Phylogenetic analysis of 23 accessions of Indonesian banana cultivars based on Internal Transcribed Spacer 2 (ITS2) region

Karlia Meitha[1], Intan Fatmawati[1], Fenny Martha Dwivany[*,1,2,4], Agus Sutanto[3], Sigit Nur Pratama[1], Husna Nugrahapraja1 and Ketut Wikantika[2,4]

[1]School of Life Sciences and Technology, Institut Teknologi Bandung, Jl. Ganesha No. 10, Bandung 40132, Indonesia
[2]Bali International Research Center for Banana, Jl. Raya Kampus UNUD, Badung, Bali 80361, Indonesia
[3]Indonesia Tropical Fruit Research Institute, Jl. Raya Solok Aripan KM 8, Solok, Sumatera Barat 27356, Indonesia
[4]Center for Remote Sensing, Institut Teknologi Bandung, Jl. Ganesha No. 10, Bandung, 40132, Indonesia
*Corresponding author: fenny@sith.itb.ac.id

ABSTRACT Pisang Kepok (*Musa* spp. [ABB 'Saba' subgroup]) has several unique characteristics, such as tolerance to drought and *Fusarium* Foc (TR4) disease. Currently, the genetic diversity of Pisang Kepok in Indonesia is not well identified, although it is widely cultivated. Information on genetic diversity is essential for developing breeding strategies to achieve efficient cultivar improvement in the future. Aims of this research were to analyze the genetic variation of Pisang Kepok from some islands in Indonesia and to determine the genetic relationship between Pisang Kepok and other accessions banana cultivars based on ITS2 region, as a basis for future research in improving banana quality through molecular breeding. We have conducted the multiple sequence alignment and built the phylogenetic tree analysis using the Bayesian Inference Phylogeny method of one million generations (ngen = 1,000,000). The ITS2 region showed two clade ingroups: first clade consists of banana with B genome (*balbisiana*), while the second clade consists of banana with only A genome (*acuminata*). In general, all accessions of Pisang Kepok cultivars were clustered in the B genome of bananas cultivars. In addition, the ITS2 sequences and secondary structures among Pisang Kepok from various regions are identical, suggesting that there was no genetic variation in the ITS2 region of Pisang Kepok from multiple areas in Indonesia.

KEYWORDS Pisang Kepok; *Musa* spp. (ABB 'Saba' subgroup); ITS2; phylogenetic; ITS2 secondary structure

1. Introduction

The increasing global population demands a dependable source of food that is highly nutritious, cheap, and easy to cultivate in most developing countries. Banana is considered as a promising crop due to its adaptability in a wide range of climate, which currently is the staple for hundreds of millions in the tropics and subtropics. The highest banana consumption per capita is recorded in Uganda, Rwanda, and Cameroon, exceeding 200 kg of bananas and plantains. In the rural areas of these countries, banana supplies up to 25% of the daily calorie intake. In the Philippines and Brazil, the population consumes around 60 kg/year (FAO 2019) either from local production or imported.

Humans have been cultivating bananas since the emergence of agriculture (Denham et al. 2003), and in 2017, global production reached 144 million tons (FAO 2019). The domestication of the *Musa* genus originated in Southeast Asia, as inferred by the chloroplast and nuclear genes Li et al. (2013), resulting in a large variety of plantains and bananas. Indonesia is considered as the center of origin of a diverse banana species (Simmonds 1962), highly rich in genetic resources of both wild and cultivated bananas Daniells et al. (2001). There are at least 325 recorded *Musa* cultivars in Indonesia (Valmayor et al. 2000), those are widespread in Sumatra, Bali, Nusa Tenggara, Java, Kalimantan, Sulawesi, Maluku, and Papua.

Bananas play multiple roles in the lives of Indonesian, supplying nutrition when consumed as fresh or cooked, and also serving as a key element in their culture, such as for religious ceremonies in Bali (Hidayat et al. 2018). Unfortunately, banana production in Indonesia is sometimes troublesome that decreases both the quality and quantity of harvest due to biotic and abiotic stresses. For many years, the banana plantation in Indonesia has been struck by Tropical Race 4 (TR4) of the *Fusarium* wilt fungus (Molina et al. 2009; Jumjunidang and Soemargono 2012; Maryani et al. 2019). A soil pathogen that attacks vascular tissue and inhibits transport of water and sugars (FAO 2019). The increasing heat and diminishing rainfall in tropics and subtropics also threaten the production of this

TABLE 1 Banana accessions used in this study. They represented AA, AAA, AB, B, and ABB genomes of banana.

No	Local Name	Scientific Name	Sampling Location
1	Pisang hias	*Musa ornata*	Bali
2	Pisang hias	*Musa velutina*	Bali
3	Kepok	*Musa* spp. (ABB subgroup 'Saba')	Bandung, West Java
4	Kepok	*Musa* spp. (ABB subgroup 'Saba')	Bali
5	Kepok	*Musa* spp. (ABB subgroup 'Saba')	Flores, NTT
6	Kepok	*Musa* spp. (ABB subgroup 'Saba')	Sorong, Papua
7	Kepok tanjung	*Musa* spp. (ABB group)	Bali
8	Sobo	*Musa* spp. (ABB group)	Bali
9	Poh	*Musa* spp. (ABB group)	Bali
10	Sabe macan	*Musa* spp. (ABB group)	Bali
11	Siam	*Musa* spp. (ABB group)	Bali
12	Bali	*Musa* spp. (ABB group)	Bali
13	Ambon hijau	*Musa acuminata* (AAA subgroup 'Cavendish')	Bali
14	Ambon kuning	*Musa acuminata* (AAA subgroup 'Gross Michel')	Bali
15	Ambon Lumut	*Musa acuminata* (AAA subgroup 'Cavendish')	Bali
16	Cavendish	*Musa acuminata* (AAA subgroup 'Cavendish')	Bali
17	Tembaga	*Musa acuminata* (AAA subgroup 'Green Red Banana')	Bali
18	Rojo molo	*Musa acuminata* (AAA subgroup 'Green Red Banana')	Bali
19	Mas	*Musa acuminata* (AA subgroup 'Sucrier')	Subang, West Java
20	Mas marlin	*Musa acuminata* (AA subgroup 'Berlin')	Bali
21	Bile	*Musa* spp. (AB group)	Bali
22	Klutuk	*Musa balbisiana* Colla (B group)	Bali
23	Klutuk wulung	*Musa balbisiana* Colla (B group)	Tasikmalaya, West Java

world's popular food crop (Baral 2019). Thus, the selection of banana cultivar with a high tolerance to *Fusarium* and drought is required.

One of the most cultivated bananas in the tropics is Pisang Kepok or *Musa* spp. (ABB) subgroup 'Saba'. In Indonesia, the cultivation is mostly done in traditional farms that are widely dispersed in their archipelago. Saba banana has been long serving as an important source of nutrition for the locals. In addition to that, from the molecular breeding standpoint, Saba banana is a valuable genetic resource for generating cultivars tolerant to drought (Ravi et al. 2013) and/or *Fusarium* wilt disease (Foc TR 1) cultivated in the greenhouse (Rebouças et al. 2018). Pisang Ke-

pok or Saba banana is a triploid of the hybridized result of *Musa balbisiana* (BB genome) and *Musa acuminata* (AA genome) (Ploetz et al. 2007).

Currently, the genetic diversity of Pisang Kepok in Indonesia is not well identified. Genetic diversity information is required in developing breeding strategies to achieve efficient cultivar improvement in the future (Opara et al. 2010). Poerba and Ahmad (2010) identified the genetic variability of cooking bananas and plantains by RAPD and ISSR markers from 36 accessions, with 12 of them were described as local Pisang Kepok. Ahmad et al. (2014) also studied the genetic diversity of *M. balbisiana* Colla in Indonesia based on the AFLP marker. The cultivar is well-known for its resistance to pest, disease, and drought (Morpurgo et al. 1994; Boonruangrod et al. 2008). However, the fruit contains big and plenty of seeds that make it hardly edible, locally it is also known as Pisang Batu or "stone fruit banana".

The DNA marker of Internal Transcribed Spacer 2 (ITS2) is a potential universal marker for plants and animals. It has sufficient variability to identify a broader range of plant taxa and even to distinguish closely related species (Chen et al. 2010; Yao et al. 2010). The success rate of identification, based on ITS2 region, at the species level of *Musa* is up to 82.5% (Yao et al. 2010). This DNA marker also resolves the minimal genetic variation in the East African highland bananas (Němečková et al. 2018) and contributes to revealing the disjunct genome group of *Musa ornata* in Mexico and Asia (Burgos-Hernández et al. 2017). Further, the identification of ITS secondary structure improves discrimination inter-species (Hřibová et al. 2011; Zhang et al. 2015), enabling us to select superior cultivars based on their ITS2 region. In this presented study, we analyzed the sequence of ITS2 region and secondary structure of Pisang Kepok or Saba banana from West Java, Bali, Flores, and Papua to determine the genetic variability inferred from the dispersal and years of cultivation in the subjected areas. Further, the expected variation might also indicate phenotypic modification that could be beneficial, such as stress tolerance and disease resistance.

2. Materials and Methods

2.1. Plant Materials and DNA Extraction

Seven cultivars of banana were collected from six locations dispersed in four islands of Indonesia (Table 1). Cigar leaves were sampled and placed in a 50 mL conical tube containing a sachet of 2 g of silica gel. Prior to this, the cutter blade was wiped with antiseptic tissue containing 70% alcohol. Within three to four days, after reaching our laboratory at School of Life Sciences and Technology ITB, the sachet of silica gel was removed, and the leaf samples in the conical tube were immediately dipped in liquid nitrogen and stored at -80 ºC until subsequent process.

The DNA was extracted by using CTAB method (Doyle 1990) with a modification to scale down the volume. Also, precipitation step with cold isopropanol (35%)

was made in a shorter incubation time within one or two hours. White DNA strands were apparent in most of the extractions from Saba banana, allowing spooling with a glass hook prior to the final protein precipitation step by using ammonium acetate 5 M (wash buffer). If this was not the case, then the supernatant was poured gently into a fresh 1.5 mL microtube, and wash buffer was added. Quality of the extracted DNAs was analysed by gel electrophoresis, while quantification by spectrophotometer analysis (Eppendorf Biospectrometer Kinetic, Germany). Extracted DNAs with a thick and intact band, and with a ratio of $A_{260/280nm}$ 1.8 from each sampled cultivar were selected for the next steps.

2.2. Amplification of ITS2 Region and Sequence Analysis

Amplification of ITS2 region was done by polymerase chain reaction (PCR) using the universal primer pair as described in Chen et al. (2010). The reaction was set for denaturation at 94 ºC for 30 s, specific primer annealing at 55 ºC for 30 s, and elongation at 72 ºC for 30 s for a total of 40 cycles. A pre-denaturation reaction was made at 94 ºC for 1 min at the beginning and post-elongation at 72 ºC for 1 min prior to concluding the PCR. The 50 µL of PCR reaction comprised 5 µL of DNA samples and mixed with 25 µL of GoTaq Green Master Mix (Promega Co., USA); 15 µL of Nuclease-Free Water; and 2.5 µL of each of the primer set ITS-S2F and ITS-

S3R. The primer sequences used in the reaction were ITS-S2F: 5'-ATGCGATACTTGGTGTGAAT-3' and ITS-S3R: 5'-GACGCTTCTCCAGACTACAAT-3' (Gu2).

The PCR products were then purified by using Wizard SV® Gel and PCR Clean-up System (Promega, USA) and sequenced at Macrogen (South Korea) using pair-end Sanger sequencing platform. The contigs of sequencing results were constructed using CodonCode Aligner software (CodonCode Corporation, Dedham, Massachusetts), and ITS2 region was annotated based on ITS2 database (Merget et al. 2012). Following this, annotated sequences were analyzed for identity against GenBank data by using BLASTn (Altschul et al. 1990). The sequencing results are presented in Suppl. Table 1.

2.3. Phylogenetic Tree Construction and Secondary Structure Analysis

Multiple sequence alignment was created by using ClustalX (Thompson et al. 1997) and submitting sequences of ITS2 from the seven cultivars of banana. Then, the phylogenetic tree was built on MrBayes version 3.2 (Ronquist and Huelsenbeck 2003) by using the model of general time-reversible (rates= gamma, nst=mixed). Markov-Chain Monte-Carlo (MCMC) reconstruction method was set to perform repetitions as many as 1 million generations (ngen=1,000,000). Prediction of the secondary structure of the analyzed banana ITS2 was made in ITS2 ribosomal DNA database,

TABLE 2 Comparison of sequence length and percentage of GC content in the ITS2 region of analysed bananas.

No	Local Name (Scientific Name)	ITS2 Length (bp)	% (GC content)
1	Pisang hias (*Musa ornata*)	218	70.18
2	Pisang hias (*Musa velutina*)	218	69.27
3	Kepok - West Java (*Musa* spp. (ABB subgroup 'Saba'))	220	71.36
4	Kepok - Bali (*Musa* spp. (ABB subgroup 'Saba'))	220	71.36
5	Kepok - East Nusa Tenggara (*Musa* spp. (ABB subgroup 'Saba'))	220	71.36
6	Kepok - Papua (*Musa* spp. (ABB subgroup 'Saba'))	220	71.36
7	Kepok tanjung (*Musa* spp. (ABB group))	220	71.36
8	Sobo (*Musa* spp. (ABB group))	218	71.56
9	Poh (*Musa* spp. (ABB group))	220	71.36
10	Sabe macan (*Musa* spp. (ABB group))	218	71.56
11	Siam (*Musa* spp. (ABB group))	221	72.4
12	Bali (*Musa* spp. (ABB group))	220	71.36
13	Ambon hijau (*Musa acuminata* (AAA subgroup 'Cavendish'))	218	70.18
14	Ambon kuning (*Musa acuminata* (AAA subgroup 'Gross Michel'))	218	71.1
15	Ambon lumut (*Musa acuminata* (AAA subgroup 'Cavendish'))	219	69.41
16	Cavendish (*Musa acuminata* (AAA subgroup 'Cavendish'))	219	70.32
17	Tembaga (*Musa acuminata* (AAA subgroup 'Green Red Banana'))	222	71.17
18	Rojo molo (*Musa acuminata* (AAA subgroup 'Green Red Banana'))	222	71.17
19	Mas (*Musa acuminata* (AA subgroup 'Sucrier'))	220	70.91
20	Bile (*Musa* spp. (AB group))	220	70.45
21	Mas marlin (*Musa acuminata* (AA subgroup 'Berlin'))	220	71.36
22	Klutuk (*Musa balbisiana* Colla (B group))	220	71.36
23	Klutuk wulung (*Musa balbisiana* Colla (B group))	220	71.36

http://its2-old.bioapps.biozentrum.uni-wuerzburg.de/cgi
-bin/index.pl?predict (Koetschan et al. 2010).

E-value less than 0.01 and Viridiplantae model was selected for ITS2 bio apps motif prediction, while the secondary structure was also predicted in the same program with e-value cutoff < 1e-16 and model possessing the highest transfer helices was depicted (Koetschan et al. 2010). ITS2 sequences were aligned with the ClustalO algorithm (Sievers and Higgins 2014) and integrated within SeaView 4 viewer (Gouy et al. 2010). Consensus representation of multiple sequence alignment was transformed using WebLogo 3 (http://weblogo.threeplusone.com/) (Crooks et al. 2004) and then was manually juxtaposed according to its motif prediction.

3. Results and Discussion

The 23 accessions in this study were the representation of AA, AAA, AB, BB, and ABB genome of bananas. Bananas and plantains are known to have originated in Papua New Guinea and East of Indonesia, which then further generated a diverse species and cultivars of *Musa* as dispersing to South East Asia and China (Simmonds 1962; De Langhe 2009). Currently, these regions have the greatest banana diversity in the world. Cavendish banana is the most popular globally, with most are imported from the tropics and subtropics (FAO 2020). However, this cultivar has a very narrow genetic pool as it has been propagated through a vegetative manner in agricultural practices. The cultivar is susceptible to a wide range of pathogens and to breed in a generative manner is problematic (Daniells et al. 2001, 2013).

Pisang Kepok (ABB genome) exhibits superior traits compared to Cavendish (AAA genome), which are tolerant to drought and *Fusarium* wilt disease (Foc TR1). Traditionally, Pisang Kepok or Saba banana has also been providing nutrition to the rural populations in Indonesia and The Philippines. Thus, improving the quality and quantity of Pisang Kepok production could be a promising alternative to diminish the dominance of Cavendish banana in the global market, which is also expected to be more sustainable although this would need more effort to also educate the consumers since Pisang Kepok requires a cooking process to be more tasteful. However, almost all bananas in the local market came from backyard farming with poor management. The production scale of these small-holding farming is usually quite large, collectively, but often do not meet the standards for export. Resolving the genetic variance in local banana cultivars would assist the selection of superior cultivars inbreeding, that further will help the local farmers by providing seedlings with better qualities.

3.1. Morphological characteristics of the collected Pisang Kepok

The most striking morphological character of *M. acuminata* and *M. balbisiana* is the color of bract that touches the younger layer or facing downward. Cultivars with B genome tend to display red color from the base to the tip,

while those with A genome exhibits a yellowish or cream color at the base. Further, the petiole leaves canal of bananas containing A genome, are usually widely open but closed or overlap in bananas containing B genome. The general morphologies of the collected samples are compliant with the description of Pisang Kepok or Saba banana. They possess a pseudostem with a height of 3 m or higher, waxy and green-yellowish in color, and the leaves are intermediate upright. The petiole margins are enclosed, not winged, and clasping the pseudostem, and the base of the second leaf is rounded on the sides. The rachis of Pisang Kepok is perpendicularly hanging with an ovoid inflorescence. The fruits are usually 21-25 cm in lengths, with a rounded tip and a persistent style of the stamen.

In this study, four accessions of Pisang Kepok were collected from Bandung (West Java), Denpasar (Bali), Flores (East Nusa Tenggara), and Sorong (Papua). The detachment style of bractea between Pisang Kepok from West Java and Bandung was quite distinctive. In Flores, the bractea rolled up prior to detaching, while in Bandung, they opened upward then detached. This may not be a result of genetic variance but simply a response to temperature in their respective habitat. The higher temperature in
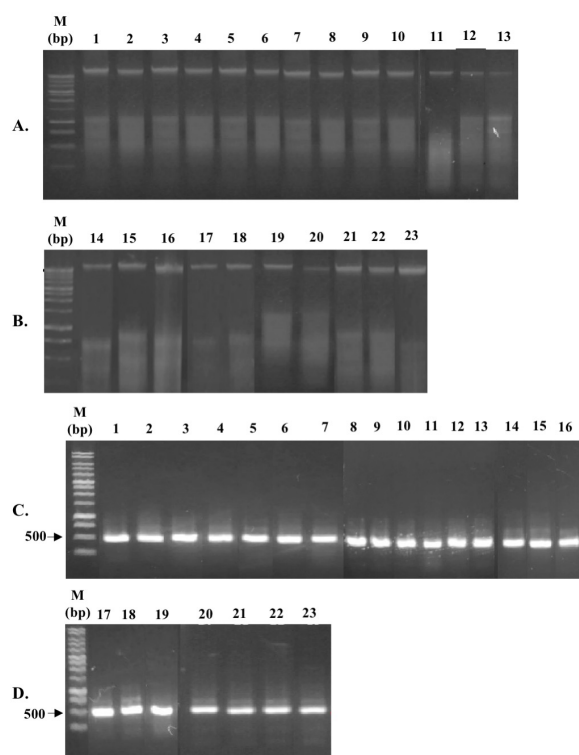

FIGURE 1 Electrophoresis results of DNA extraction and ITS2 region amplification from 23 accessions of local Indonesia bananas. (A) and (B) are DNA extraction results, while (C) and (D) are the visualization of ITS2 amplified fragments. M. 1 kb Marker ladder, 1. Pisang hias (*M. ornata*), 2. Sabe macan, 3. Mas, 4. Ambon hijau, 5. Kepok tanjung, 6. Rojo molo, 7. Tembaga, 8. Sobo, 9. Poh, 10. Bali, 11. Ambon lumut, 12. Ambon kuning, 13. Mas marlin, 14. Klutuk, 15. Pisang hias (*Musa velutina*), 16. Siam, 17. Cavendish, 18. Bile, 19. Klutuk wulung, 20. Kepok - Bali, 21. Kepok - Jawa, 22. Kepok - Flores, 23. Kepok - Papua.

```
>Kepok_Jawa   CGCTTTCGACGCTTCGCCGATGCCCCCCTCGGGGGGGGGTGGAGGCGTGTGCGGAGGATG
>Kepok_Papua  CGCTTTCGACGCTTCGCCGATGCCCCCCTCGGGGGGGGGTGGAGGCGTGTGCGGAGGATG
>Kepok_Flores CGCTTTCGACGCTTCGCCGATGCCCCCCTCGGGGGGGGGTGGAGGCGTGTGCGGAGGATG
>Kepok_Bali   CGCTTTCGACGCTTCGCCGATGCCCCCCTCGGGGGGGGGTGGAGGCGTGTGCGGAGGATG

>Kepok_Jawa   GCCCCCCGTGCCCGAGGGTGCGGTTGGCCGAAGAGCGGGCCGTCGGTGGTTGTCGAACA
>Kepok_Papua  GCCCCCCGTGCCCGAGGGTGCGGTTGGCCGAAGAGCGGGCCGTCGGTGGTTGTCGAACA
>Kepok_Flores GCCCCCCGTGCCCGAGGGTGCGGTTGGCCGAAGAGCGGGCCGTCGGTGGTTGTCGAACA
>Kepok_Bali   GCCCCCCGTGCCCGAGGGTGCGGTTGGCCGAAGAGCGGGCCGTCGGTGGTTGTCGAACA

>Kepok_Jawa   CGACGCGTGGTGGATGCCTTGTGCGAGCCGTACGTCGTGCCTTCGGGACCCGGGCGAGGC
>Kepok_Papua  CGACGCGTGGTGGATGCCTTGTGCGAGCCGTACGTCGTGCCTTCGGGACCCGGGCGAGGC
>Kepok_Flores CGACGCGTGGTGGATGCCTTGTGCGAGCCGTACGTCGTGCCTTCGGGACCCGGGCGAGGC
>Kepok_Bali   CGACGCGTGGTGGATGCCTTGTGCGAGCCGTACGTCGTGCCTTCGGGACCCGGGCGAGGC

>Kepok_Jawa            CTCGAGGACCCAAGTCGTGGTGCGAGTCGATGCCACGGACCG
>Kepok_Papua           CTCGAGGACCCAAGTCGTGGTGCGAGTCGATGCCACGGACCG
>Kepok_Flores          CTCGAGGACCCAAGTCGTGGTGCGAGTCGATGCCACGGACCG
>Kepok_Bali            CTCGAGGACCCAAGTCGTGGTGCGAGTCGATGCCACGGACCG 220
```

FIGURE 2 Multiple sequence alignment of ITS-2 region among Kepok Jawa from Bandung (West Java), Kepok Papua from Sorong (Papua), Kepok Flores from Flores (Nusa Tenggara), and Kepok Bali from Denpasar (Bali).

Flores could prompt more rapid transpiration compared to the colder temperature in Bandung. Bractea tip is prone to turgor loss, hence curling up, as it is the farthest from water-supplying xylem in the pseudostem. Thus, we did not find any significant differences in the morphology of Pisang Kepok from West Java, Bali, East Nusa Tenggara, and Papua.

3.2. DNA extraction, ITS2 amplification, and sequence analysis

The concentration of acquired DNA was approximately 300 - 1000 ng/μL. Extracted DNA from all accessions was visualized in electrophoresis gel and resulted in a single band at length longer than 14 kb (Figure 1A, 1B). However, some shorter smeared bands were also detected that might be contamination of RNA. This was not a problem because the products of mis-annealed primer would be visible in an electrophoresis gel. The fragment of ITS2 was detected at a length of approximately 500 bp (Figure 1C, 1D), and no shorter band was visible, indicating that amplification was made based on the DNA template. The resulted fragments were longer than banana ITS2 analyzed by Hřibová et al. (2011), which were in a range of 205 – 250 bp. This was due to the amplified region of ITS2 in this study also included parts of the 5.8S and 28S regions of the ribosomal RNA genes. Following this, the amplified fragments from each accession were sequenced and analyzed using BLAST (Altschul et al. 1990). The results showed that all submitted sequences have 99-100% query cover and identity to ITS2 region of *Musa* cultivars with AA, AB, B, AAA, and ABB genome (Table 2).

The length of amplified ITS2 fragments was ranging from 218 to 222 after removing the 5.8S and 28S regions

from the sequence (Table 3). These results were comparable to Hřibová et al. (2011), documenting that *Musa* ITS2 regions were 205-250 bp long. The GC content of *Musa* ITS2 fragments in this study was in the range of 69.27% - 71.56%, which was slightly higher than the results of Hřibová et al. (2011) of 56.11% - 70.97%. We also found that *Ambon Hijau, Ambon Lumut,* and Cavendish possess diverse length and GC content of ITS2, although these three cultivars are classified as *M. acuminata* (AAA) subgroup Cavendish. However, this was not documented in *Musa* spp. (ABB) subgroup Saba, with identical length and GC content from all accessions. In general, the ABB and BB bananas in this study have ITS2 region comprising of 220 nucleotides and 70.36% GC content. Based on the multiple sequence alignment, Pisang Kepok from West Java, Bali, Flores, and Papua also exhibited an identical sequence of ITS2 regions (Figure 2). De Langhe (2009) suggests that the variance of length and GC content within a *Musa* species could be a consequence of mutation. Several types of mutation that possibly caused this variance are the insertion or deletion of a single nucleotide, nucleotide substitution, transition mutation of self-replacement to purine or pyrimidine, or transverse mutation of cross-replacement to purine or pyrimidine (Campbell and Reece 2000).

3.3. Phylogenetic analysis

The phylogenetic tree was constructed by using the Bayesian method, involving the search of the most representative tree by generating up to 88% - 90% of highly probable trees (Xiong 2006). Statistically, Bayesian is highly correlated to maximum likelihood, focusing on the prior probability and likelihood (Holder and Lewis 2003). The separation of ingroup and outgroup in Bayesian anal-
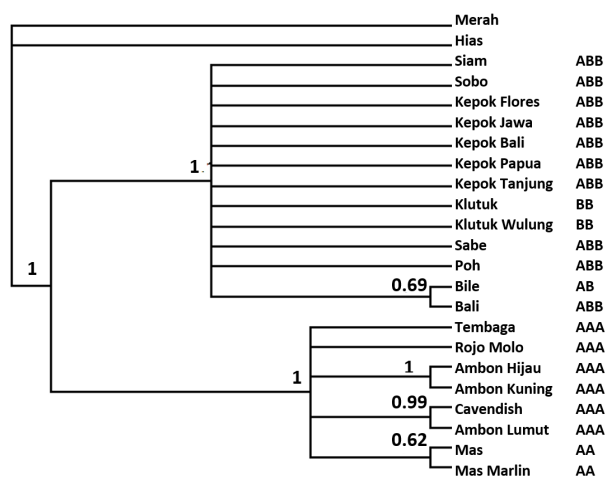
FIGURE 3 Phylogenetic tree built on nucleotide sequence of ITS2 region of 23 accessions banana in Indonesia.

ysis is based on 1.00 posterior probabilities (PP = 1.00) (Figure 3). The outgroup covers *M. ornata* and *M. velutina,* while ingroup comprises of *M. acuminata* and *M. balbisiana.* The genetic distance between *M. ornata* and *M. velutina* is relatively far to be included as ingroup, but adequately close to being the outgroup as both species are still members of *Musaceae.*

Clade I comprises of bananas with B genome (*M. balbisiana*) (PP = 1.00). The four accessions of Pisang Kepok are grouped with *Pisang Sobo, Kepok Tanjung, Poh, Sabe Macan, Bali, Siam* (ABB genome), *Klutuk, Klutuk Wulung* (BB genome), and *Bile* (AB genome). Siddiqah (2002) suggested that morphologically Pisang Kepok is tightly related to *Klutuk Wulung* and cultivars with ABB genome. Pisang Kepok and other ABB cultivars are grouped with *Klutuk Wulung* (BB genome) due to their dominant characters such as the ovoid shape of the inflorescence, ovulum are grouped in four per locus, and the detachment of bractea in an opening pattern (Siddiqah 2002). Clade I display polytomy as indicated by the insufficient data to display an additional branch. This suggested that there was no genetic variance in bananas with AB and BB genome, which was also confirmed the identical length and GC content of their respective ITS2 sequence. All accessions of Pisang Kepok are grouped together with bananas of genome AB, and ABB generated a polytomy (Dover 1982), which was a consequence of the dominance of one particular type of rDNA sequence in the hybrid of *M. acuminata* x *balbisiana* (Kovarik et al. 2004). Further, this might have led to the complete homogenization of rDNA locus from *M. balbisiana* (B genome). This also could be a consequence of chromosome replacement between two genomes (A and B genome) (Wang et al. 2019). *Bile* (AB genome) generated a sister clad with *Bali* (ABB genome) (PP = 0.69) in clade I. The reconstruction of phylogenetic tree by using Bayesian analysis showed that *M. balbisiana* is separated from the other taxa and created a polytomy.

Accessions with A genome (*M. acuminata*) are

grouped in clade II (P = 1.00). Ambon Lumut generated a sister clad with Cavendish (AAA genome) (PP = 0.99), Ambon Hijau with mbon Kuning (AAA genome) (PP = 1.00). Mas generated a sister clad with Mas Marlin (AA genome) (PP = 0.62). The ITS2 region of M. acuminata (AA or AAA genome) is more variable compared to *M. balbisiana* that created polytomy. Genetic variability and density of AA or AAA genome are located in Indonesia, while ABB genome in India, The Philippines, and Vietnam (De Langhe 2009). The hybrid breeds of AB and ABB genome are originated from *M. acuminata* that was brought by humans to the native area of *M. balbisiana* such as India, Myanmar, Thailand, and The Philippines (Price 1995). The low genetic variability of *M. balbisiana* in Indonesia might due to the narrow pool of common ancestor where they originated from. The event of backcrossing is thought to induce gene substitution in *M. acuminata* by *M.balbisiana*, that only several parts of *M.acuminata* (A genome) were left in bananas with ABB and AB genome (De Langhe et al. 2010).

3.4. *Predicted secondary structure of ITS2*

Although the comparison of ITS2 sequences has been acknowledged as a powerful tool in discriminating taxa, it still needs improvement to identify a broad range of species. One of the strategic means is by analyzing the secondary structure of ITS2, as this does not increase the number of sequenced area and may also infer function (Zhang et al. 2015). Generally, RNAs activities in cells are determined by their secondary structures, although the definitive role of ITS2 RNA in molecular regulation is still yet to reveal. ITS2 secondary structure is highly conserved in plant DNA (Jobes and Thien 1997). It consists of four helices structured spacers that are unique and useful to improve discrimination intra-species. Helix I and IV have higher diversity and only helix II and III that could be identified in general to all eukaryotes (Coleman 2007). Helix II could be determined by the content of pyrimidine-pyrimidine nodules, as for Helix III, are characterized by its content, which is the most conserved primer on 5' with UGGU sequence (Schultz et al. 2005).

The predicted secondary structures of ITS2 from all accessions in this study are presented in Table 3. Pisang Kepok from West Java, Bali, East Nusa Tenggara, and Papua exhibit the identical secondary structure of ITS2 as characterized by three bubbles in helix I, a hook-like structure between helix I and II, four bubbles in helix III, and three bubbles in helix IV (Figure 4). The identical structure suggests that these accessions are the same cultivar with no genetic variation. It was interesting that the identical secondary structure of ITS2 was also detected inter-species in *M. ornata* and *M. velutina*, although that was not the case with their sequence length and identity (Suppl. Figure 2). In this case, the RNA of ITS2 in *M. ornata* and *M. velutina* might be playing a consistent role. The hook-like structure between helix I and II was almost constantly found in accessions comprising B genome, but none in cultivars with only A genome. Ambon Hijau and
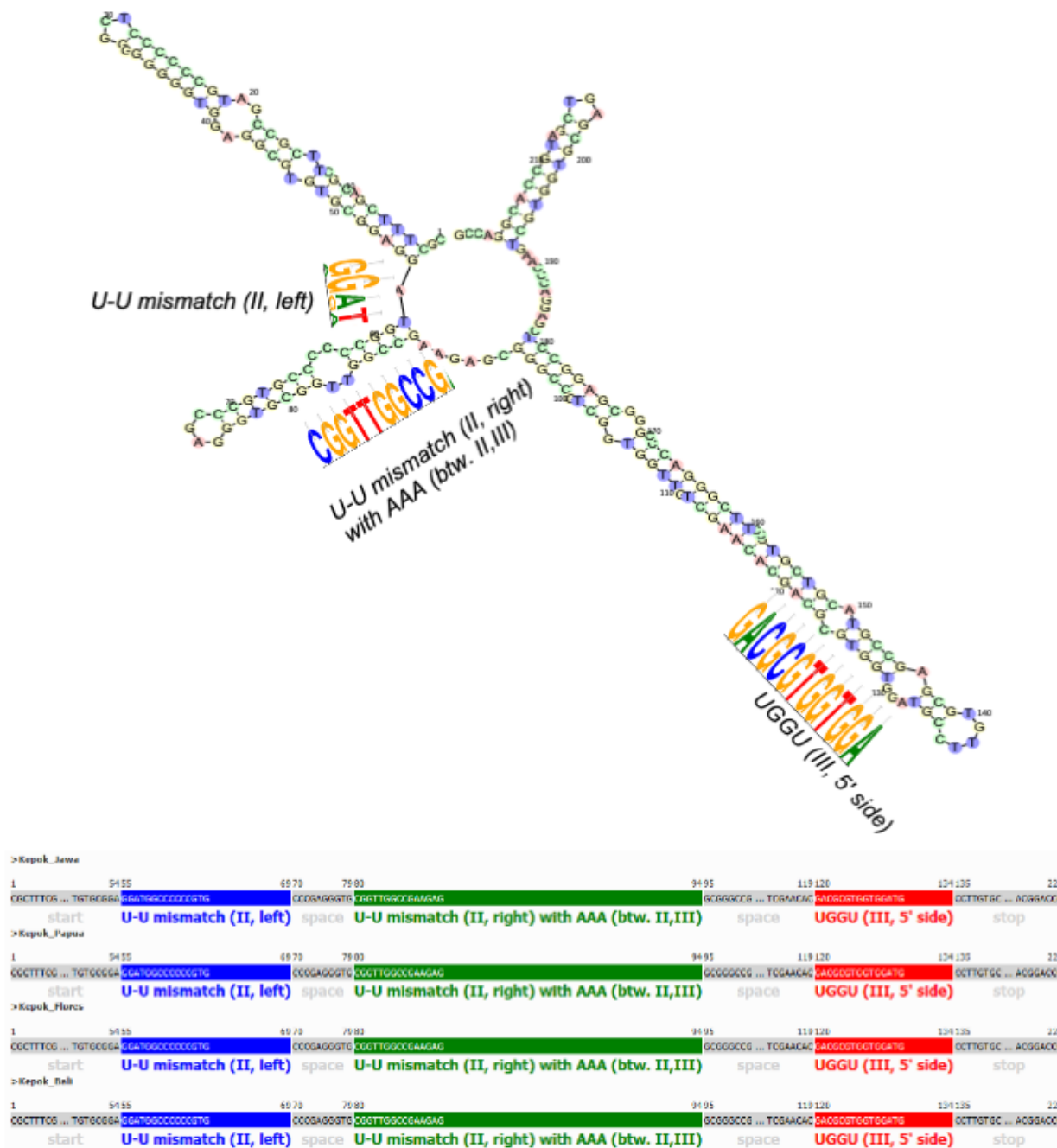
FIGURE 4 Prediction of secondary structure of ITS2 RNA from Pisang Kepok, and also is representing genome AB. The model is identical for the four accessions from West Java, Bali, East Nusa Tenggara, and Papua. It contains four conserved motifs of UGGU, U-U mismatch in between helix I and II, and U-U mismatch in between helix II and III.

Ambon Kuning also possess the identical secondary structure of ITS2, consisting of four bubbles in helix I, one in helix II, three in helix III, and also three bubbles in helix IV (Figure 4). However, this was slightly different from the secondary structure of ITS2 in Ambon Lumut that possesses only two bubbles in helix I (Figure 3)

The formation of bubbles in the secondary structure is a consequence of mismatch pairings of the nucleotide bases on the ITS2 RNA, subsequent to the generation of helices. These mismatch pairings also present in the four

specific and conserved motifs of ITS2 RNA, which were also found in the secondary structure of ITS2 from Pisang Kepok (Figure 4) and all the other bananas analyzed in this study (Suppl. Figure 1-3). The motifs are the UGGU (III, 5' side), U-U mismatch (II, left), and U-U mismatch (II, right) with AAA (btw. II, III). These motifs are predicted by 14 unique bases from the sequences of ITS2 RNA. Generally, the UGGU motif precedes the apex of the helix III or a pyrimidine-pyrimidine mismatch in helix II (Merget et al. 2012). Our data showed that UGGU

TABLE 3 The prediction of secondary structure of ITS2 region in several banana cultivars in Indonesia.

No	Local Name (Scientific Name)	Predicted Secondary Structure of ITS2	No	Local Name (Scientific Name)	Predicted Secondary Structure of ITS2
1	Pisang hias (*Musa ornata*)		13	Ambon hijau (*Musa acuminata* (AAA subgroup 'Cavendish'))	
2	Pisang hias (*Musa velutina*)		14	Ambon kuning (*Musa acuminata* (AAA subgroup 'Gross Michel'))	
3	Kepok – West Java (*Musa* spp. (ABB subgroup 'Saba'))		15	Ambon lumut (*Musa acuminata* (AAA subgroup 'Cavendish'))	
4	Kepok - Bali (*Musa* spp. (ABB subgroup 'Saba'))		16	Cavendish (*Musa acuminata* (AAA subgroup 'Cavendish'))	
5	Kepok – East Nusa Tenggara (*Musa* spp. (ABB subgroup 'Saba'))		17	Tembaga (*Musa acuminata* (AAA subgroup 'Green Red Banana'))	
6	Kepok - Papua (*Musa* spp. (ABB subgroup 'Saba'))		18	Rojo molo (*Musa acuminata* (AAA subgroup 'Green Red Banana'))	
7	Kepok tanjung (*Musa* spp. (ABB group))		19	Mas (*Musa acuminata* (AA subgroup 'Sucrier'))	
8	Sobo (*Musa* spp. (ABB group))		20	Mas marlin (*Musa acuminata* (AA subgroup 'Berlin'))	
9	Poh (*Musa* spp. (ABB group))		21	Bile (*Musa* spp. (AB group))	
10	Sabe macan (*Musa* spp. (ABB group))		22	Klutuk (*Musa balbisiana Colla* (B group))	
11	Siam (*Musa* spp. (ABB group))		23	Klutuk wulung (*Musa balbisiana Colla* (B group))	
12	Bali (*Musa* spp. (ABB group))				

motif was found in helix III in all secondary structures of ITS2 from the 23 accessions, starting from base 118, 119, or 120. We also found that each UGGU motif contains four U-U mismatches, one preceding, and another one following. The U-U mismatch (II, left) was consistently predicted between helix I and II from base 53, 54, or 55, with the two U-U mismatch positions were separated by two nucleotides on each ITS2 RNA sequence. Lastly, U-U mismatch (II, right) with AAA (btw. II, III) in accessions was predicted to be positioned between helix II and III with insertion of G in the AAA.

ITS sequences in plants are associated with environmental adaptation such as shade, salinity, and drought tolerance, although much remains unknown (Whittle 2006). The predicted secondary structures of ITS2 in banana cultivars implies variation in the sequences of ITS2 is not always supported by their secondary structure. Hence, analyzing nucleotide sequence and secondary structure is recommended to reveal the phylogenetic relationship between banana cultivars. Our data suggested that the four accessions of Pisang Kepok are not genetically distinct based on their ITS2 RNA sequence and secondary structure. Hence, they could be categorized as the same cultivar.

As the national effort to support the improvement of Indonesia bananas production, the characterization of superior cultivars is a prerequisite. This study showed that Pisang Kepok from Bandung (West Java), Denpasar (Bali), Flores (East Nusa Tenggara), and Sorong (Papua) possess identical ITS2 sequence, which indicates that they may also exhibit tolerance to *Fusarium* (Foc TR1) and drought. Hence, this could be one of the foundations further to analyze the genetic basis of these advantageous traits. This needs more research to inquire about the notion given that the analysis was performed based on ITS2 in which does not correlate to disease resistance phenotype and adaptation to a hostile environment. Traditional or molecular breeding of *Fusarium* and/or drought susceptible cultivars with Pisang Kepok could be performed immediately. However, an extended time frame is necessary if we are to transfer the responsible genes through genetic engineering methods.

An alternative meaning to improve export and increase the price of our local bananas in the international market is to investigate cultivars that are tolerant to post-harvest diseases. Currently, the local production of bananas in Indonesia is considered one of the highest in Asia, following The Philippines and India. However, Indonesian bananas are mostly unable to penetrate the global market due to the poor quality of the harvest. The seeking of these desirable traits, with the fruits that are able to endure long postharvest processing and distribution, could also be performed by the analysis of ITS2 RNA. In this case, we would need to determine the cultivar(s) that exhibit those traits, then compare their ITS2 RNA sequence and secondary structures with other candidates. If the ITS2 RNA of the candidates is highly identical, then they could be further analyzed, as mentioned earlier.

4. Conclusions

We conclude that Pisang Kepok from Bandung (West Java), Denpasar (Bali), Flores (East Nusa Tenggara), and Sorong (Papua) are the same cultivar as indicated by the identical ITS2 length, structure, and secondary structure. This also suggests that the distribution of Pisang Kepok to these four islands in Indonesia took place a short time ago and possibly was introduced by humans for agricultural purposes. Furthermore, it indicates that they might have the identical superior traits of *Fusarium* and drought tolerance, which are useful for the improvement of other banana cultivars. ITS2 has been illustrated in many studies as a robust molecular marker to some degree, particularly in this research, in clustering the bananas based on their genome group. Nevertheless, the use of ITS2 in the minor taxonomical groups such as cultivar could be meticulous and needs extra details. Provided that cultivar 'Kepok' might possess desirable traits such as tolerance to drought and disease such as in Foc TR1, with more research, it assumed that these superior traits might provide the improvement of other banana cultivars.

Acknowledgments

Authors' contributions

FMD and HN designed the study, IF carried out the laboratory work and analyzed most of the data. KW and AS assisted the initial planning and sampling of all accessions in this study. SNP predicted the structure and motif of ITS2 RNA, and then KM prepared the manuscript. All authors read and approved the final version of the manuscript.

Competing interests

The authors declare no competing interest.

References

.

Ahmad F, Megia RITA, Poerba YS. 2014. Genetic Diversity of *Musa balbisiana* Colla in Indonesia Based on AFLP Marker. HAYATI J Biosci. 21(1):39–47. doi:10.4308/hjb.21.1.39.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215(3):403–410. doi:10.1016/S0022-2836(05)80360-2.

Baral A. 2019. Bananas tackling drought and heat – with DREBs and more. Physiol Plant. 165(2):128–130. doi:10.1111/ppl.12905.

Boonruangrod R, Desai D, Fluch S, Berenyi M, Burg K. 2008. Identification of cytoplasmic ancestor gene-pools of *Musa acuminata* Colla and *Musa balbisiana* Colla and their hybrids by chloroplast and mitochondrial haplotyping. Theor Appl Genet. 118(1):43–55. doi:10.1007/s00122-008-0875-3.

Burgos-Hernández M, González D, Castillo-Campos G. 2017. Phylogenetic position of the disjunct species *Musa ornata* (Musaceae): first approach to understand its distribution. Genet Resour Crop Evol. 64(8):1889–1904. doi:10.1007/s10722-016-0479-8.

Campbell NA, Reece J. 2000. Biology. Boston: Benjamin Cummings Pearson.

Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, et al. 2010. Validation of the ITS2 Region as a Novel DNA Barcode for Identifying Medicinal Plant Species. PLoS One. 5(1):e8613. doi:10.1371/journal.pone.0008613.

Coleman AW. 2007. Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. Nucleic Acids Res. 35(10):3322–3329. doi:10.1093/nar/gkm233.

Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. Genome Res. 14(6):1188–1190. doi:10.1101/gr.849004.

Daniells J, O'Keefe V, Smyth H, Gething K, Fanning K, Telford P. 2013. Planet of the Cavendish - understanding the domination. In: Acta Horticulturae, 986. Int Soc Hort Sci (ISHS), Leuven, Belgium. p. 219–224. doi:10.17660/ActaHortic.2013.986.23.

Daniells JW, Arnaud E, Sharrock SL. 2001. Musalogue, a Catalogue of Musa Germplasm: Diversity in the Genus *Musa*. INIBAP. URL https://books.google.co.id/books?id=X3DRoQEACAAJ.

De Langhe E. 2009. Relevance of banana seeds in archaeology. Ethnobot Res Appl. 7:271–281. doi:10.17348/era.7.0.271-281.

De Langhe E, Hribová E, Carpentier S, Dolezel J, Swennen R. 2010. Did backcrossing contribute to the origin of hybrid edible bananas? Ann Bot. 106(6):849–857. doi:10.1093/aob/mcq187.

Denham TP, Haberle SG, Lentfer C, Fullagar R, Field J, Therin M, Porch N, Winsborough B. 2003. Origins of Agriculture at Kuk Swamp in the Highlands of New Guinea. Science. 301(5630):189–193. doi:10.1126/science.1085255.

Dover G. 1982. Molecular drive: a cohesive mode of species evolution. Nature. 299(5879):111–117. doi:10.1038/299111a0.

Doyle JJ. 1990. Isolation of plant DNA from fresh tissue. Focus p. 13–15.

FAO. 2019. Banana facts and figures. Technical report, Food and Agriculture Organizations of United Nation, Rome. URL http://www.fao.org/faostat/en/{#}data/QC.

FAO. 2020. Banana Market Review: Preliminary Results 2019. Technical report, Rome. URL http://www.fao.org/publications/card/en/c/CA7567EN/.

Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol. 27(2):221–224. doi:10.1093/molbev/msp259.

Hřibová E, Čížková J, Christelová P, Taudien S, de Langhe E, Doležel J. 2011. The ITS1-5.8S-ITS2 Sequence Region in the Musaceae: Structure, Diversity and Use in Molecular Phylogeny. PLoS One. 6(3):e17863. doi:10.1371/journal.pone.0017863.

Hidayat T, Kelana HW, Ismanto DIA, Meitha K. 2018. Survey on ethnobotanic value of Banana (*Musa* spp; Musaceae) in Bali Province, Indonesia. HAYATI J Biosci. 25(1):31–39. doi:10.4308/hjb.25.1.31.

Holder M, Lewis PO. 2003. Phylogeny estimation: traditional and Bayesian approaches. Nat Rev Genet. 4(4):275–284. doi:10.1038/nrg1044.

Jobes LV, Thien DB. 1997. A Conserved Motif in the 5.8S Ribosomal RNA (rRNA) Gene is a Useful Diagnostic Marker for Plant Internal Transcribed Spacer (ITS) Sequences. Plant Mol Biol Rep. 15(4):326–334. doi:10.1023/A:1007462330699.

Jumjunidang R, Soemargono A. 2012. Identification and distribution of *Fusarium oxysporum* f. sp. cubense isolates through analysis of vegetative compatibility group in Lampung province, Indonesia. ARPN J Agric Biol Sci. 7(4):297–284.

Koetschan C, Förster F, Keller A, Schleicher T, Ruderisch B, Schwarz R, Müller T, Wolf M, Schultz J. 2010. The ITS2 Database III—sequences and structures for phylogeny. Nucleic Acids Res. 38(suppl_1):D275–D279. doi:10.1093/nar/gkp966.

Kovarik A, Matyasek R, Lim KY, Skalicka K, Koukalova B, Knapp S, Chase M, Leitch AR. 2004. Concerted evolution of 18–5.8–26S rDNA repeats in *Nicotiana allotetraploids*. Biol J Linn Soc. 82(4):615–625. doi:10.1111/j.1095-8312.2004.00345.x.

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(11):e80502. doi:10.1371/journal.pone.0080502.

Maryani N, Lombard L, Poerba YS, Subandiyah S, Crous PW, Kema GH. 2019. Phylogeny and genetic diversity of the banana Fusarium wilt pathogen *Fusarium oxysporum* f. sp. cubense in the Indonesian centre of origin. Stud Mycol. 92:155–194. doi:10.1016/j.simyco.2018.06.003.

Merget B, Koetschan C, Hackl T, Förster F, Dandekar T, Müller T, Schultz J, Wolf M. 2012. The ITS2 Database. J Visualized Exp. (61):e3806. doi:10.3791/3806.

Molina A, Fabregar E, Sinohin V, Yi G, Viljoen A. 2009. Recent occurrence of *Fusarium oxysporum* f. sp. cubense tropical race 4 in Asia. In: D Jones, IV den Bergh, editors, III International Symposium on Banana: ISHS-ProMusa Symposium on Recent

Advances in Banana Crop Protection for Sustainable Production and Improved Livelihoods, volume 828. p. 109–116. doi:10.17660/ActaHortic.2009.828.10.

Morpurgo R, Lopato SV, Afza R, Novak FJ. 1994. Selection parameters for resistance to *Fusarium oxysporum* f. sp. cubense race 1 and race 4 on diploid banana (*Musa acuminata* Colla). Euphytica. 75(1):121–129. doi:10.1007/BF00024539.

Němečková A, Christelová P, Čížková J, Nyine M, den houwe I, Svačina R, Uwimana B, Swennen R, Doležel J, Hřibová E. 2018. Molecular and Cytogenetic Study of East African Highland Banana). Front Plant Sci. 9(1371):1–13. doi:10.3389/fpls.2018.01371.

Opara UL, Jacobson D, Al-Saady NA. 2010. Analysis of genetic diversity in banana cultivars (*Musa* cvs.) from the South of Oman using AFLP markers and classification by phylogenetic, hierarchical clustering and principal component analyses. J Zhejiang Univ., Sci., B. 11(5):332–341. doi:10.1631/jzus.B0900310.

Ploetz RC, Kepler AK, Daniells J, Nelson SC. 2007. Banana and Plantain–an Overview with Emphasis on Pacific Island Cultivars. URL http://www.bananenzeug.ch/wp-content/uploads/2018/06/banana-plantain-overview.pdf.

Poerba YS, Ahmad F. 2010. Genetic variability among 18 cultivars of cooking bananas and plantains by RAPD and ISSR markers. Biodiversitas. 11(3):118–123. doi:10.13057/biodiv/d110303.

Price N. 1995. The origin and development of banana and plant cultivation. In: S Gowen, editor, Banana and plantain. London: Chapman and HIll. p. 1–13.

Ravi I, Uma S, Vaganan MM, Mustaffa MM. 2013. Phenotyping bananas for drought resistance. Front Physio. 4(9):1–15. doi:10.3389/fphys.2013.00009.

Rebouças TA, Haddad F, Ferreira CF, de Oliveira SAS, da Silva Ledo CA, Amorim EP. 2018. Identification of banana genotypes resistant to Fusarium wilt race 1 under field and greenhouse conditions. Sci Hortic. 239:308–313. doi:10.1016/j.scienta.2018.04.037.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 19(12):1572–1574. doi:10.1093/bioinformatics/btg180.

Schultz J, Maisel S, Gerlach D, Müller T, Wolf M. 2005. A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. RNA. 11(4):361–364. doi:10.1261/rna.7204505.

Siddiqah M. 2002. Biodiversitas dan Hubungan Kekerabatan Berdasarkan Marker Morfologi Berbagai Plasma Nutfah Pisang.

Sievers F, Higgins DG. 2014. Clustal Omega, accurate alignment of very large numbers of sequences. In: D Russell, editor, Multiple Sequence Alignment Methods. Methods in Molecular Biology (Methods and Protocols), volume 1079. Totowa NJ: Humana Press. p. 105–116. doi:10.1007/978-1-62703-646-7_6.

Simmonds N. 1962. The evolution of the bananas. London: Longmas Inc.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25(24):4876–4882. doi:10.1093/nar/25.24.4876.

Valmayor RV, Jamaluddin SH, Silayoi B, Kusumo S, Danh LD, Pascua OC, Espino RC. 2000. Banana Cultivar Names and Synonyms in Southeast Asia. Laguna: INIBAP Regoional Office of Asia and The Pacific. p. 8–14.

Wang Z, Miao H, Liu J, Xu B, Yao X, Xu C, Zhao S, Fang X, Jia C, Wang J, et al. 2019. *Musa balbisiana* genome reveals subgenome evolution and functional divergence. Nat Plants. 5(8):810–821. doi:10.1038/s41477-019-0452-6.

Whittle CA. 2006. The influence of environmental factors, the pollen : ovule ratio and seed bank persistence on molecular evolutionary rates in plants. J Evol Biol. 19(1):302–308. doi:10.1111/j.1420-9101.2005.00977.x.

Xiong J. 2006. Essential Bioinformatics. New York: Cambrige University Press. doi:10.1017/CBO9780511806087.

Yao H, Song J, Liu C, Luo K, Han J, Li Y, Pang X, Xu H, Zhu Y, Xiao P, Chen S. 2010. Use of ITS2 region as the universal DNA barcode for plants and animals. PloS ONE. 5(10):e13102. doi:10.1371/journal.pone.0013102.

Zhang W, Yuan Y, Yang S, Huang J, Huang L. 2015. ITS2 Secondary Structure Improves Discrimination between Medicinal "Mu Tong" Species when Using DNA Barcoding. PloS ONE. 10(7):e0131185. doi:10.1371/journal.pone.0131185.