Banana Classification Using Sanger Sequencing of the Ribosomal DNA Internal Transcribed Spacers (ITS)

Hongyun Zeng , Bingzhi Huang , Linbing XU , Yuanli Wu *

Posted Date: 22 February 2024

doi: 10.20944/preprints202402.1260.v1

Keywords: Musa; germplasm resources; genotypes; ITS

Article

Banana Classification Using Sanger Sequencing of the Ribosomal DNA Internal Transcribed Spacers (ITS)

Zeng Hongyun, HUANG Bingzhi, XU Linbing and WU Yuanli *

Institute of Fruit Tree Research, Guangdong Academy of Agricultural Sciences / Key Laboratory of South Subtropical Fruit Biology and Genetic Resource Utilization, Ministry of Agriculture and Rural Affairs / Guangdong Key Laboratory of Tropical and Subtropical Fruit Tree Research, Guangzhou 510640, Guangdong, China)

* Correspondence: Corresponding author: WU Yuanli, wuyuanli@gdaas.cn

Abstract: Many types of banana (*Musa* spp.)—one of the most economically important horticultural crops—exist, and their ploidy (diploid, triploid, and tetraploid, usually) and genome types (containing A or/and B genome in most cases) differ. At present, observation and genome type detection are commonly used to identify banana germplasm resources. However, the former is tedious, while the latter cannot distinguish categories below genome types. It is therefore urgent that a simple and effective method for identifying banana germplasm resources is established. We sequenced and analyzed the ribosomal DNA internal transcribed spacer (ITS) sequences of 62 banana germplasm and found that the sequencing peaks, especially the 20 bp region near the 420 bp position (referred to as the "420 bp region"), exhibited relatively recognizable and repeatable polymorphism characteristics. Using the "420 bp region" as a marker, we were able to quickly distinguish bananas belonging to different genome type groups (comprising the AA, AB, AAA, ABB, AAB, ABBB, AAAB, and AAAA groups) or different subgroups in the same genome type group (for example, the Cavendish, Gros Michel, Red, Lakatan, and Ibota Bota subgroups in the AAA group; and the Pisang Raja, Plantain, Pome and Silk subgroups in the AAB group). Moreover, it appeared that Sanger sequencing of ITS could be used for identifying banana hybrid offspring. These results demonstrated that the banana ITS region contains rich genetic information that is useful for genotyping. In general, ITS sequencing simplifies the classification of banana germplasm resources and has potential application in several areas of *Musa* improvement.

Keywords: *Musa*; germplasm resources; genotypes; ITS

Introduction

Banana and plantain (*Musa* spp.) are crops of vital importance to hundreds of millions of people around the world. Most edible cultivars of banana are derived from two wild species, namely *M. acuminata* (A genome) and *M. balbisiana* (B genome). According to Simmonds [1], during their long-term evolution, these two wild species hybridized and interbred, evolving into modern banana through continuous natural selection and artificial selection. Simmonds and Shepherd developed a banana hybrid identification system using 15 prominent traits, among which 13 traits are related to the reproductive organs and are highly polymorphic between *M. acuminata* and *M. balbisiana* [2]. Based on these phenotypic scoring systems, the *M. acuminata* × *M. balbisiana* hybrids could be characterized into different ploidy levels (diploid, triploid, and tetraploid) and several genome groups (e.g., AA, AB, AAA, AAB, and ABB). Generally, one genome group can be divided into several subgroups; for example, the AAA genome group contains Cavendish, Gros Michel, Red, Lakatan, and Ibota Bota. According to morphological markers, highly similar cultivated varieties of banana can be classified into a subgroup, such as Cavendish banana'Brazil','Williams', and'Pei Chiao'. The identification of banana germplasm is the premise for the evaluation and utilization of banana germplasm resources. However, the management of banana germplasm resources has relied immensely on identification using local names and morphological characters, and the extent of the genetic diversity of banana has not been established with molecular markers. Most of the morphological markers are polygenic and highly influenced by the environment. Moreover,

 doi:10.20944/preprints202402.1260.v1

morphological observation is time-consuming and laborious. Despite this, morphological markers are still used in banana breeding due to the unavailability of other markers.

Molecular marker technology has been widely used in germplasm resource identification over the past decades. As described by Kaemmer, DNA oligonucleotide and amplification fingerprinting has been successfully used to detect genetic polymorphisms in 15 representative species and cultivars of the genus *Musa* comprising the AA, AAA, AAAA, AAB, ABB, and BB genotypes [3]. Risterucci [4] demonstrated the usefulness of diversity arrays technology (DArT)—a DNA hybridization-based molecular marker technique that can simultaneously detect variation at numerous genomic loci without sequence information—for genetic diversity analyses of *Musa* genotypes. Numerous studies have utilized few simple sequence repeat (SSR) markers, which have relatively limited genome coverage, in the classification of banana [5–8]. Using SSR and amplified fragment length polymorphism (AFLP) markers, several studies analyzed *Musa* genome groups [9–12]. However, due to objective and subjective reasons, the repeatability of SSRs cannot be guaranteed.

Single nucleotide polymorphisms (SNPs), another class of markers, can accurately distinguish highly similar crop germplasm resources [13]. Moreover, SNPs have higher genetic stability and are the most promising molecular marker at present for differentiating crop germplasm. They are referred to as the third-generation molecular marker. Due to their abundance and genome-wide coverage, and particularly with the advent of high-throughput genotyping methods such as genotyping-by-sequencing, SNP markers have been employed in population genetics studies in banana [14,15]. Alberto [14] used SNP markers in DNA sequencing data related to restriction enzyme sites to study and compare the chromosome structures of 36 banana varieties belonging to the ABB genotype (including different subspecies). Gardoce [16] developed a 1 K SNP genotyping panel, effectively distinguishing between genomic groups, based on the filtering of high-quality genome-wide SNPs from the Musa Germplasm Information System, and used it to assess the genetic diversity and population structure of 183 *Musa* spp. accessions. However, the further application of SNPs in banana germplasm identification needs to be explored.

Fragment variation analysis of the internal transcribed spacer (ITS) in the ribosomal RNA (rRNA) coding gene is widely used to evaluate phylogenetic relationships at a lower taxonomic level, since this region has experienced limited natural selection pressure and exhibits great variation, even among closely related species. Nwakanma [17] analyzed the ITS fragments of nine banana genotypes (AA, BB, AB, AAA, AAB, ABB, AAAA, AAAB, and AABB) by restriction fragment length polymorphism (RFLP) and found that the A and B genomes could be distinguished by Rsa I digestion. In addition, Dita [18] found that the banana *ACTIN2* gene could also be used as a molecular marker to identify the A and B genomes. The above two markers can be used to identify whether banana resources contain A and B genomes, but they are unable to determine the copy number of A and B genomes in polyploids (for example, AAB and ABB cannot be distinguished). Teo [19] used inter-retrotransposon amplified polymorphism (IRAP) to identify the A and B genomes. Subsequently, Nair [20] used copia-IRAP primers to amplify banana germplasm resources which, together with Alu I digestion, could effectively identify AAB and ABB. However, the above methods cannot be used for classification below the genome type. The ITS sequences of 36 banana species (42 accessions from the ingroup representing three genera) together with 10 ingroup accessions retrieved from the GenBank database and four outgroup accessions were used to construct the phylogeny of the banana family [21]. However, it remains unclear whether ITS can be used for the classification of cultivated banana varieties.

The National Litchi and Banana Germplasm Resources Garden in China manages and conserves more than 400 accessions of local and introduced cultivars and wild species of *Musa* in the field and through *in vitro* conservation. The genetic characterization of the germplasm collection has not been explored extensively at the molecular level. In this paper, we discovered that the ITS sequencing peaks (especially the "420 bp region") of banana exhibited recognizable and repeatable polymorphism characteristics. Based on this finding, we developed a new method for identifying banana germplasm resources and successfully divided 62 accessions of banana into 44 "ITS types".

Results

2.ITS sequencing of 62 accessions of bananas

 doi:10.20944/preprints202402.1260.v1

The ITS region of banana is composed of two spacer regions and 5.8S (Supplemental Figure 1). We utilized two primers, namely ITS L and ITS 4, the positions of which are shown in Supplemental Figure 1, to clone the ITS fragment of the cultivated variety Brazil, as described previously by Nwakanma [17]. We found that there were many "non-single peaks" when we checked the ITS sequencing peaks of 'Brazil' (Supplemental Figure 2), implying the heterozygosis of 'Brazil'. We then tested another line of 'Brazil' as a biological repeat. There were few differences among these repetitions, indicating that the peaks of ITS were repeatable (Supplemental Figure 2 and 3). Then, the ITS of other Cavendish varieties, namely Pei Chiao and Formosana (bred from Pei Chiao), was tested and compared with 'Brazil'. These lines were not easily distinguishable (Supplemental Figures 2, 4 and 5), implying that these varieties seemed to be the same subgroup of banana, as far as the ITS results were concerned.

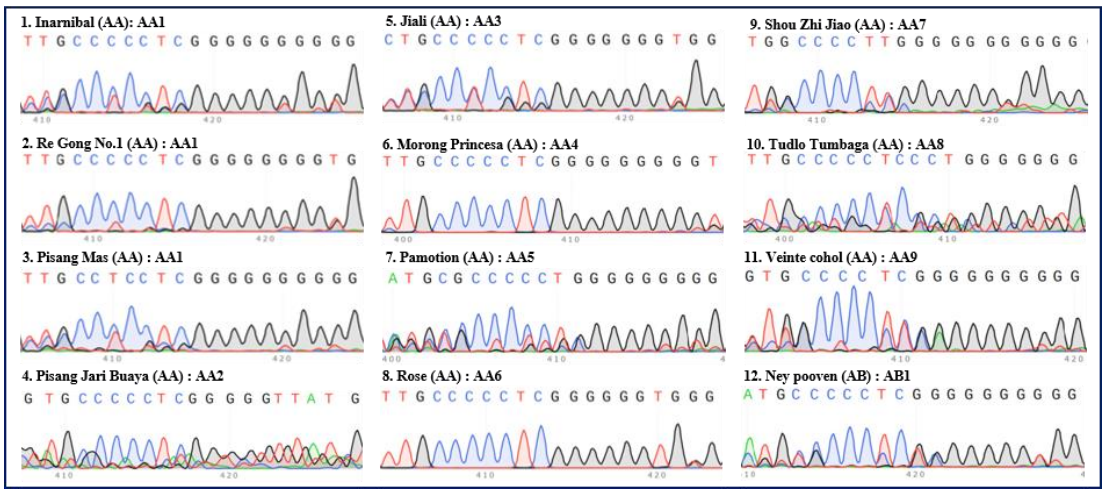

Figure 1. Internal transcribed spacers (ITS) sequence peaks (420 bp region) of 12 accessions of diploid banana. Eleven accessions of AA group banana ('Inarnibal', 'Re Gong No.1', 'Pisang Mas', 'Pisang Jari Buaya', 'Jia Li', 'Morong Princesa', 'Pamotion', 'Rose', 'Shou Zhi Jiao', 'Tudlo Tumbaga', and 'Veinte Cohol') and one accession ('Ney Pooven') of AB group banana were tested. The serial number, name, group, and ITS type of each accession are indicated in each panel. Green, red, blue, and black peaks represent "A", "T", "C", and "G", respectively.

By contrast, upon comparison of three Cavendish varieties, obvious differences were detected in the Gros Michel banana variety Gros Michel and the Ibota Bota banana variety Yangambi KM5 (Supplemental Figures 2, 6 and 7). Moreover, there was a clear distinction between several AAA group bananas and the Pisang Awak banana (ABB group) variety Guang Fen No.1 (Supplemental Figures 2, 6, 7 and 8), indicating that ITS could be a suitable marker for genotyping. In order to explore whether ITS sequencing peaks could reflect the genetic polymorphism of different cultivars, the ITS regions of 62 lines of representative banana germplasm resources were amplified by PCR and sequenced, and the peaks were compared.

2.Classification of different subgroups of bananas using the "420 bp region" of ITS

In total, there were eight genome groups (AA, AB, AAA, ABB, AAB, ABBB, AAAB, and AAAA), 42 accessions from 17 well known subgroups (Inarnibal, Sucrier, Pisang Jari Buaya, Ney Pooven, Cavendish, Gros Michel, Red, Lakatan, Ibota Bota, Pisang Awak, Da Jiao, Pelipita, Saba, Pisang Raja, Plantain, Pome and Silk), 15 accessions of unknown subgroups, and seven accessions of hybrid banana (Table 1). The analysis of 62 sequencing maps proved challenging. Hence, the "420 bp region", a window of the ITS sequencing map 20 bp in size near the position of 420 bp, similar to "NNNCCCCCTCGGGGGGGNNN" in most cases, was selected as a typical target for simplifying the comparison processes, since it exhibited representative polymorphism characteristics (Supplemental Figures 1–8). It was evident that the "420 bp region" of many cultivars differed. In order to facilitate reference, recording, and comparison, different combinations of letters (indicating the genome group) and numbers were used to represent different polymorphism types of the ITS "420 bp region".

 doi:10.20944/preprints202402.1260.v1

Table 1. Information on the 62 accessions of banana.

Serial no.	Names	Groups	Subgroups	ITS types
1	Inarnibal	AA	Inarnibal	AA1
2	Re Gong No.1	AA	Inarnibal	AA1
3	Pisang Mas	AA	Sucrier	AA1
4	Pisang Jari Buaya	AA	Pisang Jari Buaya	AA2
5	Jia Li (mutant of Kluai lep mu nang)	AA	Unknown	AA3
6	Morong Princesa	AA	Unknown	AA4
7	Pamotion	AA	Unknown	AA5
8	Rose	AA	Unknown	AA6
9	Shou Zhi Jiao	AA	Unknown	AA7
10	Tudlo Tumbaga	AA	Unknown	AA8
11	Veinte cohol	AA	Unknown	AA9
12	Ney Pooven	AB	Ney Pooven	AB1
13	Brazil	AAA	Cavendish	AAA1
14	Costa Rica	AAA	Cavendish	AAA2
15	Da Feng No.1	AAA	Cavendish	AAA1
16	Drawf Williams	AAA	Cavendish	AAA1
17	Formosana	AAA	Cavendish	AAA1
18	Nan Tian Huang	AAA	Cavendish	AAA1
19	Nong Ke No.1	AAA	Cavendish	AAA1
20	Pei Chiao	AAA	Cavendish	AAA1
21	Philippines	AAA	Cavendish	AAA1
22	Qi Wei	AAA	Cavendish	AAA1
23	Williams	AAA	Cavendish	AAA1
24	Xian Ba	AAA	Cavendish	AAA1
25	Green Banana	AAA	Red	AAA3
26	Gui Hong Jiao No.1	AAA	Red	AAA3
27	Hong Jiao Wang	AAA	Red	AAA4
28	Red Banana	AAA	Red	AAA5
29	Gros Michel	AAA	Gros Michel	AAA6
30	Berangan	AAA	Lakatan	AAA7
31	Yangambi KM5	AAA	Ibota Bota	AAA8
32	Zhong Jiao No.9	AAA	FHIA-01×SH-3142	AAA9
33	Ai Fen No.1	ABB	Pisang Awak	ABB1
34	Bu Si Fen	ABB	Pisang Awak	ABB1
35	Guang Fen No.1	ABB	Pisang Awak	ABB1
36	Jin Fen No.1	ABB	Pisang Awak	ABB1
37	Dong Guan Da Jiao	ABB	Da Jiao	ABB2
38	Gui Da Jiao No.1	ABB	Da Jiao	ABB3

 doi:10.20944/preprints202402.1260.v1

39	Hai Nan Niu Ba Jiao	ABB	Da Jiao	ABB4
40	Lian Shan Ye Sheng Da Jiao	ABB	Da Jiao	ABB3
41	Pan Yu Da Jiao	ABB	Da Jiao	ABB4
42	Qi Tou Da Jiao	ABB	Da Jiao	ABB5
43	Pelipita	ABB	Pelipita	ABB6
44	Saba	ABB	Saba	ABB7
45	Raja	AAB	Pisang Raja	AAB1
46	Horn Plantain	AAB	Plantain	AAB2
47	Nendran	AAB	Plantain	AAB3
48	Lady Finger Nelson	AAB	Pome	AAB4
49	Zhong Shan Long Ya Jiao	AAB	Silk	AAB5
50	Bengal Cai Jiao	AAB	Unknown	AAB6
51	Kluai Roi wi	AAB	Unknown	AAB7
52	Pisang Ceylan	AAB	Unknown	AAB8
53	Ji Jiao	AAB	Unknown	AAB9
54	Gui Ji Jiao No.1	AAB	Unknown	AAB10
55	King	AAB	Unknown	AAB11
56	Radjah	AAB	Unknown	AAB12
57	Fen Za No.1	ABBB	Guang Fen No.1×Musa Balbisiana	ABBB1
58	Bita-2	ABBB	Fougamou(AAB, Pisang Awak) x Musa Balbisiana 1-63	ABBB2
59	Guang Dong Tetraploid Banana	ABBB	Qi Tou Da Jiao×BB	ABBB3
60	Bita-3	AAAB	(AAAB, Laknau (AAB, Laknau) x Tjau Lagada(AA))	AAAB1
61	FHIA-01	AAAB	Prata Ana×SH-3142	AAAB2
62	SH3436-9	AAAA	Unknown	AAAA1

There were nine ITS types (“AA1–9”) among 11 accessions of AA group bananas (Table 1). While two Inarnibal bananas (‘Inarnibal’ and ‘Re Gong No.1’) and Pisang Mas banana (‘Pisang Mas’) were classified into the one ITS type (“AA1”), another eight accessions from different subgroups represented eight ITS types (“AA2–9”) (Figure 1). The ITS “420 bp region” of the AB group Ney Pooven was an exception and was referred to as “AB1” (Figure 1).

There were nine ITS types (“AAA1–9”) among 20 accessions of AAA group bananas (Table 1). All Cavendish bananas were included in the same ITS type (“AAA1”), except ‘Costa Rica’ (Figure 2). Some extra peaks were observed in the ITS “420 bp region” of ‘Costa Rica’, which was thus referred to as the “AAA2” type, despite its similarity to the “420 bp region” of the other 11 Cavendish accessions (Figure 2). Four red banana accessions were clustered into three types (“AAA3-5”), and these types were evidently similar (Figure 3). ‘Gros Michel’ (Gros Michel subgroup) belonged to type “AAA6” (Figure 3). ‘Berangan’ (Lakatan subgroup), ‘Yangambi KM5’ (Ibota Bota subgroup), and

 doi:10.20944/preprints202402.1260.v1

'Zhong Jiao No.9' ('FHIA-01'×'SH-3142') were "AAA7", "AAA8", and "AAA9" types, respectively (Figure 3).

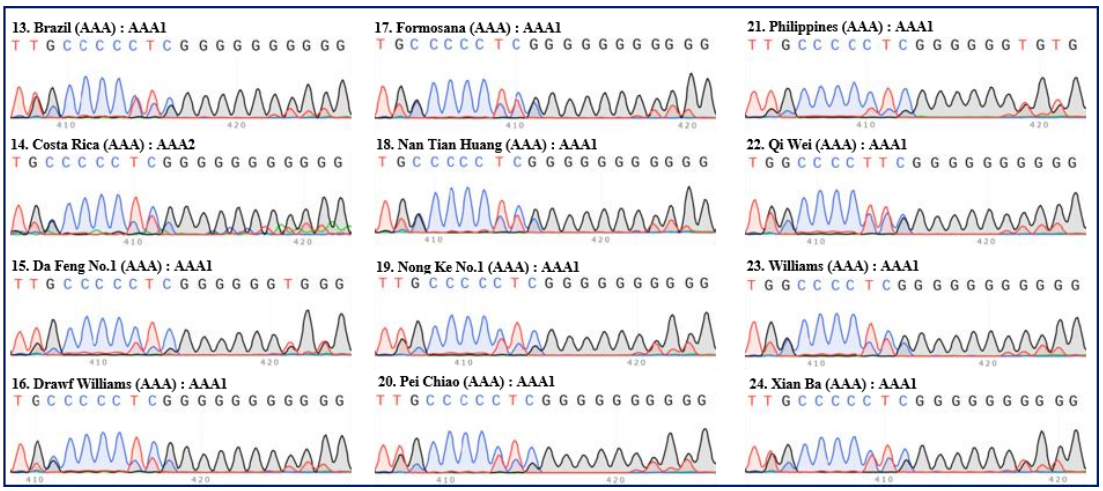

Figure 2. Internal transcribed spacers (ITS) sequence peaks (420 bp region) of 12 accessions of Cavendish banana. Twelve accessions of Cavendish ('Brazil', 'Costa Rica', 'Da Feng No.1', 'Drawf Williams', 'Formosana', 'Nan Tian Huang', 'Nong Ke No.1', 'Pei Chiao', 'Philippines', 'Qi Wei', 'Williams', and 'Xian Ba') were tested. The serial number, name, group, and ITS type of each accession are indicated in each panel. Green, red, blue, and black peaks represent "A", "T", "C", and "G", respectively.

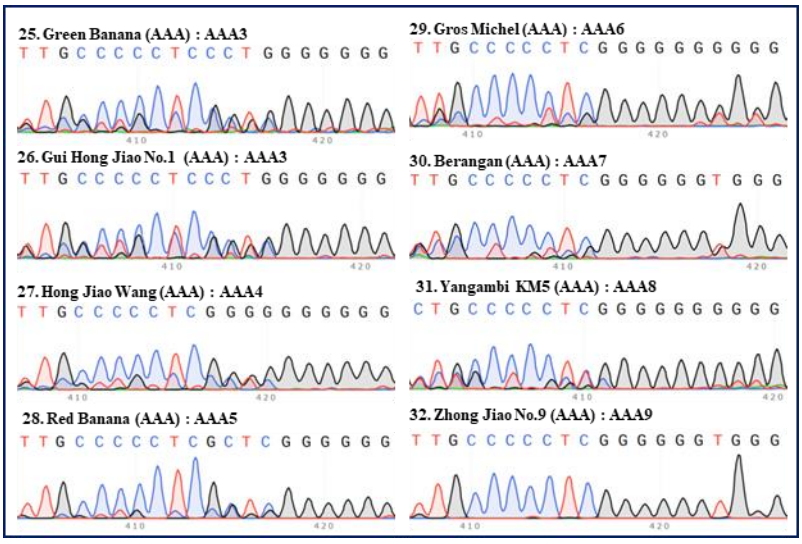

Figure 3. Internal transcribed spacers (ITS) sequence peaks (420 bp region) of eight accessions of AAA group banana, excluding Cavendish.Four accessions of Red ('Green Banana', 'Gui Hong Jiao No.1', 'Hong Jiao Wang', and 'Red Banana'), one accession of Gros Michel ('Gros Michel'), one accession of Lakatan ('Berangan'), one accession of Ibota Bota ('Yangambi KM5'), and one accessions of the hybrid banana ('Zhong Jiao No.9') were tested. The serial number, name, group, and ITS type of each accession are indicated in each panel. Green, red, blue, and black peaks represent "A", "T", "C", and "G", respectively.

There were seven ITS types ("ABB1–7") among 12 accessions of ABB group bananas (Table 1). All four Pisang Awak bananas ('Ai Fen No.1', 'Bu Si Fen', 'Guang Fen No.1', and 'Jin Fen No.1') were clustered into one type ("ABB1") (Figure 4). Six Da Jiao ('Dong Guan Da Jiao', 'Gui Da Jiao No.1', 'Hai Nan Niu Ba Jiao', 'Lian Shan Ye Sheng Da Jiao', 'Pan Yu Da Jiao', and 'Qi Tou Da Jiao') were clustered into four types ("ABB2-5"), though they did exhibit similarities (Figure 4).

Interestingly, none of the 12 accessions of AAB group bananas shared the same ITS type with the others (Table 1). Even the two plantains ('Horn Plantain' and 'Nendran') differed (typed as

 doi:10.20944/preprints202402.1260.v1

"AAB2" and "AAB3", respectively) (Figure 5). The types of 'Ji Jiao' and 'Gui Ji Jiao No.1', which should be the same type of cultivar, were similar (Figure 5).

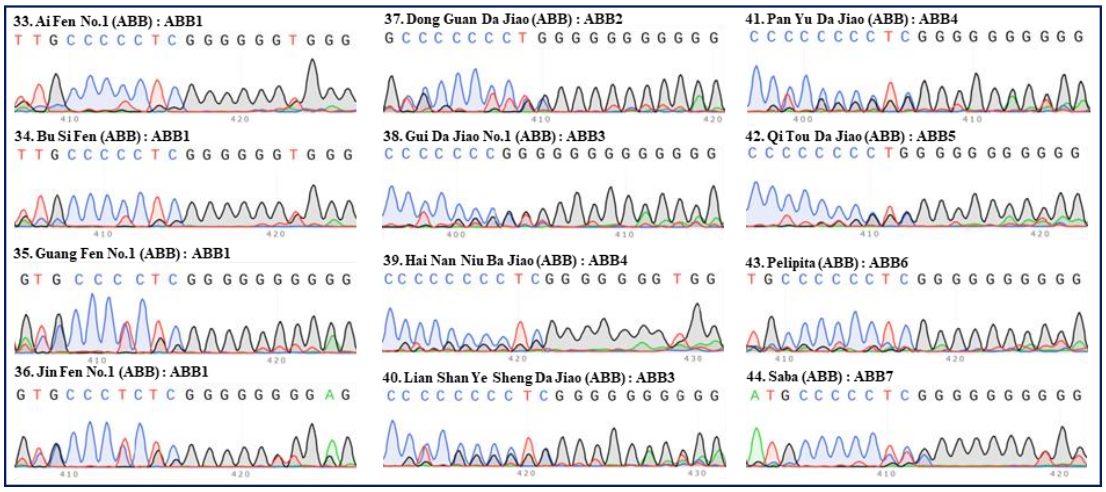

Figure 4. Internal transcribed spacers (ITS) sequence peaks (420 bp region) of 12 accessions of ABB group banana. Four accessions of Pisang Awak ('Ai Fen No.1', 'Bu Si Fen', 'Guang Fen No.1' and 'Jin Fen No.1'), six accessions of Da Jiao ('Dong Guan Da Jiao', 'Gui Da Jiao No.1', 'Hai Nan Niu Ba Jiao', 'Lian Shan Ye Sheng Da Jiao', 'Pan Yu Da Jiao' and 'Qi Tou Da Jiao'), one accessions of Pelipita ('Pelipita') and one accessions of Saba ('Saba') were tested. The serial number, name, group, and ITS type of each accession are indicated in each panel. Green, red, blue, and black peaks represent "A", "T", "C", and "G", respectively.

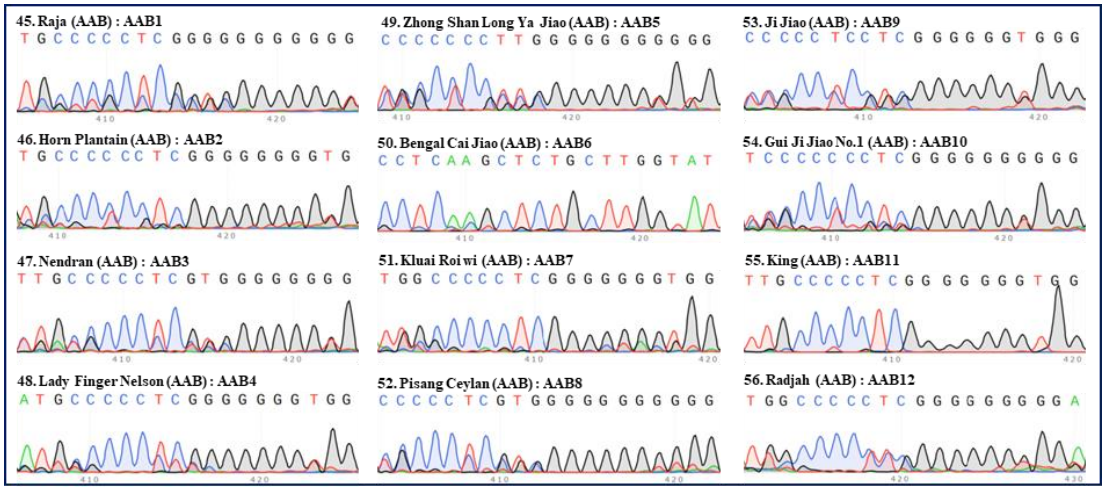

Figure 5. Internal transcribed spacers (ITS) sequence peaks (420 bp region) of 12 accessions of AAB group banana. One accessions of Pisang Raja ('Raja'), six accessions of Plantain ('Horn Plantain', and 'Qi Tou Da Jiao'), one accessions of Pome ('Lady Finger Nelson') and one accessions of Silk ('Zhong Shan Long Ya Jiao') and six accessions of unknown subgroups ('Bengal Cai Jiao', 'Kluai Roi wi', 'Pisang Ceylan', 'Ji Jiao', 'Gui Ji Jiao No.1', 'King' and 'Radjah') were tested. The serial number, name, group, and ITS type of each accession are indicated in each panel. Green, red, blue, and black peaks represent "A", "T", "C", and "G", respectively.

2.Identification of hybrid bananas using the "420 bp region" of ITS

Six accessions of tetraploid banana (including the ABBB, AAAB, and AAAA groups) were hybrid bananas, representing six ITS types ("ABBB1-3", "AAAB1-2", and "AAAA1") (Table 1). Despite the ITS "420 bp region" of 'Fen Za No.1' (typed as "ABBB1") being similar to that of 'Guang Fen No.1' ("ABB1"), a Pisang Awak banana that is the female parent of 'Fen Za No.1', it was easy to tell the difference between type ABBB1 and ABB1 (Figure 6). This implied that ITS sequencing could be applied in identifying banana hybrid offspring. In line with expectations, a tiny discrepancy was

 doi:10.20944/preprints202402.1260.v1

found between the ITS "420 bp region" of 'Guang Dong Tetraploid Banana' (typed as "ABBB3") and its female parent 'Qi Tou Da Jiao' ("ABB1"), and 'Zhong Jiao No.9' (typed as "AAA11") and its female parent 'FHIA-01' (typed as "AAAB2") (Figure 6).

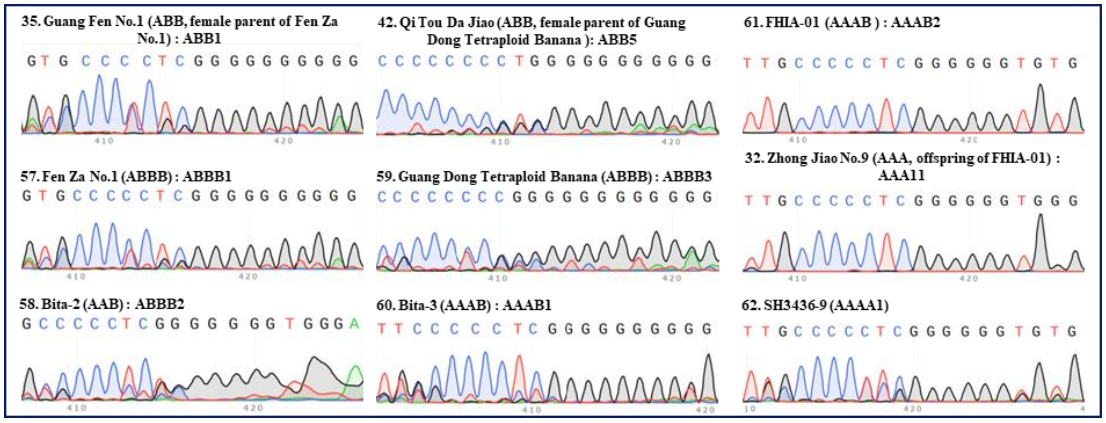

Figure 6. Internal transcribed spacers (ITS) sequence peaks (420 bp region) of several accessions of tetraploid banana and their related cultivars.Seven accessions of tetraploid banana ('Fen Za No.1', 'Bita-2', 'Guang Dong Tetraploid Banana', 'Bita-3', 'FHIA-01', 'Zhong Jiao No.9', 'SH3436-9') were tested. Results of 'Guang Fen No.1' (female parent of 'Fen Za No.1'), 'Qi Tou Da Jiao' (female parent of 'Guang Dong Tetraploid Banana'), 'Zhong Jiao No.9' (offspring of 'FHIA-01') were shown for comparisons. Serial number, name, group, and ITS type of each accession was indicated in each panel. The serial number, name, group, and ITS type of each accession are indicated in each panel. Green, red, blue, and black peaks represent "A", "T", "C", and "G", respectively.

Discussion

The genetic background of modern cultivated banana is complex, which is primarily because (1) during evolution, the ancestors of cultivated banana were formed by different wild species and interspecific hybridization, which endowed them with a rich genetic diversity; and (2) similarly, the long-term vegetative reproduction of cultivated banana varieties resulted in the accumulation of significant genetic diversity. Therefore, the differences in ITS sequences in cultivated banana can reflect the differences in cultivation types to some extent. In the present analysis, the superposition results of the SNP sets from the genomes of banana in the first-generation sequencing peak map of the ITS fragment were used as a fingerprint of an ITS SNP polymorphism. It was confirmed by our experiments that the ITS sequencing peaks, and particularly the "420 bp region", of different banana cultivars could accurately reflect their genetic background to some extent. Using the "420 bp region" as a marker, 62 accessions of banana were clustered into 44 types. The improvement of banana cultivars through crossbreeding is promising [22]. Our method appears to be effective for identifying the offspring and therefore will be useful for the early detection of hybrid banana.

A simple and reliable genotyping method for banana clones, hybrids, species, and relatives will facilitate germplasm management and support breeding initiatives toward a marker-based approach. Various molecular marker technologies have been applied in banana: RFLP, variable number of tandem repeats (VNTR), random amplified polymorphic DNA (RAPD), inter simple sequence repeats (ISSR), ITS, IRAP, AFLP, SSR, DArT, and EcoTILLING [23]. However, these methods are generally complicated. Even though a variety of tools can be used, the classification of banana remains a challenge. For example, the classification of cultivated ABB is necessary, since the more popular names (Saba, Pisang Awak, Peyan, Bluggoe, Monthan) represent a cluster of closely related cultivars generated by somatic variation [14]. The nomenclature of the entire ABB group is difficult to resolve given that the only source of information is the local name of each variant in Asia [14]. This difficulty was confirmed by Saraswathi [24], who combined morpho-taxonomic descriptors and SSR markers and attempted to discriminate the Indian subgroups. Using DArT and SSR markers on a wider sample, researchers confirmed that the classification was consistent for accessions belonging to the subgroups Pelipita, Klue Teparod, and Pisang Awak [15,25]. However, Sardos [15] found that accessions classified as belonging to the subgroups Saba, Monthan, Bluggoe, Ney Mannan, or Peyan

 doi:10.20944/preprints202402.1260.v1

were often misclassified. Herein, using ITS sequencing, we revealed that the classification of most accessions was in accordance with traditional classification. For example, 11 of 12 Cavendish banana used in this study were clustered into one ITS type (**Figure 2**), and all four Pisang Awak bananas were clustered into one ITS type (**Figure 4**). Though there were some exceptions regarding consistency with the traditional classification, this could be clarified using further analysis.

The method outlined in this paper can effectively improve the efficiency of banana germplasm identification as an auxiliary means of character identification and genome type identification. However, it needs to be improved in the future. Sequencing peaks with weak signals occurred frequently and needed to be repeatedly verified. ITS sequencing might also be useful in the identification of other crops. Using this method, further sequencing peak maps of other sequences can be mined and developed as molecular markers for germplasm identification.

Material and methods

4.Materials

The banana germplasm used in this paper was obtained from the National Litchi and Banana Germplasm Resources Garden. As most edible cultivars are derived from *M. acuminata* (AA genome) and *M. balbisiana* (BB), the study was restricted to accessions of these two species and those derived from them. Two biological repeats of each germplasm were tested.

4.Rapid extraction of banana genomic DNA

A small piece of banana leaf (about 1 cm^2) was placed in a 1.5 mL centrifuge tube and ground with an electric grinding rod for about 30 s, following which 600 μL DNA extract solution [100 mM Tris-HCL, 50 mM EDTA (pH 8.0), 500 mM NaCl, 2% SDS, 1% PVP40] was added. The samples were then placed in a water bath at 65°C for 15 min and centrifuged at 12,000 rpm for 5 min, and then the supernatant was placed into a new 1.5 mL centrifuge tube. An equal volume of isopropanol was added, the tube was shaken upside down, centrifuged at 12,000 rpm for 5 min, and then the supernatant was discarded. Five-hundred microliters of 70% ethanol was then added, centrifuged at 12,000 rpm for 5 min, and then the ethanol was discarded. The centrifuge tube was dried in a 65°C oven for 10 min. One-hundred microliters of sterile water was added and the tube was placed in a −20°C refrigerator for storage. Upon preparing the PCR reaction system, 1–2 μL of DNA solution was added.

4.PCR, sequencing, and analysis

The banana ITS region was amplified using the ITSL (TCGTAACAAGGTTTCCGTAGGTG) and ITS4 (TCCTCCGCTTATTGATATGC) primers, as described by Nwakanma [17]. The PCR reaction conditions included pre-denaturation at 95°C for 5 min, denaturation for 30 s, annealing at 55°C for 30 s, extension at 72°C for 30 cycles, and extension for 10 min. Following the PCR reaction, the product was sent to the sequencing company for sequencing with the ITSL primers (the ITS PCR product was about 700 bp in size). For each DNA template, PCR and sequencing were repeated at least twice. After obtaining the sequencing results, the software Snap Gene was used to open the .ab1 file and check the sequencing peak diagram. If there was visible sample pollution, an abnormal peak, or an insufficient sequencing length, the sample was retested. Regions such as "NNNCCCCCTCGGGGGGGNNN" (only the highest peak was counted here) of the ITS sequencing peak maps near the position at 420 bp were found, and screenshots were captured for comparison. To ensure the stability and reliability of the results, the results of different biological repetitions and technical repetitions were compared for each banana germplasm resource to ensure consistency among multiple repetitions.

Acknowledgments: This work was supported by the Natural Science Foundation of Guangdong Province (2022A1515110492; 2023A1515012955) and Guangzhou Science and Technology Plan Project (2023A04J0795). We appreciated Dr. Liu Yu from Very Genome Co., Ltd. (Guangzhou, China) for experimental instruction.

Author Contributions: HYZ conceived and designed the experiments. HYZ and BZH performed the experiments. YLW, BZH, HYZ and LBX contributed reagents/materials/analysis tools. HYZ analyzed the data and wrote the manuscript. All authors reviewed and edited the manuscript.

 doi:10.20944/preprints202402.1260.v1

Conflicts of Interest: The authors have no conflict of interest to declare.

References

1. Simmonds N.W. The evolution of the bananas. *Longmans Green* 1962, *170*.
2. Simmonds N.W.; Shepherd K. Taxonomy and origins of cultivated bananas. *Bot. J. Linn. Soc* 1955, *55*, 302–
3. Kaemmer, D.; Afza, R.; Weising, K.; Kahl, G.; Novak, F. Oligonucleotide and amplification fingerprinting of wild species and cultivars of banana (Musa spp.). *Nat. Biotechnol.* 1992, *10*, 1030–
4. Risterucci, AM.; Hippolyte, I.; Perrier, X.; Xia, L.; Caig, V.; Evers, M.; Huttner, E.; Kilian, A. Development and assessment of diversity arrays technology for high-throughput DNA analyses in Musa. *Theor. Appl. Genet.* 2009, *119*, 1093–103.
5. Hinge V.; Shaikh I.; Chavhan R.; Deshmukh A.; Shelake R.; Ghuge S.; Dethe A.; Penna S.; Kadam U. Assessment of genetic diversity and volatile content of commercially grown banana (Musa spp.) cultivars. *Sci. Rep-UK* 2022, *12*,
6. Ravishankar K.; Vidhya L.; Cyriac A.; Rekha A.; Goel R.; Sharma N.K. Development of SSR markers based on a survey of genomic sequences and their molecular analysis in banana (Musa spp.). *J. Hortic. Sci. Biotech.* 2012, *87*, 84–
7. Rotchanapreeda T.; Wongniam S.; Swangpol S.; Chareonsap P.; Sukkaewmanee N.; Somana J. Development of SSR markers from Musa balbisiana for genetic diversity analysis among Thai bananas. *Plant Syst. Evol.* 2016,
8. Suthanthiram B.; Arumugam C.; Uma S.; Saraswathi M. MusatransSSRDB (a transcriptome derived SSR database)–An advanced tool for banana improvement. *J. Biosciences* 2019,
9. Ahmad F.; Megia R.; Poerba Y.S. Genetic diversity of Musa balbisiana Colla in Indonesia based on AFLP marker. *HAYATI J. Biosci.* 2014, *21*, 39–47.
10. Jesus O.; Silva S.; Amorim E.; Ferreira C.; de C.; José M.; Silva G.; Figueira A. Genetic diversity and population structure of Musa accessions in ex situ conservation. *BMC Plant Biol.* 2013.
11. Mertens A.; Bawin Y.; Vanden Abeele S.; Kallow S.; Vu D.T.; Le L.; Tuong V.D.; Swennen R.; Vandelook F.; Panis B.; et al. Genetic diversity and structure of Musa balbisiana populations in Vietnam and its implications for the conservation of banana crop wild relatives. *PLOS ONE* 2021, *16*, e0253255.
12. Sardos J.; Christelová P.; Cizkova J.; Paofa J.; Sachter-Smith G. & Janssens S.; Rauka G.B.; Ruas M.; Daniells J.; Dolezel J.; Roux N. Collection of new diversity of wild and cultivated bananas (Musa spp.) in the Autonomous Region of Bougainville, Papua New Guinea. *Genet. Resour. Crop Ev.* 2018, *65*.
13. Canal W.M. SNP identification in crop plants. *Curr. Opin. Plant Biol.* 2009, *12*, 211–217.
14. Alberto C.; Julie S.; Yann H.; Guillaume M.; Catherine B.; Nicolas R.; Rony S.; Sebastien C.C.; Mathieu R. Unravelling the complex story of intergenomic recombination in ABB allotriploid bananas. *Ann Bot-London* 2021, *127*, 7–20.
15. Sardos J.; Rouard M.; Hueber Y.; Cenci A.; Hyma K.; Houwe I.; Hribova E.; Courtois B.; Roux N. A genome-wide association study on the seedless phenotype in banana (Musa spp.) reveals the potential of a selected panel to detect candidate genes in a vegetatively propagated crop. *PLOS ONE* 2016, *11*, e0154448.
16. Gardoce, R.; Manohar, A.N.; Mendoza, JV.; Tejano M.; Nocum, J.D.; Lachica, G.; Gueco, L.; Cueva, F.; Lantican, D. A novel SNP panel developed for targeted genotyping-by-sequencing (GBS) reveals genetic diversity and population structure of Musa spp. germplasm collection. *Mol. Genet. Genomics* 2023, *298*.
17. Nwakanma D.C.; Pillay M.; Okoli B.E.; Tenkouano A. PCR-RFLP of the ribosomal DNA internal transcribed spacers (ITS) provides markers for the A and B genomes in *Musa L. Theor. Appl. Genet.*, 2003, *108*, 154–159.
18. Dita, M.A.; Waalwijk C.; Buddenhagen I.W.; Souza M.T.; Kema H.J. A molecular diagnostic for tropical race 4 of the banana fusarium wilt pathogen. *Plant Pathol.* 2010, *59*, 348–357.
19. Teo C.; Tan S.; Ho C.; Faridah Q.; Othman Y.; Heslop-Harrison J.; Kalendar R.; Schulman A. Genomic constitution and classification using retrotransposon-based markers in the orphan crop banana. *J. Plant Biol.* 2005, *48*, 96–105.
20. Nair A.S.; Teo C.H.; Schwarzacher T.; Harrison P.H. Genomic classification of banana cultivars from south India using IRAP markers. *Euphytica* 2005, *144*, 285–290.
21. Li L.; Häkkinen M.; Yuan Y.; Hao G.; Ge X. Molecular phylogeny and systematics of the banana family (Musaceae) inferred from multiple nuclear and chloroplast DNA fragments, with a special reference to the genus Musa. *Mol. Phylogenet. Evol.* 2010, *57*, 1–10.
22. MORÁN J.F. Improvement of cavendish banana cultivars through conventional breeding. *Acta Hortic.* 2013, *986*, 205–208.
23. Biswas M.K.; Yi G. Genes and markers: application in banana crop improvement. *In Banana: genomics and transgenic approaches for genetic improvement*; Mohandas, S., Ravishankar, K., Eds.; Springer Press: Singapore, Singapore, 2016; 35-50.
24. Saraswathi M.S.; Uma S.; Vadivel E.; Durai P.; Siva S.A.; Rajagopal G.; Sathiamoorthy S. Diversity analysis in Indian cooking bananas (Musa, ABB) through morphotaxonomic and molecular characterization. *Acta Hortic.* 2011.

 doi:10.20944/preprints202402.1260.v1

25. Christelová P.; Langhe E.; Hribova E.; Cizkova J.; Sardos J.; Hušáková M.; Houwe I.; Sutanto A.; Kepler A.; Swennen R.; et al. Molecular and cytological characterization of the global Musa germplasm collection provides insights into the treasure of banana diversity. *Biodivers. Conserv.* 2017, *26*.