Indian J. Genet., 80(3) 301-307 (2020)
DOI: 10.31742/IJGPB.80.3.9

Haplotype network analysis of wild banana relatives *Ensete glaucum*, *Musa acuminata* and *Musa balbisiana* based on cpDNA *rbcL* sequences in *ex-situ* collection

Lia Hapsari*, Dewi Ayu Lestari and Rasyadan Taufiq Probojati[1]

Purwodadi Botanic Garden – Research Center for Plant Conservation and Botanic Gardens, Indonesian Institute of Sciences, Pasuruan, East Java, 67163, Indonesia; [1]Department of Biology, Faculty of Mathematic and Natural Sciences, State University of Malang, Malang, East Java, 65145, Indonesia

(Received: March 2020; Revised: July 2020; Accepted: July 2020)

Abstract

Conservation and genetic assessment of wild banana relatives is important for future breeding purposes. Haplotype network analysis was conducted to wild banana relatives comprised *Ensete glaucum*, *Musa acuminata* and *Musa balbisiana*using *rbcL* gene sequences. Sequences characterization showed high conservation level (91%), low indels (1.83%), and low parsimony informatives (3.51%). However, it was powerfull to separate the dataset at families, genera and species level; and moderately to separate at intraspecies level of wild bananas. Haplotype and nucleotide diversity of wild bananas were high. At intraspecies level, both *M. acuminata* and *M. balbisiana* showed high haplotype diversity but low nucelotide diversity among haplotypes; *M. acuminata* has higher value than *M. balbisiana*. No haplotype and nucleotide diversity in *E. glaucum*. Wild bananas were separated into seven haplotypes, with four haplogroups. Mutational pathway revealed that *E. glaucum* haplotype became root; and was closely related to *M. balbisiana* than *M. acuminata*. *M. acuminata* var. *malaccensis* haplotype became root within all *M. acuminata* varieties, and each haplotype differed by single point mutation.

Key words: Conservation genetic, DNA sequences, molecular, polymorphism, phylogenetic

Introduction

Zingiberales order comprised some noticeable taxa, including gingers (Zingiberaceae), bird of paradise (Strelitziaceae), heliconias (Heliconiaceae), and bananas (Musaceae) (Kress et al. 2001). In particular of banana family, it recognized of genera, namely *Ensete* Bruce ex Horan, *Musella* (Fr.) C.Y. Wu and *Musa* L. Cultivated bananas were assumed to have derived from complex hybridization and domestication process of wild species, mainly *Musa acuminata* Colla and/or *Musa balbisiana* Colla (De Langhe et al. 2009).Elucidating the phylogenetic and haplotype network within wild banana relatives and their closely related species is important for future breeding purposes of modern bananas.

Banana improvement supported by genomic research with collection and assessment of wild banana relatives diversity are important to ensure the modern bananas future. It provides large supply of resistance genes for biotic and abiotic stresses, and other desired traits (Heslop-Harrison 2011). Nevertheless, it has been under considerable threats due to habitat destruction, fragmentation and conversion of tropical forests, and other anthropogenic disturbances (Ford-Lloyd et al. 2011). Therefore, it is important to prioritize the collection, effective conservation, improving the availability, and providing related assessments of wild bananas for use in plant breeding.

Haplotype networks are connected graphs with cycles used in the analysis of population genetic data to visualise genealogical relationships at intraspecific level, as well as to make inference about gene flow, biogeography and history of populations. Haplotype network considered more informative than conventional phylogenetic trees to display intraspecific DNA sequence variation (Mardulyn 2012; Paradis 2018). Aims of this study was to evaluate haplotype diversity and distribution of wild banana relatives in Purwodadi

*Corresponding author's e-mail: hapsari.lia@gmail.com; lia.hapsari@lipi.go.id

Published by the Indian Society of Genetics & Plant Breeding, A-Block, F2, First Floor, NASC Complex, IARI P.O., Pusa Campus, New Delhi 110 012; Online management by www.isgpb.org; indianjournals.com

Botanic Garden *ex-situ* collection using *rbcL* gene sequences; and its implications for conservation and breeding purposes. *RbcL* gene located in chloroplast genome is one of valuable molecular barcode for genetic diversity and phylogenetic studies (Hollingsworth et al. 2009). Previous molecular studies using *rbcL*were reported often used and suitable for Monocots, particularly in Zingiberales order (Kress et al. 2001), but still limited on bananas (Musaceae).

Materials and methods

Plant materials

Plant materials comprised of thirteen accessions of wild bananas collection of Purwodadi Botanic Garden/ PBG, located in Pasuruan, East Java as in group. It consisted of three species i.e., *E. glaucum*, *M. balbisiana* and *M. acuminata*. Three accessions of *E. glaucum* were collected from East Java and West Java populations. Three accessions of *M. balbisiana* were comprised of two morphologically different specimens i.e., Pisang Klutuk Ijo (green) and Pisang Klutuk Wulung (black-purple). Seven accessions of *M. acuminata* were comprised of five varieties and two un-identified varieties. Three species of Heliconiaceae family and two species of Strelitziaceae family were used as an outgroup, retrieved from Genbank (Table 1, Fig. 1).

DNA extraction, amplification and sequencing

The samples were fresh young furled leaves, one individual per accession. Whole genome DNAs were extracted using PromegaWizard® Genomic DNA Purification Kit followed the manufacturer's protocols for plant. Primer pairs used for amplification i.e. *rbcL*-1F (5'-ATGTCACCACAAACAGAAAC-3') and *rbcL*724R (5'-TCGCATGTACCTGCAGTAGC-3') referred to CBOL (2009). Amplified products were purified and sequenced at *1st BASE Laboratories* (Malaysia) using ABI PRISM 3730xl Genetic Analyzer. All *rbcL* nucleotide sequences from this study have been submitted to Genbank database (Table 1).

Haplotype network analysis

Sequences evaluation were conducted using ABI sequences Scanner v.10. Multiple sequences alignments were performed using ClustalW program in MEGA7.0.26. Nucleotides variation were analyzed with Bioedit 7.0.5.3 and DnaSP 6.12.03. Haplotype diversity and distribution map was performed through median joining analysis using Haplotype Network 5.0.1.1 (Paradis 2018). Phylogenetic reconstructions and pairwise distance analyses were also employed using MEGA7.0.26, Kimura 2 parameter evolution model, Neighbor Joining/NJ algorythm, 1000 bootstrap replicates (Kumar et al. 2016).

Results and discussion

Molecular characterization

DNA sequences produced within wild banana relatives were 713 to 723 bp. BLASTn to Genbank database showed homology with *rbcL* sequences of species from Zingiberales order (similarity 96-99%). The total aligned and selected sequences were 712 bp. It shows high conservation level (655 monomorphic sites; 91.99%); with low indels (13 sites; 1.83%), and low parsimony informatives (25 sites; 3.51%). Further, the dataset were shared high similarity of *rbcL* sequences (96.93-100%) among ingroup and outgroup, and even lower within ingroup (98.12-100%). Likewise, previous studies showed high *rbcL* sequences similarity 98-100% among Fe'i bananas populations and *Ensete* spp. in Maluku Islands (Hiariej et al. 2015); also within *Heliconia* species 99.98-100% (Hapsari et al. 2019).

Nucleotide composition were high in AT (56.99%) than GC (43.01%). Chloroplast genes (including *rbcL*) are dominated by a genomic bias towards a high AT than GC. Low GC content means low spots of mutation and recombination rates (Kiktev et al. 2018). About 34 mutations were identified, mostly due to substitution of transversion (58.82%) than transision (41.18%), with ratio transition/transversion (Ti/Tv) 0.7. Nucleotide substitution rate within *rbcL* on land plants were varied, the most rapid rate was found in Poaceae family, followed by families in the orders Orchidales, Liliales, Bromeliales, and Arecales (Newmaster et al. 2006).

Haplotype diversity and distribution

Haplotype diversity represents probability of two randomly selected alleles to be different from each other, whilst nucleotide diversity reflects nucleotide divergence of two individuals with respect to one locus (Sharma et al. 2013). Haplotype network analysis overall dataset (18 OTU) resulted in twelve haplotypes with high haplotype diversity (Hd>0.5) and high nucleotide diversity (π0.5%). Separate analyses, Strelitziaceae and Heliconiaceae as outgroup showed very high haplotype diversity (Hd=1.00). Haplotype and nucleotide diversity of wild bananas (Musaceae) were considered high, but the values were lower than Strelitziaceae and Heliconiaceae. Further, this study revealed that haplotype diversity at intraspecies level

Table 1. Plant materials examined

Sample code	Field coll. number	Species name/local name	Population locality	Genbank acc. number
		IN GROUP		
E1	At Nursery	*Ensete glaucum* (Roxb.) Cheesman	Bogor, West Java	MN822062
E2	At Nursery	*Ensete glaucum* (Roxb.) Cheesman	Bogor, West Java	MN822063
E3	At Nursery	*Ensete glaucum* (Roxb.) Cheesman	Pasuruan, East Java	MN822064
M1	XXIV.D.1	*Musa balbisiana* Colla/ Pisang Klutuk Ijo	Pasuruan, East Java	MK238285
M2	XXIV.B.19	*Musa balbisiana* Colla/ Pisang Klutuk Wulung	Kebumen, Central Java	MN822065
M3	At Nursery	*Musa balbisiana* Colla/ Pisang Klutuk Wulung	Pati, Central Java	MN822066
M4	XXIV.A.29	*Musa acuminata* var. *flava* (Ridl.) Nasution	Tuban, East Java	MK238286
M5	XXIV.D.12	*Musa acuminata* var. *rutilifes* Back. Nasution	Tuban, East Java	MN822067
M6	At Nursery	*Musa acuminata* var. *nakaii* Nasution	Pasuruan, East Java	MN822068
M7	XXIV.A.27	*Musa acuminata* var. *alasensis* Nasution	Solok, West Sumatera	MN822069
M8	XXIV.A.9	*Musa acuminata* var. *malaccensis* (Ridl.) Nasution	Bogor, West Java	MN822070
M9	At Nursery	*Musa acuminata* var.1/ Pisang Tengkarong	Ambon, Maluku	MN822071
M10	At Nursery	*Musa acuminata* var.2/ Pisang Hatalai	Ambon, Maluku	MN822072
		OUT GROUP		
H1	V.D.II.26	*Heliconia chartacea* Lane ex Barreiros	Brazil	MK238301
H2	V.D.II.7	*Heliconia metallica* Planch. & Linden ex Hook.	Brazil	MK238300
H3	V.D.II.17	*Heliconia psittacorum* L.f.	West Java	MK238294
S1	V.D.II.1	*Ravenala madagascariensis* Sonn.	Malagasy	MK238283
S2	V.D.II.13	*Phenakospermum guyannense* (A.Rich.) Endl. ex Miq.	Brazil	MK238284

E = *Ensete*, M = *Musa*, H = *Heliconia*, S = Strelitziaceae; Coll. = Collection

of both *M. acuminata* and *M. balbisiana* were high but low nucleotide diversity among haplotypes; *M. acuminata* has higher value than *M. balbisiana*. Meanwhile, no haplotype and nucleotide diversity observed in *E. glaucum* populations (Table 2).

The lineage pattern resulted from haplotype network analysis was clearly separated into 3 distinct groups following its family. Further, within Musaceae family it was clearly separated until species level into 3 distinct subgroups (Fig. 2). Likewise, phylogenetic analysis (NJ) also resulted into the same pattern of tree topology; and supported by high bootstraps (85-

Table 2. Haplotype and nucleotide diversity of monophyletic groups

OTU	Member of OTU	Number of haplotype	Hd±SD	π(%)±SD(%)	Eta
Overall dataset	18	12	0.95±0.03	1.69±0.25	34
Strelitziaceae	2	2	1.00±0.50	2.00±1.00	14
Heliconiaceae	3	3	1.00±0.27	1.24±0.41	13
Musaceae	13	7	0.90±0.05	0.91±0.12	15
E. glaucum	3	1	0.00±0.00	0.00±0.00	0
M. balbisiana	3	2	0.67±0.31	0.09±0.05	1
M. acuminata	7	4	0.81±0.13	0.19±0.04	3

OTU = operational taxonomic unit; Hd = haplotype diversity; π = nucleotide diversity; SD = standard deviation; Eta = total number of mutation/polymorphic site

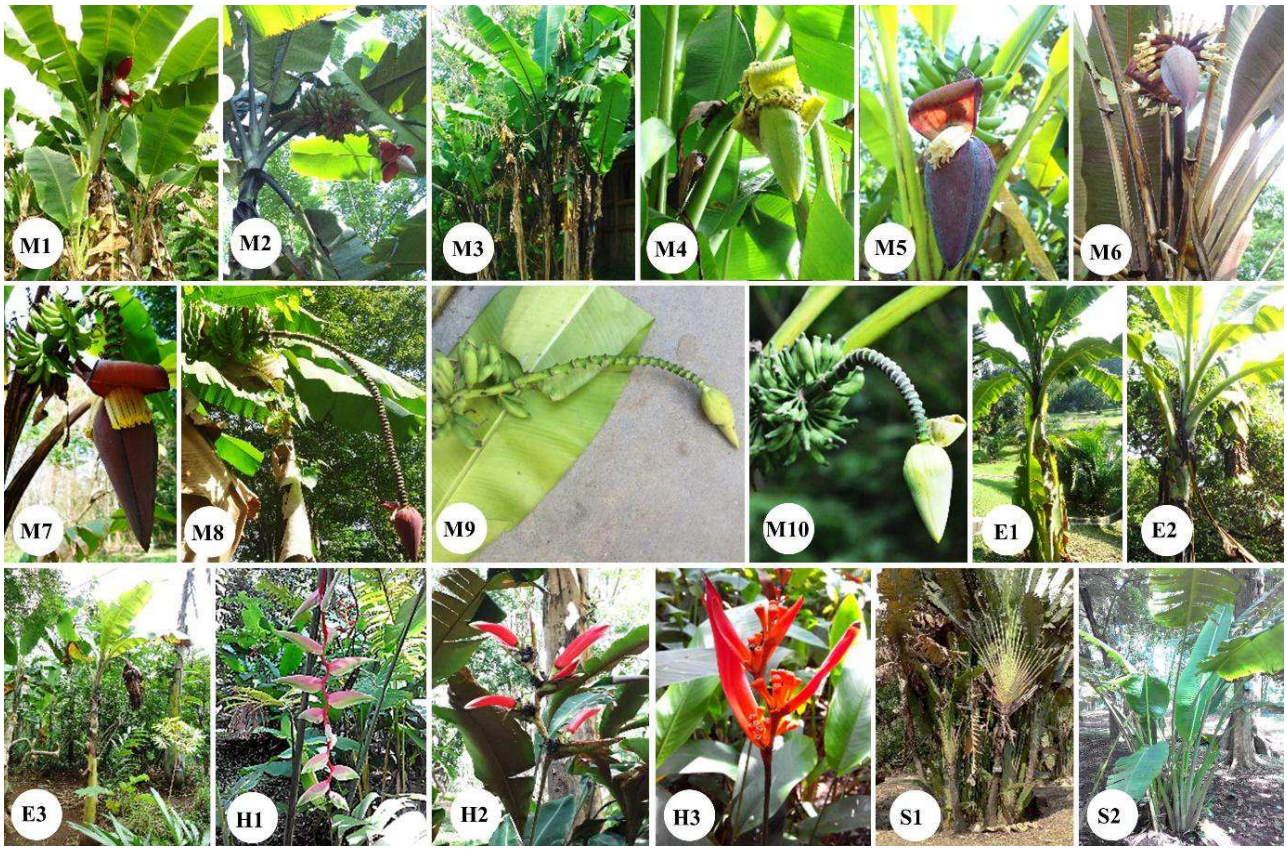

Fig. 1. Morphological appearances of wild banana relatives examined

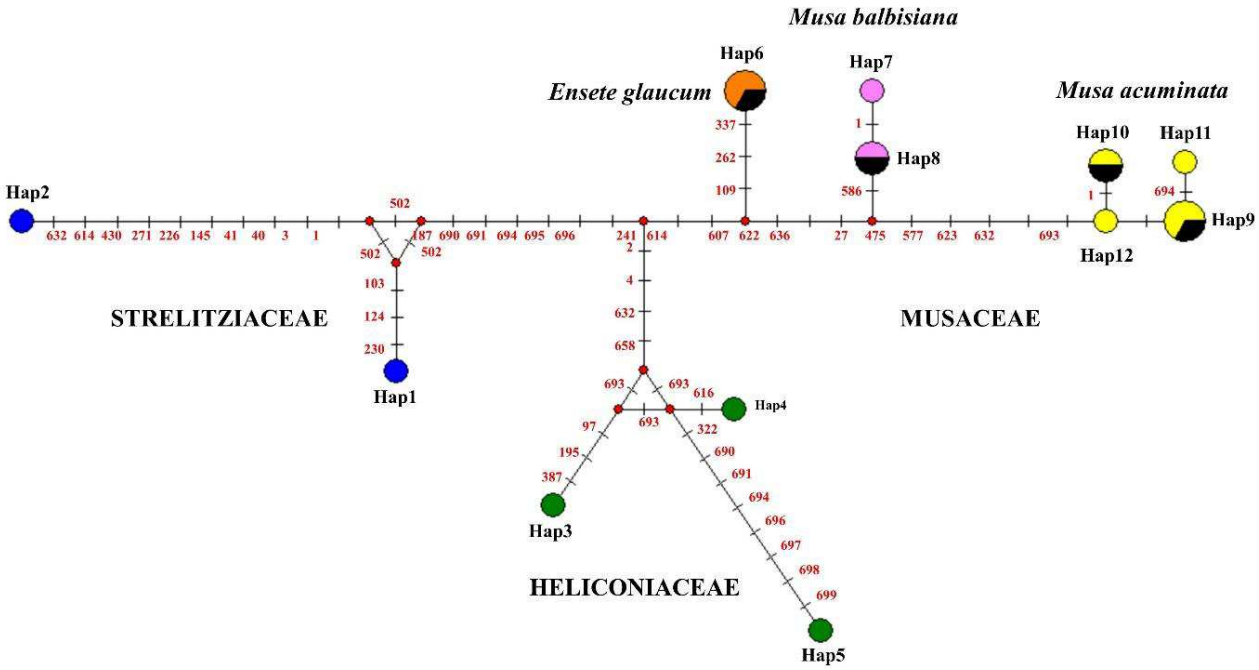

Fig. 2. Haplotype network of wild banana relatives. Haplotype distribution: Hap1=S1; Hap2=S2; Hap3=H1; Hap4=H2; Hap5=H3; Hap6=E1, E2, E3; Hap7=M1; Hap8=M2, M3; Hap9=M4, M6, M10; Hap10=M5, M9; Hap11=M7; and Hap12=M8

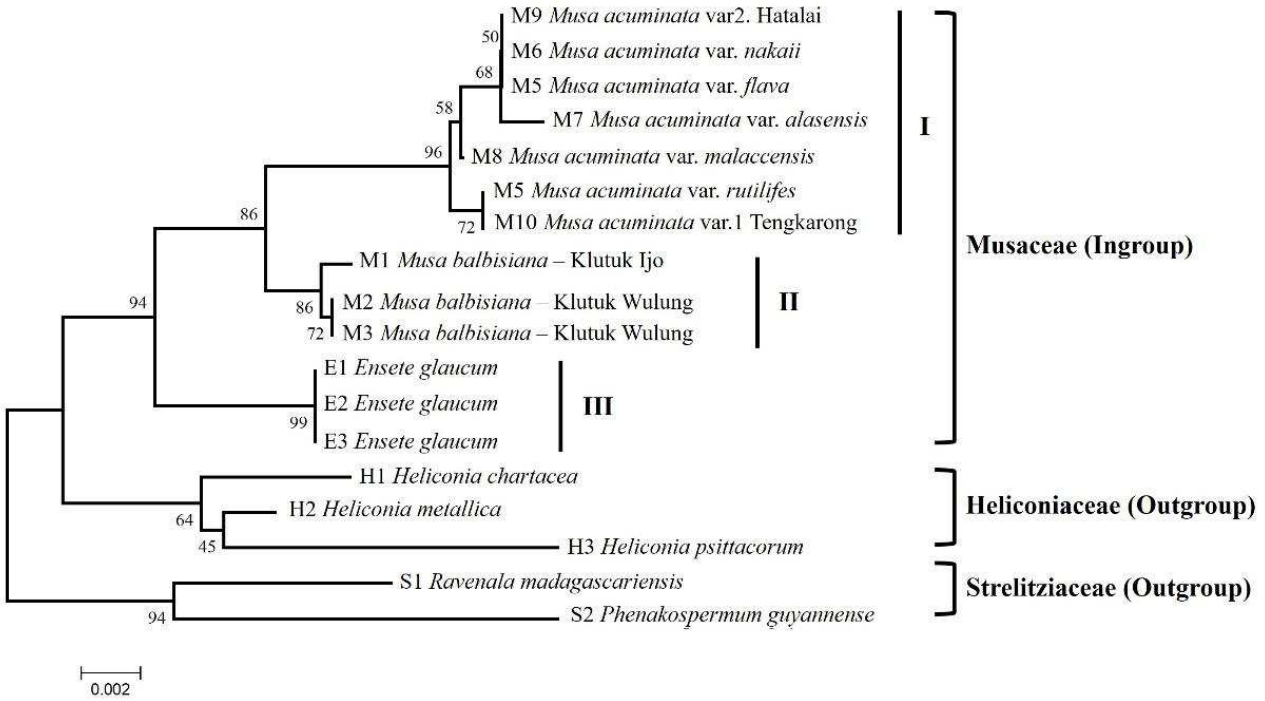

Fig. 3. Phylogenetic tree of wild banana relatives using NJ algorythm

99%) unless the separation of Heliconiaceae, yet it was still possible to identify the Heliconiaceae as a monophyletic group (Fig. 3). No bootstraping was recognised in haplotype network analysis. Nonetheless, major additional feature of haplotype networks are to add putative recombination events and alternative mutational pathways thus more informative to visualize the genealogy until intraspecies level (Fig. 2); in which it is not shown in the phylogenetic tree (Fig. 3). Strelitziaceae and Heliconiaceae comprised of five haplotypes and separated with Musaceae by 11-22 nucleotides. Musaceae family was separated into seven haplotypes, with four haplogroups. The occurrence of several polymorphic nucleotides between haplotype pairs indicates that the considered region has a high substitution rate (Nanayakkara et al. 2018). The highest substitution rate within Musaceae was between *Ensete glaucum* to *M. acuminata* haplotypes (11 nucleotides). *Ensete glaucum* was closely related to *M. balbisiana*, separated by 7 polymorphic nucelotides (Fig. 2).

In details, three specimens of *E. glaucum* originated from East Java and West Java were identical and clustered in a Haplogroup 6; and the gene flow had become root within Musaceae. *M. balbisiana* was comprised of two haplotypes. It was confirmed that Pisang Klutuk Ijo was slightly different and separated with Pisang Klutuk Wulung in Haplotype 7 by a single point mutation (nucleotide 1, T→6). Two populations of Pisang Klutuk Wulung were identical and clustered in Haplotype 8. *Musa acuminata* varieties were comprised of four haplotypes, became root within all *M. acuminata* varieties examined and each haplotype was differed only by a single point mutation (nucelotide 1/T→6,693/A→T, 694/A→T). Later, *M. acuminata* var. *rutilifes* and unidentified Pisang Tengkarong were identical and grouped in Haplotype 10; *M. acuminata* var. *flava*, *M. acuminata* var. *nakaii*, and unidentified Pisang Hatalai were considered identical and grouped in Haplotype 9. Whilst, *M. acuminata* var. *alasensis* was separated in Haplotype 11 (Fig. 2).

The present study revealed that although *rbcL* sequences were highly conserved with low parsimony informatives but it is powerfull to separate the dataset at families, genera and species level; and suggested thatit is moderately suitable to separate within species (intraspecies) level of wild bananas. It has low discrimination power at population level. However, haplotype network analysis can give better visualization of mutational pathway until intraspecific level. Further assessments using other genetic markers both from chloroplast and/or nuclear genomes

which are more powerfull in discrimination should be conducted. Nuclear DNA barcode suggested to be used in bananas i.e. nrITS (Li et al. 2010). Further, *rbcL* sequences combined with other sequences from chloroplast genome (*matK,trnH-psbA*, *trnL-F*, *rps16*, *rpoB*, etc.) generated phylogenetic trees with higher resolving power to infraspecific level compared to *rbcL* data alone (Dong et al. 2012; Lestari et al. 2018).

Implication for conservation and further banana breeding

Elucidating the haplotype diversity and distribution of wild banana relatives are important to provide valuable information to clarify the taxonomic status of collection, consideration for conservation management strategy and further banana breeding purposes. Wild banana accessions in *ex-situ* collection which molecularly revealed identical (placed at the same haplotype 8, 9, 10), if resources are limited, it suggested to be chosen one of them as a representative collection. However, if the morphology is convincingly distinguished, then another identification assessment should be conducted to confirm the taxonomic status further.

All wild banana relatives examined in this study are valuable genetic resources with high potentials.They have several important traits such as, they may be grown in wide range of conditions, relatively drought-tolerant, preliminary diseases resistant (virus, bacterial, fungi), potentialy utilized for vegetables, fibres, medicines, animal fodder, ornamental, roofing and packaging. They also have ecological importance to stabilize soils and microclimates (Majumdar et al. 2013; Hapsari and Masrum 2012; Hapsari 2014; Chen et al. 2019). The wild bananas from *ex situ* field collections were contained unique variation, although considered less genetically diverse. Thus, more efforts are needed in conservation of wild bananas from wild/natural populations from Indonesia as part of the origin center and from multiple individuals in one population. To search for more variability and desired resistance characters, concerted efforts should be made to identify the traits and breed suitable genotypes to mitigate climate change challenges.

Authors' contribution

Conceptualization of research (LH); Designing of the experiments (LH); Contribution of experimental materials (LH, DAL); Execution of field/lab experiments and data collection (LH); Analysis of data and interpretation (LH, RTP); Preparation and editing of the manuscript (LH, DAL, RTP).

Declaration

The authors declare no conflict of interest.

Acknowledgements

The authors greatly acknowledge the authorities of Purwodadi Botanic Garden for providing the plant materials for this study. Sincere thanks were also addressed to Plant Physiology Laboratory of Biology Department, Brawijaya University for the molecular and genetic facilities.

References

CBOL (The Consortium for the Barcode of Life). 2009. A DNA barcode for land plants CBOL Plant Working Group1 Communicated. PNAS, **106**(31): 12794-12797. www.pnas.org/cgi/doi/10.1073/pnas.0905845106.

Chen A., Sun J., Matthews A., Armas-Egas L., Chen N., et al. 2019. Assessing variations in host resistance to *Fusarium oxysporum* f.sp. *cubense* race 4 in *Musa* species, with a focus on the subtropical race 4. Front. Microbiol., **10**: 1062. https://doi.org/10.3389/fmicb.2019.01062.

De Langhe E., Vrydaghs L., de Maret P., Perrier X. and Denham T. 2009. Why bananas matter: An introduction to the history of banana domestication. Ethnobot. Res. Appl., **7**: 165-177. www.ethnobotany-journal.org/vol7/i1547-3465-07-165.pdf.

Dong W., Liu J., Yu J., Wang L. and Zhou S. 2012. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. PloS ONE, **7**(4): e35071. https://doi.org/10.1371/journal.pone.0035071

Hapsari L., Trimanto, and Wahyudi D. 2019. Species diversity and phylogenetic analysis of *Heliconia* spp. collections of Purwodadi Botanic Garden (East Java, Indonesia) inferred by *rbcL* gene sequences. Biodiversitas, **20**(5): 1266-1283.https://doi.org/10.13057/biodiv/d200505.

Hapsari L. 2014. Wild *Musa* species collection of Purwodadi Botanic Garden: Inventory and its morpho-taxonomic review. JTLS, **4**(1): 70-81.

Hapsari L., and Masrum A. 2012. Preliminary screening resistance of *Musa* germplasm for banana bunchy top disease in Purwodadi Botanic Garden, Pasuruan, East Java. Bot. Gard. Bull., **15**: 57-70.

Heslop-Harrison J. S. 2011. Genomics, banana breeding and super domestication. In: Van den Bergh et al. (eds). Proc. Int'l ISHS-ProMusa Symp. on Global Perspectives on Asian Challenges. Acta Hort., **897**: ISHS.

Hiariej A., Arumingtyas E. L., Widoretno W. and Azrianingsih R. 2015. Genetic kinship of tongkat langit banana (*Musa troglodytarum* L.) from Moluccas based on *rbcL* gene sequence. Indian J. Genet., **75**(4): 526-528. https://doi.org/10.5958/0975-6906.2015.00085.1.

Hollingsworth P. M., Forrest L. L., Spouge J. L., Hajibabaei M., et al. 2009. A DNA barcode for land plants. PNAS, **106**(31): 12794-12797. https://doi.org/10.1371/journal.pone.0019254.

Ford-Lloyd B. V., Schmidt M., Armstrong S. J., Barazani O. Z. et al. 2011. Crop wild relatives-undervalued underutilized and under threat? BioScience, **61**: 559-565. https://doi.org/10.1525/bio.2011.61.7.10.

Kiktev D. A., Sheng Z., Lobachev K. S. and Petes T. D. 2018. GC content elevates mutation and recombination ratesin the yeast *Saccharomyces cerevisiae*. PNAS, **115**(30): E7109-E7118. www.pnas.org/cgi/doi/10.1073/pnas.1807334115.

Kress W. J., Prince L., Hahn W. J. and Zimmer EA. 2001. Unraveling the evolutionary radiation of the families of the Zingiberales using morphological and molecular evidence. Syst. Biol., **50**(6): 926-944.

Kumar S., Stecher G. and Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. Evol., **33**: 1870-1874. https://doi.org/10.1093/molbev/msw054.

Lestari D. A., Azrianingsih R. and Hendrian. 2018. Phylogenetics of Annonaceae species from East Java collections of Purwodadi Botanic Garden based on coding and non-coding DNA sequences. JTBB, **3**: 1-7. https://doi.org/10.22146/jtbb.28308.

Li L-F., Hakkinen M., Yuan Y-M., Hao G. and Ge X-J. 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*. Mol. Phyl. Evol., **57**: 1-10. https://doi.org/10.1016/j.ympev.2010.06.021.

Majumdar K., Sarkar A., Deb D., Majumder J. and Datta B. K. 2013. Distribution record of *Ensete glaucum* (Roxb.) Cheesm. (Musaceae) in Tripura, Northeast India: a rare wild primitive banana. AJCB, **2**(2): 164-167.

Mardulyn P. 2012. Trees and/or networks to display intraspecific DNA sequence variation? Mol. Biol., **21**: 3385-3390.

Nanayakkara N., Herath V, and Jayatilake D. V. 2018. Haplotype diversity analysis of bacterial leaf blight resistance gene *Xa21* in rice. Trop. Agr. Res., **30**(1): 56-68. https://doi.org/10.4038/tar.v30i1.8278.

Newmaster S. G., Fazekas A. J., and Ragupathya S. 2006. DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. Canadian J. Bot., **84**: 335-341. https://doi.org/10.1139/b06-047 .

Paradis E. 2018. Analysis of haplotype networks: The randomized minimum spanning tree method. Methods Ecol. Evol., **9**: 1308-1317.

Sharma M., Fomda B. A., Mazta S., Sehgal R., Singh B. B. and Malla N. 2013. Genetic diversity and population genetic structure analysis of *Echinococcus granulosus* sensu stricto complex based on mitochondrial DNA signature. PLoS ONE, **8**(12): e82904. https://doi.org/10.1371/journal.pone.0082904.