh	AAB	3x	LIPI-559	RCB-BRIN	A1	A1	A2	MZ395235	MZ318115	MZ318141
D15	*Musa* AAB (Pisang Raja subgroup)	Pisang Raja	AAB	3x	LIPI-360	RCB-BRIN	B1	B1	B1	MZ395236	MZ318116	MZ318142
D16	*Musa* AA (Sucrier subgroup)	Mas Kirana	AA	3x	LIPI-550	RCB-BRIN	A1	A3	A1	MZ395237	MZ318117	MZ318143
D17	*Musa* AAB (Plantain subgroup)	Tanduk Galek	AAB	3x	LIPI-588	RCB-BRIN	A1	A1	A4	MZ395238	MZ318118	MZ318144
D18	*Musa* AAA (Red subgroup)	Pisang Udang	AAA	3x	LIPI-261	RCB-BRIN	A1	A1	A3	MZ395239	MZ318119	MZ318145
D19	*Musa*	Pisang Ustrali		4x	LIPI-093	RCB-BRIN	A1	A1	A4	MZ395240	MZ318120	MZ318146
D20	*Musa* AAA (Cavendish subgroup)	Cavendish	AAA	3x	LIPI-217	RCB-BRIN	A1	A3	A5	MZ395241	MZ318121	MZ318147
D21	*Musa* AAA (Gros Michel subgroup)	Pisang Ambon	AAA	3x	LIPI-250	RCB-BRIN	A1	A3	A5	MZ395242	MZ318122	MZ318148
D22	*Musa balbisiana* Colla	Pisang Klutuk Sukun	BB	2x	P1980041; vak.XXIV.D.1	PBG-BRIN	B1	B2	B1	MZ395243	MZ318123	MZ318149
D23	*Musa balbisiana* Colla	Pisang Klutuk Wulung	BB	2x	P197707103; vak. XXIV.B.19-ab	PBG-BRIN	B1	B1	B1	MZ395244	MZ318124	MZ318150
D24	*Musa boornensis*	Pisang Hutan		2x	P2016090046; vak.XXIV.D.109	PBG-BRIN	Bo1	Bo1	Bo1	MZ395245	MZ318125	MZ318151
D25	*Musa acuminata* var. *flava*	Pisang Hutan	AA	2x	-	PBG-BRIN	A1	A2	A2	MZ395246	MZ318126	MZ318152
D26	*Musa acuminata* var. *nakaii*	Cau Kole	AA	2x	-	PBG-BRIN	A1	A1	A3	MZ395247	MZ318127	MZ318153
GenBank 1	*Musa acuminata*	-	AA	2x	-	-	A1	-	-	AF378770	-	-
GenBank 2	*Musa acuminata* subsp. *malaccensis*	-	AA	2x	-	-	A1	-	-	MN822070	-	-
GenBank 3	*Musa acuminata* var. *flava*	-	AA	2x	-	-	A1	-	-	MK238286	-	-

continued

Table 2 Continued

Sample code	Scientific name	Local name	Genome composition	Ploidy level	Accession no.	Collection	Chlorotype			GenBank accession no.		
							*rbc*L	*mat*K	*trn*L-F	*rbc*L	*mat*K	*trn*L-F
GenBank 4	*Musa acuminata* var. *rutilifes*	-	AA	2x	-	-	A1	-	-	MN822067	-	-
GenBank 5	*Ensete glaucum*	-		2x	-	-	G1	-	-	MN822062	-	-
GenBank 6	*Ensete glaucum*	-		2x	-	-	G1	-	-	MN822063	-	-
GenBank 7	*Ensete glaucum*	-		2x	-	-	G1	-	-	FJ871853	-	-
GenBank 8	*Musella lasiocarpa*	-		2x	-	-	L1	-	-	AF243844	-	-
GenBank 9	*Musa balbisiana*	-	BB	2x	-	-	B1	-	-	KJ506057	-	-
GenBank 10	*Musa acuminata* var. *zebrina*	-	AA	2x	-	-	-	A1	-	-	KC904687	-
GenBank 11	*Musa acuminata*	-	AAA	2x	-	-	-	A3	-	-	KX619465	-
GenBank 12	*Musa acuminata* subsp. *malaccensis*	-	AA	2x	-	-	-	A3	-	-	KC904703	-
GenBank 13	*Musa balbisiana*	-	BB	2x	-	-	-	B1	-	-	MG041499	-
GenBank 14	*Musa balbisiana*	-	BB	2x	-	-	-	B1	-	-	MG041489	-
GenBank 15	*Musa balbisiana*	-	BB	2x	-	-	-	B1	-	-	KC904717	-
GenBank 16	*Musa balbisiana*	-	BB	2x	-	-	-	B1	-	-	KC904718	-
GenBank 17	*Ensete glaucum*	-		2x	-	-	-	G1	-	-	GQ374836	-
GenBank 18	*Ensete glaucum*	-		2x	-	-	-	G1	-	-	KX619473	-
GenBank 19	*Ensete glaucum*	-		2x	-	-	-	G1	-	-	FJ871677	-
GenBank 20	*Musa velutina*	-		2x	-	-	-	V1	-	-	FJ871653	-
GenBank 21	*Musa velutina*	-		2x	-	-	-	V1	-	-	KX619467	-
GenBank 22	*Musella lasiocarpa*	-		2x	-	-	-	L1	-	-	AF478909	-
GenBank 23	*Musa acuminata* subsp. *burmannicoides*	-	AA	2x	-	-	-	-	A1	-	-	FJ428170
GenBank 24	*Musa acuminata* subsp. *malaccensis*	-	AA	2x	-	-	-	-	A1	-	-	KT257572
GenBank 25	*Musa acuminata* subsp. *burmannica*	-	AA	2x	-	-	-	-	A1	-	-	FJ428169
GenBank 26	*Musa acuminata* subsp. *microcarpa*	-	AA	2x	-	-	-	-	A3	-	-	FJ428174
GenBank 27	*Musa acuminata*	-	AA	2x	-	-	-	-	A3	-	-	KP208935
GenBank 28	*Musa acuminata* subsp. *zebrina*	-	AA	2x	-	-	-	-	A3	-	-	FJ428173
GenBank 29	*Musa balbisiana*	-	BB	2x	-	-	-	-	B1	-	-	KT257585
GenBank 30	*Musa balbisiana*	-	BB	2x	-	-	-	-	B1	-	-	FJ621280
GenBank 31	*Musa beccarii*	-		2x	-	-	-	-	Bc1	-	-	AF431635
GenBank 32	*Musa borneensis*	-		2x	-	-	-	-	Bo2	-	-	FJ621265
GenBank 33	*Musellalasiocarpa*	-		2x	-	-	-	-	L1	-	-	KT257602
GenBank 34	*Ensete glaucum*	-		2x	-	-	-	-	G1	-	-	KT257600
GenBank 35	*Ensete glaucum*	-		2x	-	-	-	-	G1	-	-	FJ428154
GenBank 36	*Ensete glaucum*	-		2x	-	-	-	-	G2	-	-	GQ374803

Table 3 Parts of *trn*L-F sequence alignment showing large bases insertion in 11 accessions of *Musa balbisiana*

Sample code	Scientific name	Base position																																									
		290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333
GenBank 1	*Ensete glaucum* KT257600	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
GenBank 2	*Musella lasiocarpa* KT257062	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D1	*Musa acuminata* ssp. *malaccensis* 'Pisang Hutan'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D3	*Musa* AA (Sucrier subgroup) 'Pisang Rejang'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D4	*Musa acuminata* 'Pisang Rejang'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D5	*Musa acuminata* var. *breviformis* 'Cau Kole'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D6	*Musa acuminata* 'Pisang Rejang'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D8	*Musa* AB (Ney Poovan subgroup) 'Ney Poovan'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D12	*Musa* AAA 'Pisang Barangan'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D14	*Musa* AAB (Silk subgroup) 'Pisang Raja Sereh'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D16	*Musa* AA (Sucrier subgroup) 'Mas Kirana'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D17	*Musa* AAB (Plantain subgroup) 'Tanduk Galek'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D18	*Musa* AAA (Red subgroup) 'Pisang Udang'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D19	*Musa* 'Pisang Ustrali'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D20	*Musa* AAA (Cavendish subgroup) 'Cavendish'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D21	*Musa* AAA (Gros Michel subgroup) 'Pisang Ambon'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D24	*Musa boornensis* 'Pisang Hutan'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D25	*Musa acuminata* var. *flava* 'Pisang Hutan'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D26	*Musa acuminata* var. *nakaii* 'Cau Kole'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	T	T	A	T	A	A	T	A	T	G	A	A

continued

Table 3 Continued

Sample code	Scientific name	Base position																																									
		290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333
GenBank 3	*Musa balbisiana* KT257585	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	A	A	T	T	A	T	A	A	T	A	T	G	A	A	A	A	A	T	T	A	T	A	A	T	A	T	G	A	A
GenBank 4	*Musa balbisiana* FJ621280	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	A	A	T	T	A	T	A	A	T	A	T	G	A	A	A	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D2	*Musa balbisiana* Colla 'Pisang Klutuk Wulung'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	A	A	T	T	A	T	A	A	T	A	T	G	A	A	A	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D7	*Musa* ABB (Saba subgroup) 'Pisang Kepok'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	A	A	T	T	A	T	A	A	T	A	T	G	A	A	A	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D9	*Musa balbisiana* Colla 'Pisang Klutuk'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	A	A	T	T	A	T	A	A	T	A	T	G	A	A	A	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D10	*Musa balbisiana* Colla 'Pisang Klutuk Sukun'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	A	A	T	T	A	T	A	A	T	A	T	G	A	A	A	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D11	*Musa balbisiana* Colla 'Pisang Klutuk Sukun'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	A	A	T	T	A	T	A	A	T	A	T	G	A	A	A	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D13	*Musa* ABB (Pisang Awak subgroup) 'Pisang Siem'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	A	A	T	T	A	T	A	A	T	A	T	G	A	A	A	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D15	*Musa* AAB (Pisang Raja subgroup) 'Pisang Raja'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	A	A	T	T	A	T	A	A	T	A	T	G	A	A	A	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D22	*Musa balbisiana* Colla 'Pisang Klutuk Sukun'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	A	A	T	T	A	T	A	A	T	A	T	G	A	A	A	A	A	T	T	A	T	A	A	T	A	T	G	A	A
D23	*Musa balbisiana* Colla 'Pisang Klutuk Wulung'	A	T	T	A	T	A	A	T	T	A	T	A	T	A	T	A	A	T	T	A	T	A	A	T	A	T	G	A	A	A	A	A	T	T	A	T	A	A	T	A	T	G	A	A

Table 4 Banana species, varieties and cultivars with their BLAST results to the NCBI GenBank

Sample Code	Scientific Name	Genome composition	*mat*K				*rbc*L				*trn*L-F			
			Result of BLASTn	Max score	Maximal percentage identity (MPI)	Mismatch/ alignment length	Result of BLASTn	Max score	Maximal percentage identity (MPI)	Mismatch/ alignment length	Result of BLASTn	Max score	Maximal percentage identity (MPI)	Mismatch/ alignment length
D1	*Musa acuminata* ssp. *malaccensis*	AA	*Musa acuminata*	1391	100%	0/753	*Musa acuminata* sub. *malaccensis*	1173	100%	0/635	*Musa acuminata* sub. *malaccensis*	1615	100%	0/874
D2	*Musa balbisiana* Colla	BB	*Musa balbisiana*	1378	100%	0/746	*Musa balbisiana*	1173	100%	0/635	*Musa balbisiana*	1642	100%	0/889
D3	*Musa* AA (Sucrier subgroup) 'Pisang Rejang'	AA	*Musa acuminata*	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata* sub. *malaccensis*	1615	100%	0/874
D4	*Musa acuminata*	AA	*Musa acuminata*	1391	100%	0/753	*Musa acuminata* var. *flava*	1173	100%	0/635	*Musa acuminata* sub. *malaccensis*	1611	99.89%	1/874
D5	*Musa acuminata* var. *breviformis*	AA	*Musa acuminata*	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1615	100%	0/874
D6	*Musa acuminata*	AAA	*Musa acuminata*	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1615	100%	0/874
D7	*Musa* ABB (Saba subgroup) 'Pisang Kepok'	ABB	*Musa balbisiana*	1391	100%	0/753	*Musa balbisiana*	1173	100%	0/635	*Musa balbisiana*	1504	100%	0/814
D8	*Musa* AB (Ney Poovan subgroup) 'Ney Poovan'	AB	*Musa* AB group	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1609	99.89%	1/874
D9	*Musa balbisiana* Colla	BB	*Musa balbisiana*	1391	100%	0/753	*Musa balbisiana*	1173	100%	0/635	*Musa balbisiana*	1642	100%	0/889
D10	*Musa balbisiana* Colla	BB	*Musa balbisiana*	1391	100%	0/753	*Musa balbisiana*	1173	100%	0/635	*Musa balbisiana*	1642	100%	0/889
D11	*Musa balbisiana* Colla	BB	*Musa balbisiana*	1391	100%	0/753	*Musa balbisiana*	1173	100%	0/635	*Musa balbisiana*	1642	100%	0/889
D12	*Musa* AAA	AAA	*Musa* AB group	1391	100%	0/753	*Musa acuminata* var. *flava*	1173	100%	0/635	*Musa acuminata*	1609	99.89%	1/874
D13	*Musa* ABB (Pisang Awak subgroup) 'Pisang Siem'	ABB	*Musa balbisiana*	1391	100%	0/753	*Musa balbisiana*	1173	100%	0/635	*Musa balbisiana*	1642	100%	0/889
D14	*Musa* AAB (Silk subgroup) 'Pisang Raja Sereh'	AAB	*Musa* AAB group	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1609	99.89%	1/874
D15	*Musa* AAB (Pisang Raja subgroup) 'Pisang Raja'	AAB	*Musa balbisiana*	1391	100%	0/753	*Musa balbisiana*	1173	100%	0/635	*Musa balbisiana*	1642	100%	0/889
D16	*Musa* AA (Sucrier subgroup) 'Mas Kirana'	AA	*Musa acuminata*	1393	100%	0/754	*Musa acuminata*	1173	100%	0/635	*Musa acuminata* sub. *malaccensis*	1615	100%	0/874
D17	*Musa* AAB (Plantain subgroup) 'Tanduk Galek'	AAB	*Musa* AB group	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1554	100%	0/841
D18	*Musa* AAA (Red subgroup) 'Pisang Udang'	AAA	*Musa* AB group	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1615	100%	0/874
D19	*Musa*		*Musa* AB group	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1554	100%	0/841

continued

Table 4 Continued

Sample Code	Scientific Name	Genome composition	*mat*K				*rbc*L				*trn*L-F			
			Result of BLASTn	Max score	Maximal percentage identity (MPI)	Mismatch/ alignment length	Result of BLASTn	Max score	Maximal percentage identity (MPI)	Mismatch/ alignment length	Result of BLASTn	Max score	Maximal percentage identity (MPI)	Mismatch/ alignment length
D20	*Musa* AAA (Cavendish subgroup) 'Cavendish'	AAA	*Musa acuminata*	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1615	100%	0/874
D21	*Musa* AAA (Gros Michel subgroup) 'Pisang Ambon'	AAA	*Musa acuminata*	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1476	100%	799
D22	*Musa balbisiana* Colla	BB	*Musa balbisiana*	1378	100%	0/746	*Musa balbisiana*	1173	100%	0/635	*Musa balbisiana*	1642	100%	0/889
D23	*Musa balbisiana* Colla	BB	*Musa balbisiana*	1378	100%	0/753	*Musa balbisiana*	1173	100%	0/635	*Musa balbisiana*	1635	99.89%	1/889
D24	*Musa boornensis*		*Musa boornensis*	1375	99.73%	2/751	*Musa boornensis*	1173	100%	0/635	*Musa boornensis*	1611	99.89%	1/874
D25	*Musa acuminata* var. *flava*	AA	*Musa* AB group	1386	99.87%	1/753	*Musa acuminata* var. *flava*	1173	100%	0/635	*Musa acuminata*	1602	99.77%	2/874
D26	*Musa acuminata* var. *nakaii*	AA	*Musa* AB group	1391	100%	0/873	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1607	99.89%	1/874
D16	*Musa* AA (Sucrier subgroup) 'Mas Kirana'	AA	*Musa acuminata*	1393	100%	0/754	*Musa acuminata*	1173	100%	0/635	*Musa acuminata sub. malaccensis*	1615	100%	0/874
D17	*Musa* AAB (Plantain subgroup) 'Tanduk Galek'	AAB	*Musa* AB group	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1554	100%	0/841
D18	*Musa* AAA (Red subgroup) 'Pisang Udang'	AAA	*Musa* AB group	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1615	100%	0/874
D19	*Musa*		*Musa* AB group	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1554	100%	0/841
D20	*Musa* AAA (Cavendish subgroup) 'Cavendish'	AAA	*Musa acuminata*	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1615	100%	0/874
D21	*Musa* AAA (Gros Michel subgroup) 'Pisang Ambon'	AAA	*Musa acuminata*	1391	100%	0/753	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1476	100%	799
D22	*Musa balbisiana* Colla	BB	*Musa balbisiana*	1378	100%	0/746	*Musa balbisiana*	1173	100%	0/635	*Musa balbisiana*	1642	100%	0/889
D23	*Musa balbisiana* Colla	BB	*Musa balbisiana*	1378	100%	0/753	*Musa balbisiana*	1173	100%	0/635	*Musa balbisiana*	1635	99.89%	1/889
D24	*Musa boornensis*		*Musa boornensis*	1375	99.73%	2/751	*Musa boornensis*	1173	100%	0/635	*Musa boornensis*	1611	99.89%	1/874
D25	*Musa acuminata* var. *flava*	AA	*Musa* AB group	1386	99.87%	1/753	*Musa acuminata* var. *flava*	1173	100%	0/635	*Musa acuminata*	1602	99.77%	2/874
D26	*Musa acuminata* var. *nakaii*	AA	*Musa* AB group	1391	100%	0/873	*Musa acuminata*	1173	100%	0/635	*Musa acuminata*	1607	99.89%	1/874

BLAST = basic local alignment search tool, NCBI = National Centre for Biotechnology Information

ACKNOWLEDGMENT

The authors greatfully acknowledge the Director of the Research Center for Biology-LIPI for providing banana samples. The authors thank Poerba YS for her suggestions on selecting samples prior to this study. The authors would also like to thank the Director of Kebun Raya Purwodadi for granting permission, and Hapsari L for helping to collect banana samples. The authors acknowledge the facilities, and the scientific and technical assistance of the Integrated Laboratory of Bioproducts (iLaB), Indonesian Institute of Sciences. This work was supported by Productive-Innovative Research Program (RISPRO) by Indonesia Endowment Fund for Education (LPDP), Ministry of Finance, in a collaboration with Ministry of Research and Technology/National Research and Innovation Agency (Kemenristek-BRIN), under the National Research Priority (PRN) program (grant number 273/E1/PRN/2020 and B-4528/IPH/KS.02.04/VII/2020).

REFERENCES

Ainiyah RK, Wahyunindita V, Pratama WN, Pratiwi IA, Utami ESW & Hariyanto S. 2020. DNA barcoding: study of bananas (*Musa* spp.) wild and cultivars group from East Java inferred by *rbc*L gene sequences. *Ecology, Environment and Conservation Paper* 26: S7–S13.

Ahmad F. 2021. *Genetics and Diversity of Indonesian Bananas*. PhD thesis. Wageninen University & Research, Wageningen.

Cheesman E. 1948. Classification of the bananas: critical notes on species: *Musa nagensium*. *Kew Bulletin*: 325–328.

D'yachenko E, Filyushin M, Pronina E & Kochieva E. 2015. Variability of the *trn*L plastid gene's intron in the Faboideae species (Fabaceae). *Russian Journal of Genetics: Applied Research* 5: 220–226.

De Langhe E, Vrydaghs L, De Maret P, Perrier X & Denham T. 2009. Why bananas matter: an introduction to the history of banana domestication. *Ethnobotany Research and Applications* 7: 165–177.

Dhivya S, Ashutosh S, Gowtham I, Baskar V, Harini AB, Mukunthakumar S & Sathishkumar R. 2020. Molecular identification and evolutionary relationships between the subspecies of *Musa* by DNA barcodes. *BMC Genomics* 21: 1–11.

Dwivany FM, Ramadhan MR, Lim C et al. 2020. Bali bananas (*Musa* spp. L.) genetic relationship based on Internal Transcribed Spacer 2 (ITS-2). *Pertanika Journal of Tropical Agricultural Science* 43: 583–597.

Häkkinen M & Väre H. 2008. Typification and checklist of *Musa* L. names (Musaceae) with nomenclature notes. *Adansonia* 30: 63–112.

Häkkinen M. 2013. Reappraisal of sectional taxonomy in *Musa* (Musaceae). *Taxon* 62: 809–813.

Hapsari L, Azrianingsih R & Arumingtyas EL. 2018. Genetic variability and relationship of banana cultivars (*Musa* L.) from East Java, Indonesia based on the Internal Transcribed Spacer region nrDNA sequences. *Journal of Tropical Biology & Conservation (JTBC)* 15: 101–120.

Hebert PD, Cywinska A, Ball SL & Dewaard JR. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London Series B: Biological Sciences* 270: 313–321.

Heslop-Harrison JS & Schwarzacher T. 2007. Domestication, genomics and the future for banana. *Annals of Botany* 100: 1073–1084.

Kitching I, Forey PL, Humphries CJ & Williams D. 1998. *The Theory and Practice of Parsimony Analysis. Cladistics Second Edition*. Oxford University Press, Oxford.

Kress WJ. 1990. The phylogeny and classification of the Zingiberales. *Annals of the Missouri Botanical Garden*: 698–721.

Kress WJ & Erickson DL. 2007. A two-locus global DNA barcode for land plants: the coding *rbc*L gene complements the non-coding trnH-psbA spacer region. *PLoS one* 2: e508.

Kress WJ, Prince LM & Williams KJ. 2002. The phylogeny and a new classification of the gingers (Zingiberaceae): evidence from molecular data. *American Journal of Botany* 89: 1682–1696.

Kress WJ & Specht CD. 2006. The evolutionary and biogeographic origin and diversification of the tropical monocot order Zingiberales. *Aliso: A Journal of Systematic and Evolutionary Botany* 22: 621–632.

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA & Janzen DH. 2005. Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences* 102: 8369–8374.

Kuhn TS, Mooers AØ & Thomas GH. 2011. A simple polytomy resolver for dated phylogenies. *Methods in Ecology and Evolution* 2: 427–436.

Li LF, Häkkinen M, Yuan YM, Hao G & Ge XJ. 2010. Molecular phylogeny and systematics of the banana family (Musaceae) inferred from multiple nuclear and chloroplast DNA fragments, with a special reference to the genus *Musa*. *Molecular Phylogenetics and Evolution* 57: 1–10.

Li LF, Wang HY, Zhang C, Wang XF, Shi FX, Chen WN & Ge XJ. 2013. Origins and domestication of cultivated banana inferred from chloroplast and nuclear genes. *PLoS One* 8: e80502.

Nasution RE. 1991. A taxonomic study of the *Musa acuminata* Colla with its intraspecific taxa in Indonesia. *Memoir of Tokyo University of Agriculture* 32: 1–122.

Nasution RE & Yamada I. 2001. *Pisang-Pisang Liar di Indonesia*. Puslitbang Biologi-LIPI, Bogor.

Pirie MD, Vargas MPB, Botermans M, Bakker FT & Chatrou LW. 2007. Ancient paralogy in the cpDNA *trn*L-F region in Annonaceae: implications for plant molecular systematics. *American Journal of Botany* 94: 1003–1016.

Ruas M, Guignon V, Sempere G et al. 2017. MGIS: managing banana (*Musa* spp.) genetic resources information and high-throughput genotyping data. *Database 2017* 46: 1–12. https://doi.org/10.1093/database/bax046.

Simmonds NW. 1962. *The Evolution of the Bananas*. Longmans, London.

Simmonds NW & Shepherd K. 1955. The taxonomy and origins of the cultivated bananas. *Botanical Journal of the Linnean Society* 55: 302–312.

Sefc KM, Lefort F, Grando MS, Scott KD, Steinkelinier H & Thomas ME. 2001. Microsatellite markers for grapevine: a state of the art. Pp 1–30 in Roubelakis-Angelakis (ed). *Molecular Biology & Biotechnology of Garpevine*. Kluwer Academic Publishers, The Netherlands.

Sulistyaningsih LD. 2016. The diversity of wild banana species (Genus *Musa*) in Java. *Makara Journal of Science* 20: 40–48.

Sulistyaningsih LD, Megia R & Widjaja EA. 2014. Two new records of wild bananas (*Musa balbisiana* and *Musa itinerans*) from Sulawesi. *Makara Journal of Science* 18: 1–6.

Taberlet P, Gielly L, Patou G & Bouvet J. 1991. Universal primers for amplifification of three noncoding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.

Swofford DL. 1998. *PAUP: Phylogenetic Analysis Using Parsimony (and Other Methods)*. Sinauer Associates, Sunderland, MA.

Valmayor R, Jamaluddin S, Silayoi B, Kusumo S, Danh L, Pascua O & Espino R. 2000. *Banana Cultivar Names and Synonyms in Southeast Asia*. International Network for the Improvement of Banana and Plantain — Asia and the Pacific, Los Banos.

Volkaert H. 2011. Molecular analysis reveals multiple domestications of edible bananas. *Acta Horticulturae* 897: 143–152. https://doi.org/10.17660/ActaHortic.2011.897.15.

Wang Z, Miao H, Liu J et al. 2019. *Musa balbisiana* genome reveals subgenome evolution and functional divergence. *Nature Plants* 5: 810–821.

Wilkie P, Poulsen AD, Harris D & Forrest LL. 2013. The collection and storage of plant material for DNA extraction: the teabag method. *Gardens' Bulletin Singapore* 65: 231–234.

Yulita KS. 2013. Secondary structures of chloroplast *trn*L intron in Dipterocarpaceae and its implication for the phylogenetic reconstruction. *Hayati Journal of Biosciences* 20: 31–39.