Journal of Tropical Biology and Conservation 15:101–120, 2018 ISSN 1823-3902
E-ISSN 2550-1909

Research Article

Genetic Variability and Relationship of Banana Cultivars (*Musa* L.) From East Java, Indonesia based on the Internal Transcribed Spacer Region nrDNA Sequences

Lia Hapsari[1*], Rodiyati Azrianingsih[2], Estri Laras Arumingtyas[2]

[1]Purwodadi Botanic Garden, Indonesian Institute of Sciences, Pasuruan, Jl. Raya Surabaya – Malang Km 65, Purwodadi, Pasuruan, East Java, Indonesia, 67163
[2]Department of Biology, Faculty of Mathematics and Natural Sciences, University of Brawijaya, Jl. Veteran, Lowokwaru, Malang, East Java, Indonesia, 65145
*Corresponding author: hapsari.lia@gmail.co.id; lia.hapsari@lipi.go.id

Abstract

A genetic variability and relationship study subjected to 41 banana cultivars from East Java based on nrDNA sequences of the ITS region was conducted. It would be useful to consider the genomic identification, genetic conservation strategy, and for further banana improvement. ITSL and ITS4 primers were used to amplify the ITS region. Results show that the ITS region DNA sequences length of 41 banana cultivars examined varied from 631 bp to 651 bp. It showed high variability with conservation level G+C content of 62.79%. Total aligned and selected ITS region DNA sequences was 656 bp comprising 346 positions (52.74%) as conserved region, 223 positions (33.99%) as variable sites (polymorphic) and 87 positions (13.26%) alignment gaps. About 143 positions (64.13%) of the variable positions were potentially parsimony informative and 80 positions (35.87%) were singleton variables. The singletone variation sequences spesific to certain banana cultivars may be proposed as identification barcodes. The haplotype diversity was very rich (Hd=1.00), resulted 41 haplotypes with none of haplogroup. Haplotype distribution map revealed the lineage pattern of banana cultivars from East Java. They were presummably derived from common ancestors and the same population in East Java mainland which then experienced an evolution process, dispersed by human migration both in and out, and became isolated to the islands. Genetic relationship reconstruction using NJ algorithm resulted in a tree and classification better than MP algorithm. It was clustered according to its geomic group, into 3 main clades *i.e.* AA/AAA, AAB and ABB. The ITS region nrDNA sequences was proven powerful in classifying until cultivar level of bananas. All 41 banana cultivars examined are recommended for genetics conservation.

Keywords: banana cultivar, East Java, genetic diversity, Internal Transcribed Spacer, molecular marker

Received 17 July 2017
Reviewed 09 April 2018
Accepted 18 July 2018
Published 15 October 2018

Introduction

The diversity of bananas (*Musa*ceae) in Indonesia is very high, being part of the primary origin and diversity centre for both wild seeded species and edible cultivated varieties (Nasution & Yamada, 2001). Wild seeded bananas are not economically utilized much but are linked to their role as genetic resources to improve banana quality in the future (Nasution, 1991). Meanwhile, edible banana is a popular fruit plant worldwide; and it has importance in food security (Hapsari, 2011a), socio-economic and cultural values particularly in rural communities and in developing countries (Megia, 2005; Suhartanto et al., 2009; Hapsari et al., 2017). Edible banana cultivars contain high nutrient values (high carbohydrates, total sugar, vitamin C and potassium; moderate protein and low fat); it is a recommended food for people of all ages (Hapsari & Lestari, 2016). There are approximately at least 325 cultivars recognized in Southeast Asia (Valmayor et al, 2002), of which about 200 cultivars are available in Indonesia (Nasution & Yamada, 2001); and not less than 90 cultivars reported in East Java (Hapsari et al., 2015a; Hapsari et al., 2017). All those local banana cultivars are valuable resources with their own potential characteristics which are necessary to be conserved for further use.

Banana cultivars were putatively derived from natural hybridization between wild diploid *Musa* species. *Musa acuminata* (A genome, x=11) and *Musa balbisiana* (B genome, x=11) are believed to be the ancestors of most banana cultivars (Simmonds & Shepherd, 1955); also *Musa schizocarpa* (S genome, x=11) and *Musa textilis* (T genome, x=11) in a few of them (Simmonds, 1959; Singh et al., 2001). Elucidating the phylogenetic and domestication of bananas is important to provide valuable information for further banana improvement. Studies on phylogeny and domestication of bananas revealed that current banana cultivars are the result of evolution through hybridization, mutation, domestication and adaptation that occurred over thousands of years; the processess are very complex involving multiple stages and separated by time and places. It led to great phenotype and genotype variation amongst cultivars in the region *e.g.* AA, AAA, AAB, ABB, BB, AT, AS, *etc.* (Carreel et al., 2002; De Langhe et al., 2009; Hřibová et al., 2011).

Morphological characteristics have long and often been used in many diversity studies of various organisms for classification, identification, description, characterization and to reconstruct relationships. In addition to morphological characters, various molecular techniques based on PCR by utilizing the deoxyribo nucleic acid (DNA) data has also been conducted to confirm and support the morphology results (De Jesus et al., 2013; Zulfahmi, 2013).

Molecular marker represents the presence of the nucleotide sequences which encodes a trait or specific characteristic. It provides information about the existence of conserved sequences in the genome which can be used in evolutionary biology to study how different organisms are related and how they evolved. To date, optimization and assortment of molecular markers has been developed to visualize the DNA polymorphism amongst species until sub species level (Chase et al., 1993; Hidayat & Pancoro, 2008; Ubaidillah & Sutrisno, 2009).

Source of DNA can be obtained from nuclear, chloroplasts, and mitochondrial genomes. Nuclear genome that is often used for genetic analysis is ribosomal DNA (rDNA). Ribosomal DNA is the coding region of the genome RNA component of the ribosome. It has three types of units including 18S (Small Sub-Unit), 5.8S and 26S (Large Sub Unit) that is encoded by a single transcription and separated into units by an internal transcribed spacer (ITS) (Baldwin et al., 1995; Vanderpoorten et al., 2006). The ITS region as a non-coding region has mild functional problems therefore its evolution occured in a more neutral and natural manner (Álvarez & Wendel, 2003). ITS region is easy to be amplified using universal primers, has high sensitivity due to its small size (300-800 bp) and high copy number in the genome (100-200 copies) (Baldwin et al.,1995). It has proven to be a useful source of characters for genetic variability and phylogenetic studies in many angiosperm families (Baldwin et al.,1995; Álvarez & Wendel, 2003); including Musaceae (De Jesus et al., 2013; Nwakanma et al., 2003; Irishs et al., 2009; Li et al, 2010; Liu et al, 2010; Ravishankar, et al., 2011; Hřibová et al., 2011; Ekasari et al., 2012; Jingyi et al., 2013; Sulistyaningsih et al., 2014)

The present study aims to analyze and evaluate the genetic variability and relationship amongst banana cultivars from East Java based on sequences of ITS region nrDNA which contains entire ITS region (ITS1 and ITS2) including the intervening 5.8S sub-unit (ITS1-5.8S-ITS2). The natural landscape of East Java Province is interesting, it is fragmented into several islands, including the mainland and some small/very small islands. Better knowledge and understanding of genetic variability and relationship of bananas from East Java will be useful to consider its genetic conservation strategy and for further banana cultivar improvement. Genetic variability is the basic material for adaptation, flexibility and responsiveness of evolution in facing various pressures including environmental stresses due to climate change, new pests or diseases, pollution, habitat destructions and fragmentation, and/or other ever-changing environmental conditions (Sumarno & Zuraida, 2008).

Nowadays, proper documentation of genetic variability of biological resources through bio-informatic tools are important in supporting both *in-situ* and *ex-situ* conservation. Storage of genetic information through a centralized GenBank DNA database will allow biodiversity data to be preserved and provide intellectual property protection and establish commercial benefits to owners of biological resources. This study also aimed to fill in the gaps of bio-informatics data of local banana cultivars from East Java, Indonesia.

Material and Methods

Plant materials

A total of 41 banana cultivars originating from East Java were sampled from the living collections at Purwodadi Botanic Garden – Indonesian Institute of Sciences, Pasuruan, East Java. The locality of banana cultivars from areas covering 15 regencies in East Java Province (Figure 1). It represented four genomic groups *i.e.* AA, AAA, AAB and ABB (Table 1). The leaves sample was selected from one individual per cultivar.

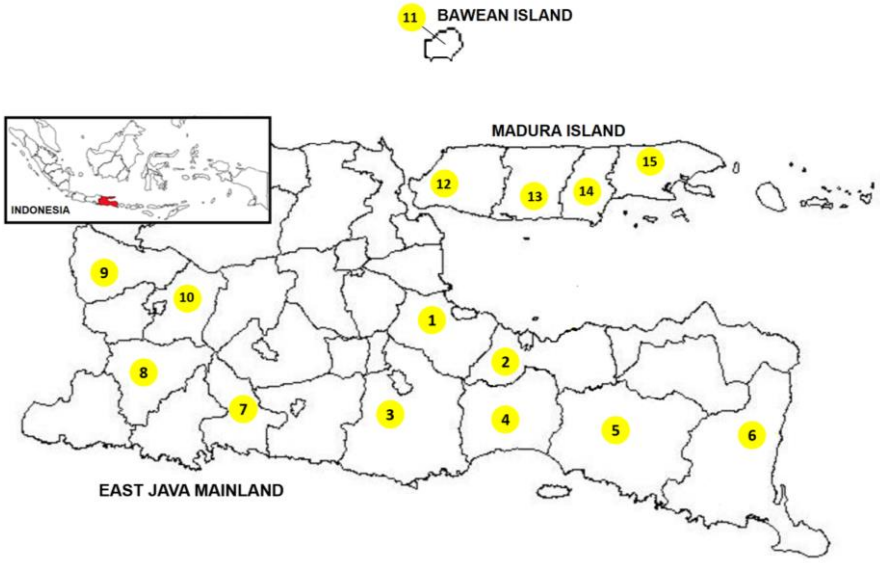

Figure 1. Map of banana cultivars locality from East Java Province: **1.** Pasuruan, **2.** Probolinggo, **3.** Malang, **4.** Lumajang, **5.** Jember, **6.** Banyuwangi, **7.** Tulungagung, **8.** Ponorogo, **9.** Ngawi, **10.** Madiun, **11.** Sangkapura (Bawean Island), **12.** Bangkalan (Madura Island), **13.** Sampang (Madura Island), **14.** Pamekasan (Madura Island), **15.** Sumenep (Madura Island).

Table 1. List of 41 banana cultivars from East Java examined

No.	Banana cultivars	Genomic group	Locality	Locality code
1	Sri	ABB	Pasuruan	PAS
2	Kates	ABB	Tutur, Pasuruan	PAS
3	Raja Prentel	ABB	Nongkojajar, Pasuruan	PAS
4	Sobo Londo	ABB	Purwodadi, Pasuruan	PAS
5	Sobo Awu	ABB	Purwodadi, Pasuruan	PAS
6	Raja Wesi	ABB	Ngawi	NGWI
7	Ebung	ABB	Siman, Ponorogo	PNRG
8	Gajih Bali	ABB	Kedungjajang, Jember	JBR
9	Susu Gabug	ABB	Sangkapura, Bawean Island, Gresik	BWN
10	Kusta Putih	ABB	Batuputih, Sumenep, Madura Island	SMNP
11	Sabeh Biru	ABB	Camplong, Sampang, Madura Island	SMPNG
12	Belindang	AAB	Batuputih, Sumenep, Madura Island	SMNP
13	Ongkap	AAB	Karang Penang, Sampang, Madura Island	SMPNG
14	Dokare	AA	Purwosari, Pasuruan	PAS
15	Jambe	AA	Tulungagung	TLGAG
16	Grito	AAA	Krucil, Probolinggo	PROB
17	Lilin	AA	Kabat, Banyuwangi	BWI
18	Satroli	AA	Krucil, Probolinggo	PROB
19	Raja Sri	AA	Glagah, Banyuwangi	BWI
20	Mas	AA	Kalibaru, Banyuwangi	BWI
21	Mas Kripik	AA	Senduro, Lumajang	LMJG
22	Gading	AA	Tragah, Bangkalan, Madura Island	BGKLN
23	Mas	AA	Galis, Bangkalan, Madura Island	BGKLN
24	Masang	AA	Tragah, Bangkalan, Madura Island	BGKLN
25	Madu	AA	Batuputih, Sumenep, Madura Island	SMNP
26	Pakak Merah	AA	Omben, Sampang, Madura Island	SMPNG
27	Kayu	AAA	Nongkojajar, Pasuruan	PAS
28	Santen	AAA	Nongkojajar, Pasuruan	PAS
29	Nangka	AAA	Purwodadi, Pasuruan	PAS
30	Agung	AAA	Krucil, Probolinggo	PROB
31	Kongkong	AAA	Lawang, Malang	MLG
32	Morosebo	AAA	Sarangan, Madiun	MDUN
33	Kidang	AAA	Kalisat, Jember	JBR
34	Embug	AAA	Ledokombo, Jember	JBR
35	Celakat	AAA	Glagah, Banyuwangi	BWI
36	Lumut	AAA	Batuputih, Sumenep, Madura Island	SMNP
37	Jabol	AAA	Lenteng, Sumenep, Madura Island	SMNP
38	Rosok	AAA	Arosbaya, Bangkalan, Madura Island	BGKLN
39	Nangkah	AAA	Waru, Pamekasan, Madura Island	PMKSN
40	Elang	AAA	Pegantenan, Pamekasan, Madura Island	PMKSN
41	Pakak Santen	AAA	Kadur, Pamekasan, Madura Island	PMKSN

Genomic group reference: Hapsari et al., 2015

Molecular analysis

Total genomic DNAs were extracted from fresh cigar leaf tissues using Promega Wizard® Genomic DNA Purification Kit (Madison, Wisconsin, USA) following the manufacturer's protocols. Amplification of the ITS region (ITS1 + 5.8S + ITS2) was accomplished using primer pairs of ITSL (5'-TCG TAA CAA GGT TTC CGT AGG TG-3') and ITS4 (5'-TCC TCC GCT TAT TGA TAT GC-3') (White et al., 1990; Hsiao et al., 1994; Nwakanma et al., 2003). PCR reactions were conducted in a 30 µL volume contains of 15 µL of DreamTaq Green PCR Master Mix (2x) from Thermo Scientific, California, USA (Taq DNA polymerase, 2x DreamTaq Green buffer, 0.4 mM each of dNTPs and 4 mM $MgCl_2$), 3 µl of 5 pmol each of forward and reverse primers and 3 µl of nuclease-free water. PCR thermal cycling program used for ITS amplification consists of initial denaturation temperature at 95 °C for 3 minutes; followed by 25 cycles of denaturation for 30 seconds at 95 °C, annealing for 30 seconds at 53 ºC, and extension for 30 seconds at 72 ºC. Final extension carried out for 7 minutes at 72 ºC. Amplified products were then purified and sequenced at 1st BASE Laboratories Sdn Bhd, Malaysia using ABI PRISM 3730xl Genetic Analyzer developed by Applied Biosystems, USA.

Data analysis

ITS region DNA sequences data were evaluated using ABI sequences Scanner v.10. Multiple sequences alignments were performed using ClustalW program followed by visual adjustment (MEGA5.03 software). We used unrooted genetic relationship to reconstruct both the network and trees. Unrooted trees (no outgroup) are useful (and informative) to establish the conservancy and the variability among a set of sequences, also to draw a network of relationships among taxonomical units, in accordance with the purpose of this study. In addition, it is also not necessary to root a haplotype network by an outgroup. Outgroups provide the connection of the ingroup populations (species, genera, *etc.*) to a broader phylogeny. However, the relationships among the haplotype network will not change (Olalde et al., 2002). Genetic variability including nucleotide mutations were analyzed with DnaSP ver. 5.10.01. Median Joining analysis was employed using Haplotype Network 4.6.1.2. to reconstruct haplotype distribution map. Genetic relationship reconstructions were performed using MEGA5.03 based on evolution model of Kimura 2 parameter (K2P) using Neighbor Joining (NJ) and Maximum Parsimony (MP) algorithms with 1000 bootsrap replications; a pairwise distance analysis was also performed to generate genetic distances (Tamura et al., 2011; Seltman et al., 2003). Bootsrap support was categorized as strong (>85%), moderate (70-85%), weak (50-69%), and poor (<50%) (Kress et al., 2002).

Results and Discussion

ITS region amplifications & DNA sequences

Amplification of the ITS region using ITSL and ITS4 primers were successfully carried out on the 41 banana cultivars examined. Visualization on 1.5% agarose gel electrophoresis was shown by the presence of a specific DNA band in the sample lane at the length of 600 bp to 700 bp (Figure 2). DNA sequences length of ITS region in Angiosperms varied between 400 bp to 800 bp (Baldwin et al., 1995). The amplicon size of banana cultivars were similar to monocot plants such as rice, sorghum and wheat *i.e.* 591 bp, 588 bp and 603 bp respectively (Nwakanma et al., 2003; Hsiao et al., 1994).

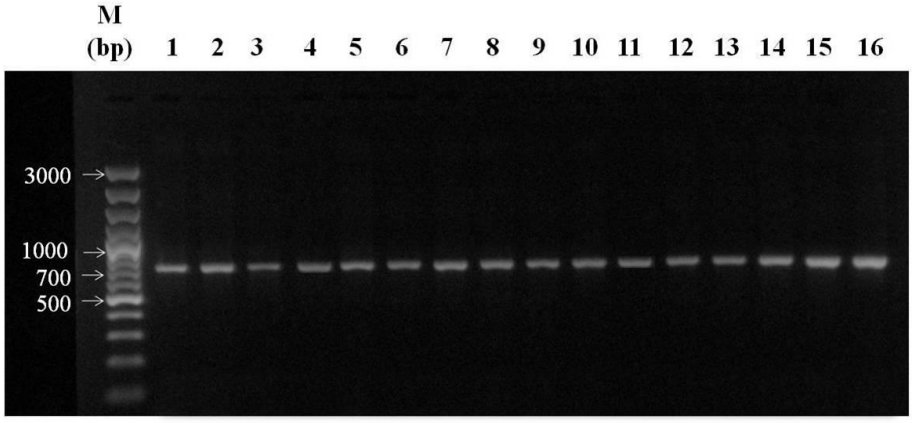

Figure 2. Electrophoregram amplicons of ITS region of some banana cultivars examined on agarose gel 1,5 %. M= marker. Lane 1= Susu Gabug, 2=Madu, 3=Pakak Santen, 4=Kusta Putih, 5=Belindang, 6=Lumut, 7=Jabol, 8=Nangkah, 9=Elang, 10=Sabeh Biru, 11=Pakak Merah, 12=Ongkap, 13=Rosok, 14=Masang, 15=Gading, 16=Mas.

Sequencing on ITS region amplicons of 41 banana cultivars from this study produced DNA sequences with sizes of 631 bp to 651 bp (Table 2). The ITS sequence size of banana cultivars was in accordance with the ITS sequence size of wild bananas which ranged 599 bp to 697 bp (Sulistyaningsih et al., 2014). Based on Basic Local Alignment Search Tool (BLAST) on NCBI, all data DNA sequences of 41 banana cultivars were homologues with ITS region in Musaceae family (similarity ≥92%). It was not found with contaminant of endophytic fungi. All 41 banana cultivars ITS sequences were subjected to genetic variability analysis and relationship reconstruction.

Table 2. Nucleotide composition of 41 banana cultivars from East Java

No.	Banana cultivars	Total sequences	Base (%)			
			T(U)	C	A	G
1	Sri (ABB) PAS	650	17,4	29,2	20,2	33,2
2	Kates (ABB) PAS	645	17,5	30,1	20,0	32,4
3	Raja Prentel (ABB) PAS	650	17,2	29,2	19,5	34,0
4	Sobo Londo (ABB) PAS	643	17,4	29,4	19,0	34,2
5	Sobo Awu (ABB) PAS	648	17,4	29,5	19,9	33,2
6	Raja Wesi (ABB) NGWI	650	18,5	29,4	20,2	32,0
7	Ebung (ABB) PNRG	649	18,0	30,7	21,0	30,4
8	Gajih Bali (ABB) JBR	650	17,8	29,2	20,5	32,5
9	Susu Gabug (ABB) BWN	648	17,1	29,0	19,8	34,1
10	Kusta Putih (ABB) SMNP	648	17,1	29,2	19,3	34,4
11	Sabeh Biru (ABB) SMPNG	648	17,1	29,3	19,3	34,3
12	Belindang (AAB) SMNP	651	16,9	29,3	20,4	33,3
13	Ongkap (AAB) SMPNG	650	16,8	29,2	20,3	33,7
14	Dokare (AA) PAS	649	17,4	29,0	20,3	33,3
15	Jambe (AA) TLGAG	640	17,2	29,7	20,6	32,5
16	Grito (AAA) PROB	650	15,7	30,0	20,9	33,4
17	Lilin (AA) BWI	650	17,2	28,6	19,5	34,6
18	Satroli (AA) PROB	645	18,3	30,1	18,4	33,2
19	Raja Sri (AA) BWI	649	18,0	29,0	19,6	33,4
20	Mas (AA) BWI	649	17,1	29,6	19,4	33,9
21	Mas Kripik (AA) LMJG	648	17,6	29,8	19,1	33,5
22	Gading (AA) BGKLN	650	17,8	28,5	20,2	33,5
23	Mas (AA) BGKLN	650	17,1	29,4	19,7	33,8
24	Masang (AA) BGKLN	631	17,1	29,3	19,2	34,4
25	Madu (AA) SMNP	647	17,0	29,5	19,9	33,5
26	Pakak Merah (AA) SMPNG	643	17,4	28,1	20,5	33,9
27	Kayu (AAA) PAS	646	16,4	30,5	20,9	32,2
28	Santen (AAA) PAS	647	16,8	29,2	21,6	32,3
29	Nangka (AAA) PAS	647	18,1	29,1	21,2	31,7
30	Agung (AAA) PROB	645	17,7	30,5	19,4	32,4
31	Kongkong (AAA) MLG	648	14,8	29,8	20,4	35,0
32	Morosebo (AAA) MDUN	643	15,6	29,7	20,4	34,4
33	Kidang (AAA) JBR	651	17,4	29,8	20,0	32,9
34	Embug (AAA) JBR	640	16,6	29,8	19,2	34,4
35	Celakat (AAA) BWI	651	16,6	30,1	20,3	33,0
36	Lumut (AAA) SMNP	646	18,1	29,4	22,9	29,6
37	Jabol (AAA) SMNP	649	16,8	31,1	20,2	31,9
38	Rosok (AAA) BGKLAN	630	15,7	29,4	20,6	34,3
39	Nangkah (AAA) PMSKN	640	16,4	29,1	20,3	34,2
40	Elang (AAA) PMSKN	643	16,3	30,3	19,1	34,2
41	Pakak Santen (AAA) PMKSN	648	16,5	29,8	21,0	32,7
	Average	646,5	17,1	29,5	20,1	33,3

Notes: G = Guanin, A= Adenin, C= Cytosin, T= Thymin, U= Uracil

Genetic variability

The total aligned and selected ITS region DNA sequences length of 41 banana cultivars was 656 bp positions. Of those, 346 positions (52.74%) were identified as conserved region (invariable/monomorphic), 223 positions (33.99%) were potential variable sites (polymorphic) and 87 positions (13.26%) were alignment gaps or missing data. About 143 positions (64.13%) of the variable positions were potentially parsimony informative and 80 positions (35.87%) were singleton variables. The ITS region DNA sequences of 41 banana cultivars showed high variability with neutral conservation level so that become valuable characters to reconstruct genetic relationships.

The nucleotide composition of ITS region in 41 banana cultivars were high in G+C bases content with an average of 62.62% in *Musa* ABB group, 62.80% in *Musa* AAB group, 62.89% in *Musa* AA group and 62.85% in *Musa* AAA group. At average, G+C bases content of 41 banana cultivars was 62.79% (Table 2). ITS region as non-coding region or intron was known to have high G+C content because it was associated with their relative functions in transcription and translation. DNA sequences with higher G+C content are hotspots of mutation, C base is often methylated and occured errors during multiplication. DNA methylation is an addition reaction of a methyl group at the 5' end of C base covalently. Methylation was induced substitution mutation in the form of transition *i.e.* C base to T/U (Ubaidillah & Sutrisno, 2009; Nusifera, 2007). Methylation events and/or the addition of other alkyls in the DNA are the major causes of mutations in many organisms.

DNA sequences with singletone variable positions is DNA sequence variation where point mutation occured only in one operating taxonomic unit (OTU) (Hidayat & Pancoro, 2008). Mutation events analysis in singleton variable positions revealed that variation of ITS sequences amongst banana cultivars in this study were mostly contributed by point mutations in form of substitution (Table 3). About 73 positions of the singleton variables comprised of two variants point mutations and in 7 positions there were three variants point mutations. The subtitution events were 51 sites in form of transversion *i.e.* the substitution of a (two rings) purine for a (one ring) pyrimidine or vice versa; and 49 sites in form of transition *i.e.* a point mutation that changes a purine nucleotide to another purine (A ↔ G) or a pyrimidine nucleotide to another pyrimidine (C ↔ T). About 8 mutation events from C to T occured pressumable due to methylation (Table 3). It was reported that only minor proportion of indels affected to the variation of ITS sequences in Angiosperms (Baldwin et al., 1995).

Table 3. Singletone variations of ITS region DNA sequences of 26 banana cultivars

No.	Banana cultivars	Base position	Base mutation	Total
1	Lumut (AAA) SMNP	92, 101, 104, 108, 113, 116, 136, 168, 178,218, 133, 316, 418, 329	C → T, G → A, G → A, G → A, G → A, G → A, C → T, G → A, G → A, G → C, G → T, G → T, A → G	14
2	Elang (AAA) PMSKN	358, 530, 593, 537, 638, 543, 563, 545, 565, 462, 474, 556, 596	C → G, C → A, C → A, C → T, C → T, G → A, G → A, A → C, A → C, T → G, G → C, G → C, A → G	13
3	Nangka (AAA) PAS	243, 244, 341, 429, 245, 377, 269, 281, 362	G → T, G → C, G → C, G → C, C → T, C → T, C → A, C → G	9
4	Santen (AAA) PAS	536, 574, 564, 573, 598, 593, 596, 638	G → T, G → T, G → A, C → G, C → G, C → A, A → G, C → T	8
5	Kates (ABB) PAS	1, 6, 7, 264, 311, 530	T → C, G → A, A → G, C → A, A → C, C → T	6
6	Masang (AA) BGKLN	471, 630, 590, 624, 634, 636	A → G, A → G, C → T, C → T, T → A, T → G	6
7	Ebung (ABB) PNRG	108, 243, 110, 170, 270	G → C, G → C, G → A, G → A, G → A	5
8	Satroli (AA) PROB	218, 304, 410, 642, 647	G → T, G → C, G → C, T → C, A → G	5
9	Pakak Merah (AA) SMPNG	31, 32, 223, 379	G → T, A → G, C → T, C → T	4
10	Gajih Bali (ABB) JBR	130, 151, 157, 265	G → C, C → A, C → A, G → A	4
11	Raja Wesi (ABB) NGWI	200, 309, 464	C → T, C → A, G → C	3
12	Rosok (AAA) BGKLAN	481, 516, 572	C → G, T → C, T → C	3
13	Kongkong (AAA) MLG	499, 640, 642	A → G, T → A, T → A	3
14	Susu Gabug (ABB) BWN	14, 63	C → T, T → A	2
15	Morosebo (AAA) MDUN	82, 341	T → G, G → A	2
16	Jabol (AAA) SMNP	205, 234	G → C, G → C	2
17	Agung (AAA) PROB	482, 635	G → T, A → T	2
18	Nangkah (AAA) PMSKN	537, 620	C → G, C → T	2
19	Kidang (AAA) JBR	266	T → A	1
20	Embug (AAA) JBR	288	T → G	1
21	Lilin (AA) BWI	353	A → G	1
22	Raja Sri (AA) BWI	383	C → T	1
23	Mas (AA) BWI	569	A → C	1
24	Kayu (AAA) PAS	460	G → T	1
25	Dokare (AA) PAS	465	C → G	1
26	Jambe (AA) TLGAG	555	G → A	1

Notes: G = Guanin, A= Adenin, C= Cytosin, T= Thymin

Singletone variations were found in 26 out of 41 banana cultivars ITS sequences. Those singletone variation sequences are presented in Table 3. It may be proposed as identification barcodes for each particular banana cultivar. Pisang Lumut is a cultivar with the highest singletone variation events (14 events) followed by Pisang Elang (13 events). Interestingly, they both originate from Madura Island but in different regions, *i.e.* Sumenep and Pamekasan respectively. In addition, Pisang Nangka (Pasuruan), Pisang Santen (Pasuruan), Pisang Kates (Pasuruan) and Pisang Masang (Bangkalan, Madura) also have high singletone variations with mutation events of 9, 8, 6, and 6 respectively (Table 3).

Mutations are essential to evolution. Mutations can be caused by high-energy sources such as radiation, chemicals, high pressures of environment, or appear spontaneously during DNA replication, *etc*. Beneficial mutations allow an organism to reproduce more effectively and adapt well to a changing environment (Carlin, 2011). Banana cultivar is asexually reproductive by corms, therefore it can maintain the mutated genetic constituent. However, domestication, human selection and migration also add evolution aspects in banana cultivars. Therefore, It is presumed that banana cultivars with high singletone variables found in this study may be carrying adaptive alleles to their specific habitats, and resulted variation. Particularly in Madura Island, which has a dry ecotype in tropical weather, limy land, low rainfall, and low soil productivity (Rochana, 2012).

Haplotype diversity and distribution

Haplotype is a set of specific DNA sequences in the cluster associated with a particular gene on a chromosome that is most likely inherited together or comes from a the common ancestor (Seltman et al., 2003). Haplotype network represent evolutionary relationships (genealogies) to intraspesific level (Mardulyn, 2012). Haplotype (gene) diversity based on ITS region DNA sequences of 41 banana cultivars from East Java was very high (Hd = 1.00) with variance 0,00003 and standard deviation 0,005. It resulted in 41 haplotypes with none of the haplogroup (Figure 3). This study revealed the lineage pattern of banana cultivars from East Java Province including East Java mainland, Bawean Island and Madura Island; it was inter-connected and clustered according to its genomic group into 3 clusters.

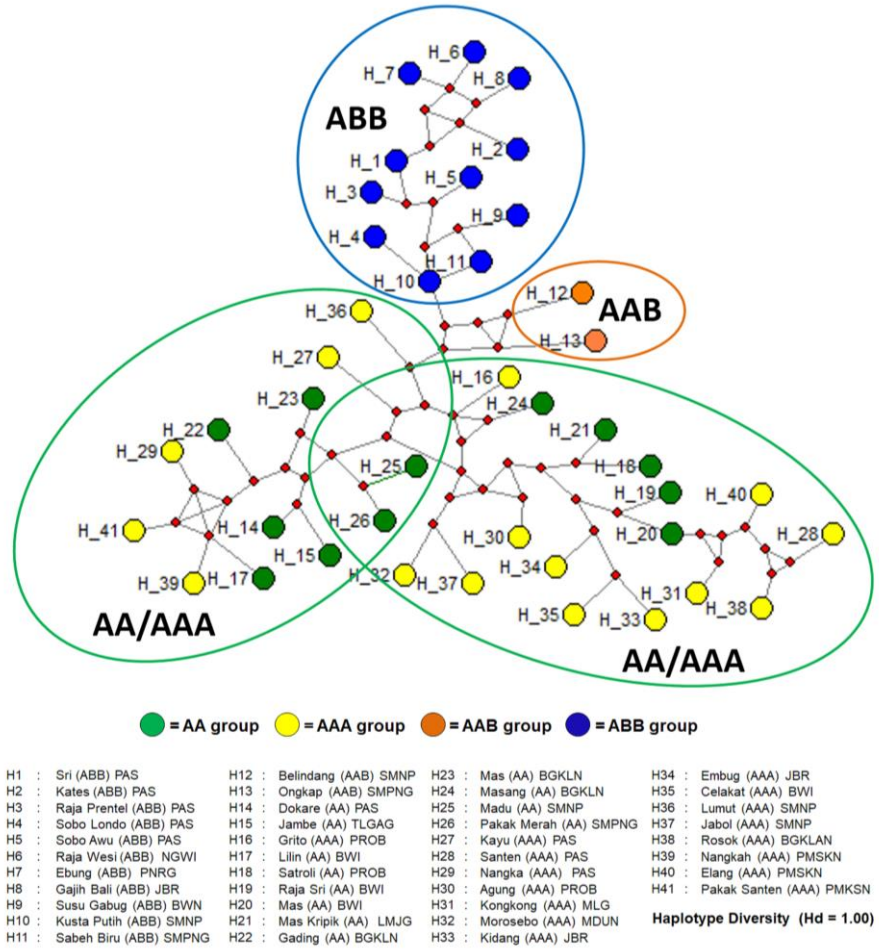

Figure 3. Haplotype distribution map of 41 banana cultivars from East Java based on ITS region DNA sequences.

Geologically, East Java mainland also the islands of Madura and Bawean (the study areas) are included as Great Sunda Island plate/ Sundaland (Verbeek & Fennema, 1896; Van Bemellen, 1949; Usman, 2012). Madura Island is a continuation of the Solo limestone mountains which is now separated by Madura Strait. Whereas Bawean Island is formed from the remains of an old volcano located near its centre which is now separated by Java Island (Verbeek & Fennema, 1896; Rochana 2012; Figure 1). Island biogeography is the study of the factors affecting species diversity of natural communities. According to the island biogeography theory by MacArthur & Wilson (1967), the

species diversity on the island was determined by the island area, also equilibrium numbers between the average rate of local extinction with the rate of migration to the islands as well as the island's isolation level. Since then, all 41 banana cultivars examined were presummably derived from common ancestors and same population in East Java mainland which then experienced evolution process, dispersed by human migration both in and out, and got isolated to the islands consequently resulting in rich haplotype diversity (Figure 3).

The haplotype distribution map showed that banana cultivars of AA were nested together with AAA group. Furthermore, the network of banana cultivars of AA and AAA groups were directly connected with banana cultivars of AAB group, and AAB group was directly connected with banana cultivars of ABB group. Hence, the AAB group bananas served as intermediate bananas which connect both B genome bananas and A genome bananas. These findings support previous studies by De Langhe et al. (2009) that during evolution, banana cultivars may first exist from clone selection of wild cultivated population and then in regions where the diffusion of plants occurred (exchange or via human migration) hybridization between wild cultivated populations and partly fertile clones from different origins led to the generation of more sterile AA diploids, and the more vigorous and nearly sterile triploids. The *Musa* AA group (edible diploid) emerged first, followed by *Musa* AAA group, *Musa* AAB group and later on *Musa* ABB group. Selection among those new diploid and triploid populations produced new cultivars. Those new cultivars can undergo somatic mutation, thus leading to new derived cultivars.

According to the haplotype network map, there are two main sub-groups of *Musa* AA group *i.e.* Sub-group 1 consist of Pisang Lilin, Gading, Dokare, Jambe, Mas Bangkalan, Madu and Pakak Merah; and Sub-group 2 consists of Pisang Mas Kripik, Raja Sri, Satroli, Mas Banyuwangi, Grito and Masang (Figure 3). Pisang Masang appeared as a derived cultivar of Pisang Mas. In addition, Pisang Mas from Bangkalan (Madura Island) was separated into a different sub-group with Pisang Mas from Banyuwangi (East Java Mainland); it is indicated that they both are genetically different due to variation although they have same name.

Musa AAA group consists of two main sub-groups *i.e.* Sub-group 1 comprises of Pisang Nangka, Nangkah, Kayu, Lumut, Morosebo and Jabol; and Sub-group 2 comprises of Pisang Celakat, Kidang, Embug, Agung, Elang, Rosok, Santen and Kongkong (Figure 3). Pisang Nangka from Pasuruan were placed in the same

sub-group and are directly connected with Pisang Nangkah from Pamekasan, Madura Island; it is indicated that they both have the same lineage, are separated by distance, but experienced less of variation. Meanwhile, Musa AAB group number was very limited to only 2 cultivars *i.e.* Pisang Belindang and Ongkap, both originating from Madura Island (Figure 3). Based on previous studies it was known that many of bananas morphologically identified as AAB, once confirmed molecularly were identified as AAA (Hapsari et al. 2015b). In addition, all banana cultivars of ABB were consistantly clustered in one sub-group, in which Pisang Kusta Putih appeared become the basic cultivar of banana ABB group studied because all of other ABB bananas were connected to it (Figure 3).

Genetic relationship

ITS region is bi-parental inherited and is more informative in revealing the evolutionary history of organisms that naturally involves crossing both inter and intra species, also occurrence of polyploidy and hybrid speciation events like in Musaceae family (Alvarez & Wendel, 2003). This study has proven that ITS region nrDNA sequences were powerful to classify to the cultivars level of bananas.

NJ algorithm analyze the data by converting the DNA sequences into distance matrix based on differences in pairs between sequences while MP algorithm using sequences character directly in which total number of character-state changes is to be preferred (Ubaidillah & Sutrisno, 2009; Tamura et al., 2011; Hall, 2008). Based on the trees resulting from this study, both algorithms can be used to explain the relationship of banana cultivars from East Java. They both resulted in trees which were separated and clustered according to genomic group. However, comparatively NJ algorithm constructed a genetic relationship tree that was better than MP algorithm in classifying and clustering of bananas from East Java (Figure 4).

Genetic relationship analysis of banana cultivars from East Java based on ITS sequences using NJ algorithm resulted in a tree which separated into 3 main clades in accordance to its genomic group classification. The first clade consists of cultivars of ABB group, and it became the outgroup and supported by strong bootstrap. The second clade consists of 2 cultivars of AAB group supported by moderate bootsrap. Third group consists of cultivars of AA and AAA groups, and supported by low bootstrap (Figure 4A). The relationship tree using MP algorithm was separated only into 2 main clades, in which the AAB group is nested in the second clade together with AA and AAA groups. The

separation of the second clade (AA, AAA and AAB groups) is supported by strong bootsrap but low bootrap support in its sub-clades separation (Figure 4B).

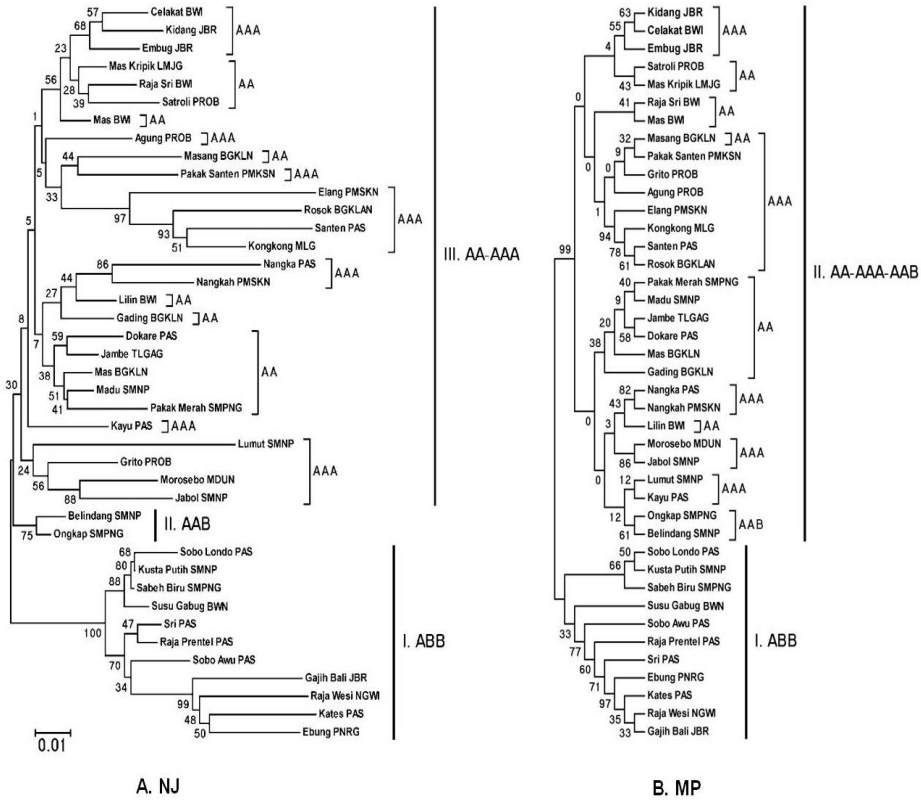

Figure 4. Relationship trees of 41 banana cultivars from East Java based on ITS region nrDNA sequences: **A)** Neighbor Joining (NJ) and **B)** Maximum Parsimony (MP)

The ABB clade was separated into 2 sub-clades according to its genetic similarity (Figure 4). Genetic similarity amongst members of Musa ABB group were 90.65% to 99.82%, in which Pisang Sabeh Biru x Kusta Putih was the closest pair while Pisang Gajih Putih x Sobo Londo was the the farthest pair. Banana cultivars with A genome were not clearly separated, *Musa* AA group

was nested together with *Musa* AAA group. However, their relationship tree had tended to separated into two main sub-groups similar to haplotype's grouping (Figure 4). Genetic similarity of AA/AAA group were 88.20% to 97.50%, in which Pisang Embug (AAA) and Mas Kripik (AA) was the closest pair whereas Pisang Santen (AAA) and Pakak Merah (AA) was the farthest pair. Pisang Belindang and Ongkap as AAB clade was related closely to AA/AAA clade with genetic similarity of 88.63% to 97.31% compared to the ABB group with genetic similarity of 89.23% to 95.28%.

The assessment of genetic variation and distance provides important information for conservation on a genetic basis. In terms of conservation genetics, it is focused mainly on the protection and maintanance of genetic variation for further evolvability (Woodruf, 2001). Banana cultivars with high genetic variability and far genetic distance or low similarity are prioritized for conservation. Referring to the genetic similarity value of Pisang Santen (AAA) x Pisang Kates (ABB), these were the farthest pair with genetic similarity 81.51% followed by Pisang Masang (AA) x Pisang Kates (ABB) with genetic similarity of 86.52%. Therefore, those banana cultivars were prioritized for conservation both *in-situ*/*on-farm* and *ex-situ*. In addition, if conservation resources are limited, any banana cultivars which are very closely related (high similarity) should be chosen with one of them as representative. Pisang Sabeh Biru (ABB) x Kusta Putih (ABB) is the closest pair with genetic similarity 99.82%, and are both from Madura Island. However, since they both had different haplotype, it is necessary for these to be conserved.

Conclusions

This study has proven that ITS region nrDNA sequences was useful as a source of characters for genetic variability and phylogenetic studies of bananas. It was powerful to classify until the cultivar level of bananas. Genetic variability of 41 banana cultivars from East Java based on ITS region DNA sequences were very rich. Haplotype analysis resulted in 41 haplotypes with none of haplogroup. Haplotype distribution and genetic relationship analyses revealed the lineage pattern of banana cultivars from East Java. It was connected and clustered according to its genomic group. They were presummably derived from the common ancestors and same population in East Java mainland which then experienced evolution process, dispersed by human migration both in and out, and got isolated to the islands. Genetic relationship analyses resulted in trees which were also clustered according to their geomic group. However, NJ

algorithm method constructed a relationship tree and classification better than MP algorithm in classifying and clustering of bananas from East Java.

ITS sequences data of 41 banana cultivars from East Java from this study were deposited at the National Center for Biotechnology Information (NCBI) with GenBank accessions numbers KT696446 to KT696490. It will allow the genetic diversity of local bananas data to remain preserved and also to provide intellectual property protection. This study recommended that all of the 41 banana cultivars from East Java are necessary for genetics conservation. Conservation efforts both *in-situ* and *ex-situ* to all local *Musa* germplasms are needed to maintain the genetic resources for further banana improvement.

Acknowledgements

This research was funded by the Indonesia Endowment Fund for Education, Ministry of Finance, Republic of Indonesia with Graduate Research Fellowship grant number PRJ-541/LPDP/2013. The authors also greatly acknowledge Purwodadi Botanic Garden, Indonesian Institute of Sciences for providing the plant materials studied and Plant Physiology Laboratory of Biology Department, University of Brawijaya for the molecular and genetic facilities. Sincere thanks to Didik Wahyudi, M.Si. for all technical help and valuable discussions during the study.

References

Álvarez I, Wendel JF. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* **29**: 417-434.

Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ. 1995. The ITS region of nuclear ribosomal DNA: A valuable source of evidence on Angiosperm phylogeny. *Annals of Missouri Botanic Garden* **82**: 247-277.

Carlin JL. 2011. Mutations are the raw materials of evolution. *Nature Education Knowledge* **3(10)**: 10.

Carreel F, Leon DG, Lagoda P, Lanaud C, Jenny C, Horry JP, Montcel TH. 2002. Ascertaining maternal and paternal lineage within *Musa* by chloroplast and mitochondrial DNA RFLP analyses. *Genome* **45(4)**: 679–692.

Chase MW, Soltis DE, Olmstead RG. 1993. Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbc*L. *Annals of Missouri Botanic Garden* **80**: 528-580.

De Jesus ON, de Oliveira e Silva S, Amorim EP, Ferreira CF, de Campos JMS, de Gaspari Silva G, Figueira A. 2013. Genetic diversity and population

structure of *Musa* accessions in *ex-situ* conservation. *Biomed Central Plant Biology* **13**: 41.

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.

Ekasari TWD, Retnoningsih A, Widianti T. 2012. Diversity analysis if banana cultivars using PCR-RFLP of the Internal Transcribed Spacer (ITS) DNA ribosom. *Jurnal MIPA* **35(1)**: 21-30. [Indonesian]

Hall BG. 2008. *Phylogenetic Trees Made Easy: A how to manual*. (3rd ed.). Sunderland, Massachusetts: Sinauer Associates, Inc.

Hapsari L, J Kennedy, DA Lestari, A Masrum, W Lestarini. 2017. Ethnobotanical survey of bananas (*Musa*ceae) in six districts of East Java, Indonesia. *Biodiversitas* **18(1)**: 160-174.

Hapsari L, Lestari DA. 2016. Fruit characteristic and nutrient values of four Indonesian banana cultivars (*Musa* spp.) at different genomic groups. *Agrivita* **38(3)**: 303-311.

Hapsari L, Masrum A, Lestari DA. 2015a. Diversity of bananas (*Musa* spp.) in Madura Island, East Java: exploration and inventory. *Journal of Biodiversity and Environmental Sciences* **6(3)**: 256-264.

Hapsari L, Wahyudi D, Azrianingsih R, Arumingtyas EL. 2015b. Genomic Identification of bananas (*Musa* spp.) from East Java assessed by PCR-RFLP of the Internal Transcribed Spacer ribosomal DNA. **International Journal of Biosciences 7(3)**: 42-52.

Hapsari L. 2011a. *Indonesian banana cultivars Purwodadi Botanic Garden's collection*. In: Proceeding of International Conference on Food Safety & Food Security. Gajah Mada University. Pp.115-119.

Hapsari L. 2011b. Two decades of banana collection (*Musa*ceae) Purwodadi Botanic Garden (1990-2010). *Jurnal Berkala Penelitian Hayati Edisi Khusus* **5A**: 147-151. [Indonesian]

Hidayat T, Pancoro A. 2008. Study of molecular phylogenetic and its role in providing basis information to improve the genetic resources quality of orchid. *Jurnal Agrobiogen* **4(1)**: 35-40. [Indonesian]

Hřibová E, JĆížková J, Christelová P, Taudin S, de Langhe S, Doleżel J. 2011. The ITS1-5.8S-ITS2 sequence region in the Musaceae: structure, diversity and use in molecular phylogeny. *PLoS ONE* **6(3)**: e17863.

Hsiao C, Chatterton NJ, KH Asay, KB Jensen. 1994. Phylogenetic relationships of ten grass species: An assessment of phylogenetic utility of the Internal Transcribed Spacer region in the nuclear ribosomal DNA in monocots. *Genome* **37**: 112–1204.

Irishs BM, Crespo A, Goenaga R, Niedz R, Ayala-Silva T. 2009. Ploidy level and genomic composition of *Musa* spp. accessions at the USDA-ARS Tropical Agriculture Research Station. *The Journal of Agriculture of the University of Puerto Rico* **93(12)**: 1-21.

Jingyi W, Xueting C, Zilong M, Yaoting W. 2013. Analysis of nuclear ribosomal ITS sequences in *Musa* (Musaceae). *Chinese Agricultural Bulletin* **29(25)**: 6-11.

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.

Li LF, Häkkinen M, Yuan Y-M, 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.

Liu AZ,Kress WJ, Li DZ. 2010. Phylogenetic analyses of the banana family (*Musa*ceae) based on nuclear ribosomal (ITS) and chloroplast (*trn*L-F) evidence. *Taxon* **59(1):** 20-28.

MacArthur RH, Wilson EO. 1967. *The Theory of Island Biogeography*. Princeton, New Jersey: Princeton University Press.

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

Megia R. 2005. *Musa* as genomic model. *Hayati* **12(4)**: 167-170. [Indonesian]

Nasution RE, Yamada I. 2001. *Wild bananas in Indonesia*. Bogor: Puslitbang Biologi, Lembaga Ilmu Pengetahuan Indonesia. [Indonesian]

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

Nusifera S. 2007. DNA Methylation and genomic imprinting (A review of epigenetic mechanism). *Jurnal Agronomi* **11(1)**: 51-58. [Indonesian]

Nwakanma C, Pillay M, Okoli BE, Tenkuano A. 2003. PCR-RFLP of the ribosom DNA Internal Transcribed Spacer (ITS) provide markers for the A and B genomes in *Musa* L. *Theoretical and Applied Genetics* **108**: 154-159.

Olalde M, Herrán A, Espinel S, Goicoechea PG. 2002. White oaks phylogeography in the Iberian Peninsula. *Forest Ecology and Management* **156**: 89–102.

Ravishankar KV, Ajitha-Kumar R, Mathiazhagan M, Ambika D-MS. 2011. Apomictic seed development in *Ensete superbum* induced by pollen of wild banana sp. *Musa balbisiana*. *Current Science* **101(4)**: 493-495.

Rochana T. 2012. Madurese people: A review anthropology. *Humanus* **9**:46-51

Seltman H, Roeder K, Devlin B. 2003. Evolutionary-based association analysis using haplotype data. *Genetic epidemiology* **25**: 48-58.

Simmonds NW, Shepherd K. 1955. The taxonomy and origins of the cultivated banana. *Botanical Journal of the Linnean Society* **55**: 302-312.

Simmonds NW. 1959. *Bananas*. New York: Longman Inc.

Singh HP, Uma S, Sathiamoorthy S. 2001. *A tentative key for identification and classification of Indian bananas*. Tiruchirapalli, India: National Research Centre for Banana (NRCB)

Suhartanto MR, Sobir, Harti H, Nasution MA. 2009. *Banana development as the support of national food security*. Presentation presented at Proceedings

of the Research Results Seminar. Bogor: Bogor Agricultural University.

Sulistyaningsih LD, Megia R, Widjaja EA. 2014. Phylogenetical study of wild banana species (*Musa* L.) in Sulawesi inferred from Internal Transcribed Spacer region of nuclear ribosomal DNA sequences. *Biotropia* **21(1)**: 13-24.

Sumarno, Zuraida N. 2008. Management of plant germplasms integrated with breeding program. *Buletin Plasma Nutfah* **14(2)**: 57-67. [Indonesian]

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. *Molecular Biology and Evolution* **28(10)**: 2731-2739.

Ubaidillah R, Sutrisno H. 2009. *Introduction to Biosystematic: theory and practice*. Jakarta: LIPI Press. [Indonesian]

Usman E. 2012. Bawean Island as geology tourism object. *Mineral dan Energi* **10(3)**:95-101 [Indonesian]

Valmayor RV, Jamaluddin SH, Silayoi B, Kusumo S, Danh LD, Pascua OC, Espino RRC. 2000. *Banana cultivar names and synonyms in Southeast Asia*. Los Banos, Laguna, Philippines: International Network for the Improvement of Banana and Plantain (INIBAP) – Asia and the Pacific Office

Van Bemmelen RW. 1949. *The Geology of Indonesia*. (Volume IA, Chapter III). The Hague, The Netherlands: Martinus Nijhoff

Vanderpoorten AL, Goffinet B, Quandt D. 2006. *Utility of the Internal Transcribed Spacers of the 18S-5.8S-26S nuclear ribosomal DNA in land plant systematics with special emphasis on Bryophytes*. In: Sharma AK (ed.). Plant genome: Biodiversity and evolution. (Volume 2, Part B. Lower Groups). Enfield: Science Publishers

Verbeek RDM, Fennema R. 1896. *Java et Madoura*. The Netherlands: JG Stemler cz.

White TJ, Bruns TD, Lee SB, Taylor JW. 1990. *Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics*. In: Innis MA, Gelfrand DH, Snisky JI, White TJ (eds.). PCR protocols: A guide to methods and applications. New York: Academic Press

Woodruff DS. 2001. Populations, species, and genetic conservations. *Encycopedia of Biodiversity* **4**: 811-829.

Zulfahmi. 2013. DNA markers for plant genetic analyses. *Jurnal Agroteknologi* **3(2)**: 41-52. [Indonesian]