Genetics and diversity of Indonesian

BANANAS

Fajarudin Ahmad

Genetics and diversity of Indonesian bananas

Fajarudin Ahmad

Genetics and diversity of Indonesian bananas

Fajarudin Ahmad

Contents

Chapter 1.

General Introduction

Fruits are usually a supplementary part of our daily diet, but some can be essential for daily consumption, such as dates in the Middle-East (Zohary & Spiegel-Roy, 1975) or chestnut for Southern European countries (Conedera et. al., 2004). Banana is another example of an essential versatile fruit, at least in some African countries, where it represents 25% of the people's daily intake (IITA, 2000). In other tropical countries, banana might not be a staple, but it grows easily, across all seasons and can vegetatively be reproduced by its corm, allowing traditional farmers or villagers to simply grow these plants in their backyards for self-consumption or as a cash crop. Thus, this fruit is well known and available during all seasons. Besides the fruits, different parts of the plant are appreciated. For instance, leaves are used for food wrapping, fiber production, and colorization (Kennedy, 2009). In many countries, bananas are an important commodity. "Cavendish" bananas are the prime export fruits that are shipped to the Far East and the Middle East, Europe and the United States. Finally, banana is also a symbol of local wisdom rituals, as it is utilized in parts of Indonesia during cultural, religious or wedding ceremonies, as well as for medicine (Hapsari et al., 2017; Sulistyaningsih & Wawo, 2011). All these properties together make the banana a prevalent crop that is disseminated all over the tropics, making it undoubtedly the most important fruit in the world.

Back to just after the human hunting and gathering age, banana is among the oldest domesticated plants in human history, with the first indication of its usage approximately 7,000 years ago (Lentfer, 2009). It is believed that it was firstly cultivated in a region comprising Malesia, Eastern Indonesia and Papua New Guinea (Perrier et al., 2011). Dissemination of this fruit to adjacent areas is assumed to have happened around 600 BC when Asian traders travelled to Southeast Asia and brought the fruits with them to India and Africa. Linguistic arguments for such connections are for instance that the word "banana" is derived from Arabic word "banan" which means finger. This may have inspired Arabic traders during their first banana endeavors across continents (Koeppel, 2008; Simmonds, 1962). Later on, European explorers reached India in 327 BC, discovered bananas as a tasty fruit and introduced them to western countries. From there explorers may have taken them along to Latin America resulting in a global dissemination.

Banana morphology and taxonomy

Banana is a relatively tall herb generally measuring 2-8 meter, although plants in the wild can be as high as 16-18 m (Lentfer, 2009). The plant consists of a subterranean stem or corm, which has a root system and a meristematic part, the latter that can develop

into a lateral shoot (also known as sucker), pseudostem, or flower and fruit (Figure 1). As banana is a monocotyledonous plant it is not a tree and hence its aboveground cylindrical structure is a pseudostem formed by the overlapping leaf parts (Simmonds, 1962).

Taxonomically, banana belongs to the genus *Musa*, in the family of the Musaceae. This family along with the Lowiaceae, Strelitziaceae, Heliconiaceae, Costaceae, Zingiberaceae, Cannaceae and Marantaceae, belongs to the order of the Zingiberales (Kress & Specht, 2006). Initially, banana was commonly utilized as an ornamental plant and the classification was based on just simple morphological characters (morphotaxonomy), such as size and the availability of edible fruit. Linnaeus provided the first description of banana, *Musa paradisiaca*, in 1753 as well as the second description, *M. sapientum*, in 1759. Later, taxonomists realized that these two species were cultivated or hybrid bananas and hence may not be appropriate for banana classification (Cheesman, 1948). Kurz (1865) hypothesized a bispecific origin of cultivars, with the "A" genome donated by *M. acuminata* and the "B" genome by *M. balbisiana*. Later, Sagot (1887) suggested three general groups of bananas. The first group includes the giant bananas such as *Ensete*, the second the bananas with fleshy fruit known as edible bananas and the third group includes the ornamental bananas with upright and bright inflorescences. Then, Baker (1893) subdivided the genus *Musa* into three subgenera, *i.e.* Physocaulis for giant bottle shaped bananas that produce many flowers and inedible fruits, Rhodochlamys for the ornamental bananas and Eumusa for the edible bananas. Cheesman (1947) used cytogenetic features for banana taxonomy. He divided the genus into four sections based on morphological characteristics and chromosome numbers: Australimusa (2n=2x=20), Callimusa (2n=2x=20) (except *M. beccarii*, (2n=2x=18)), Rhadochlamys (2n=2x=22), and Eumusa (2n=2x=22). More recently, Argent (1976) described a banana from Papua New Guinea that he placed in an additional section for this genus, Ingentimusa, with a chromosome number of 2n=2x=14.

Towards a better understanding of banana classification and genome structure

Currently, morphotaxonomy is still the basis of classification according to a list of key morphological characters as described by The International Plant Genetic Resources Institute (IPGRI) (1996). Evidently, morphotaxonomical characters sometimes have limitations to identify the genetic composition of species. Saraswathi et al. (2011) showed that they could not describe correctly the genome composition of 45

accessions, and additionally used microsatellite markers to describe these as ABB bananas. Retnoningsih et al. (2011) used microsatellite markers, which revealed that nine triploid AAA accessions were erroneously identified as AAB or ABB bananas. Hence, molecular markers contribute significantly to the contemporary classification of banana. Besides, precise genetic information is worthwhile for banana improvement programs. Furthermore, molecular markers are useful for genetic diversity studies such as allele composition, the phylogeography of cultivated bananas (Volkaert, 2011), molecular assisted breeding to select banana plants without endogenous banana streak virus sequences (Umber et al., 2016), and the generation of linkage maps and chromosome rearrangement studies (Hippolyte et al., 2010; Martin et al., 2017).

The two commonest basic molecular markers deployed in genetic studies are isozymes and nowadays DNA-based markers. Horry & Jay (1988), Jarret & Litz (1986) and Nasution (1991) used malate dehydrogenase (MDH), phosphoglucomutase (PGM), glutamate oxaloacetate transaminase (GOT), shikimate dehydrogenase (SKDH), and peroxidase to distinguish various ploidy levels of banana and plantains. Horry & Jay (1988) divided Southeast Asia into two areas where edible bananas evolved; (i) Papua and Papua New Guinea and (ii) Sunda, based on anthocyanin composition of wild and cultivated bananas bracts. Nasution (1991) used isozymes for a foundational genetic diversity study of Indonesian wild *M. acuminata*.

Contemporary genetics studies of banana use DNA based markers. They provide more polymorphism, enable the detection of codominant inheritance, are highly reproducible and relatively easy to use, cheap and fast (Kumar et al., 2009). Over time, many markers have been developed such as RFLP (Gawel & Jarret, 1991; Gawel et al., 1992), AFLP (Ude et al., 2002; Wong et al., 2002), PCR-RFLP (Nwakanma et al., 2003), ITS (Hřibová et al., 2011) and ITS with TrnL-F (Liu et al., 2010). Clearly, these techniques enabled a thorough re-evaluation of banana taxonomy. Genome size determination by flow cytometry and the more recently multilocus sequencing (Christelová et al., 2011) and diversity array technology (DArT) (Sardos et al., 2016) further complemented genome wide marker technologies and revealed that only two infrageneric groups were resolved that agree with chromosome number x=11 and x=10/9/7. Later, Häkkinen (2013) merged the previously identified sections Eumusa and Rhodoclamys with 2n=2x=22 into the section Musa and the sections Callimusa, Australimusa and Ingentimusa with 2n=2x=20/18/14 into the section Callimusa. In total, 69 species were classified and assigned to either of these two sections.

Origin of banana cultivars

Cultivated varieties (cultivars) of banana have to a greater part resulted from inter- and intra-specific hybridizations between the two aforementioned wild species, *M. acuminata* (AA) and *M. balbisiana* (BB), resulting in various auto- and allopolyploid hybrids (Perrier et al., 2009; Simmonds & Shepherd, 1955) thousands of years ago in Southeast Asia, the center of origin and diversity of bananas (Perrier et al., 2011). However, the wild species still exist and contribute to sympatric speciation with edible bananas (Simmonds, 1962). Wild bananas are seeded with only a small amount of pulp and are hardly suitable for human consumption (Figure 2). In contrast, edible bananas have seedless fruits that are full of flesh.

Nowadays, about 1,940 cultivars have been described in 39 subgroups and in 15 genome groups in most tropical areas of the world (Crichton et al., 2016; http://www.promusa.org/Banana+cultivar+checklist). Edible cultivars can be diploid (AA, AB), triploid (AAA, AAB, ABB) or tetraploid (ABBB) (Simmonds & Shepherd, 1955). Diploid bananas are the smallest in fruit size, often known as baby banana, such as Pisang Mas (AA) (Figure 2). Triploid bananas are the most popular and the most cultivated, because of their high economic value such as "Cavendish", "Ambon Kuning/Gros Michel" (AAA) , "Kepok/Bluggoe" (ABB) and "Tanduk/Plantain" (AAB) (Valmayor et al., 2000) (Figure 2). Tetraploid bananas are rare (Heslop-Harrison & Schwarzacher, 2007), but some like "Pisang Ustrali" occur in Indonesia, as well as e.g. FHIA-12 which was generated in a breeding program in Honduras (Silva *et al.*, 2001) (Figure 2). Other, less common ancestral contributors to the A genome bananas are *M. textilis* (T genome) and *M. schizocarpa* (S genome), which were found to produce very few edible AT and AS hybrids, respectively, on Papua (Indonesia) and the Pacific islands but they are of no commercial value (Simmonds & Shepherd, 1955). The hitherto overall impression is that the A genome contributes to fruit quality and taste, whereas the B genome contributes to flour content and disease resistance (Swennen & Vuylsteke, 1990; Hohn et al., 2008).

Presumably, in the first steps towards domestication, farmers selected hybrid bananas with seedless and tasty fruits in the forests, as well as banana plants that could be propagated easily in fields or yards (Simmonds & Shepherd, 1955; Heslop-Harrison & Schwarzacher, 2007). Over time, selection contributed to developing elite varieties that were propagated as cultivars. Nowadays, cultivar selection is continued using refined mutation approaches (Chen et al., 2013), enabling clonal selection for disease resistance (Hwang & Ko, 2004), plant stature (dwarfism) (Chen et al., 2016) and productivity (Khayat, 2020; Khayat et al., 1998; Israeli et al., 1996).

Banana production and challenges

Overall banana production in the 1990's is estimated at 50 million tons from 4.7 million ha. and now total production reached 125 million tons harvested from over 6 million ha. (FAO, 2020). The ever-increasing production of banana indicates the paramount value of this fruit. A greater part of its usage is destined for local tropical markets, whereas, 16.5 million ton p.a. is for export (FAO, 2014), of which "Cavendish" is by far the most prevalent cultivar (Bakry et al., 2009) representing 47% of global production, followed by plantains (17%) and "Gros Michel" (12%) (FAO, 2003). The remaining cultivars are to a greater part cooking and desert bananas. On average between 2007-2017, India is the biggest producer with an annual production of over 27 million tons, followed by China (10 million tons) and the Philippines (8 million tons). Indonesia has the highest genetic diversity and ranks sixth with an annual production of seven million tons.

In spite of its economic success for food and income, banana is severely threatened by various bacterial, viral and fungal pathogens in both monoculture plantations as well as traditional small-scale banana farms. *Ralstonia solanacearum* and *R. syzygii* ssp. *celebesensis* are two common bacterial agents that cause wilting known as Moko disease and blood disease, respectively (Blomme et al., 2017). The affected plants may produce some fruits, but these can hardly be marketed, because their fruit pulp is hard with a reddish-brown discoloration. Banana bunchy top virus (BBTV) is caused by a luteovirus and symptoms are characterized by leaf chlorosis and severe stunting which precludes fructification (Hooks et al., 2008). This disease is easily spread to healthy plant by aphids migrating from infected plants (Magee, 1927). Not aware of its risk farmers and salesmen may neglect necessary sanitary measures, which facilitates the spreading of the virus which leads to complete crop losses (Dale, 1987). Although bacterial and viral diseases can decimate yields, good sanitation contributes to effective disease control (Dale, 1987). Fungal diseases significantly affecting banana production are black leaf streak disease (BLSD) or black Sigatoka (Fullerton & Olsen, 1995; Marín et al., 2003) and Fusarium wilt of banana (FWB) or Panama disease (Brandes, 1919; Stover, 1962; Ploetz, 2005; Ordóñez et al., 2015). The former attacks leaves and drops yields, but with extensive chemical control programs, this fungus can be controlled (Arango et al., 2016). However, chemical application is a menace to the environment and raises occupational health issues (Scholthof, 2003; Barraza et al., 2011). On the contrary, FWB usually kills plants before they can produce any fruits and effective disease management strategies are not yet available (Ploetz, 2015; García et al., 2019; Salacinas, 2019; Maryani et al., 2019). FWB is a soil-borne disease that can be spread in manifold ways by moving contaminated soil from infested fields, by water, wind and animals. Moreover, humans plays a major role in the global dissemination of the

disease (Dita et al., 2018). Recently, detailed characterization of the fungus (Maryani et al., 2019), diversity of bananas susceptibility to the disease (Houbin et al., 2004; Zuo et al., 2018; García, 2019) and an evaluation of field management against this disease in the Philippines (Salacinas, 2019) provided important information for global FWB management.

Important traits for banana breeding and sustainable production

The importance of banana for food and income is unquestionable. Similar to other crops, banana breeding targets productivity, biotic stress resistance, abiotic stress tolerance, post-harvest issues and markets demands (Allard, 1960; Tenkouano et al., 2011). So far, the below mentioned breeding programs have delivered new cultivars with improved qualitative or quantitate traits, but none meet the quality of "Cavendish" in taste and appearance, which underscores the need for a larger volume of breeding efforts, alternative control strategies and overall elevated funding.

Taken together, several specific characters, such as marketable fruit or bunch size, plant architecture, parthenocarpy, nematode resistance, BLSD resistance and FWB resistance are main traits for potentially competing new cultivars that can replace "Cavendish" (Khayat & Ortiz, 2011). Fruit or bunch size are important requirements for productivity and the logistic chain (fruits per box/container) (Bakry et al., 2009). Dwarfism is rare but can also be generated by mutation breeding (Khayat, 2020) and is preferred for tolerance to strong winds and an ergonomic harvest processes. This trait is also believed to be controlled by a single recessive gene (dw) (Gubbuk et al. 2004). Breeding for dwarfism is possible by using dwarf accessions such as "Bobby Tannap", as a parent (Ortiz & Vuylsteke, 1995b), but clearly, the dw markers await mapping to efficiently target such a character. Another important target for banana breeding is parthenocarpy (fruit development without seed set). Early genetic studies suggest that this trait is controlled by three independent complementary genes (Simmonds, 1953) and likely primarily occurs in bananas with the A genome although there are also edible AB hybrids (Heslop-Harrison & Schwarzacher, 2007).

Nematodes, BLSD and FWB are all severe threats of banana production. Endo- and ectoparasitic nematodes damage the root system, which results in yield loss due to toppling plants (Stover & Simmonds, 1987). Not a single gene for nematode resistance has been reported, but resistant cultivars, such as "Yangambi" and "Pisang Jari Buaya", are potential donors in breeding programs (Price, 1994). "Yangambi" has never been used for breeding, but "Pisang Jari Buaya" is one of the parents of "FHIA-

01" (Gaidashova et al., 2008). A better understanding of BLSD resistance is urgently required. Previous studies claim that resistance is controlled by one recessive gene and two additive genes (Ortiz & Vuylsteke, 1994), but the overall biology of the causal agent has not been taken into account resulting in field-based phenotyping against a highly diverse pathogen population (Conde-Ferráez et al., 2007; Arango et al., 2016; Kimunye et al., 2020). The wild banana *M. acuminata* ssp. *burmanica* "Calcutta-4" is an iconic source of resistance and has therefore been used for breeding (de Oliveira et al., 2001), but only recently the first indications for specific resistance genes were discovered due to the recognition of the avirulence protein PfAVR4 (Arango et al., 2016). Despite the importance of this disease, presently most attention goes to FWB, surely driven by the ongoing and expanding TR4 epidemic that now also reached Latin America (Ordóñez et al., 2015; García et al., 2019).

Fusarium wilt resistance in banana

Since FWB struck the "Gros Michel"-based industry in Central America in the previous century (Stover, 1962), and the subsequent incursion of BLSD which was firstly reported in 1963 (Rhodes, 1964), breeders have sought for sources of resistance to these disease that can be used for the genetic improvement of banana. Replacing susceptible bananas by resistant genotypes is easy for traditional farmers, and is commonly practiced, but not for the banana industry. The success story of replacing the susceptible "Gros Michel" that succumbed to FWB by resistant "Cavendish" banana varieties in the previous century is not easy to match because all plantations, logistics and markets are now tailored to the latter (Koeppel, 2008). Therefore, the occurrence of *Fusarium odoratissimum* (Maryani et al., 2019) also known as Tropical race 4 (TR4) that causes FWB in "Cavendish" varieties and many local cultivars destined for domestic markets (Zuo et al., 2018; García, 2019) is a true threat to the global industry (Drenth & Kema, 2020; Maymon et al., 2020; Özarslandan & Akgül, 2020; García et al., 2019; Damodaran et al., 2019; Chittarath et al., 2018; Maymon et al., 2018; Ordoñez et al., 2016; Ordóñez et al., 2015; García et al., 2014; Buddenhagen, 2009).

As mentioned above, FWB was hitherto considered to be caused by the fungus *Fusarium oxysporum* f.sp. *cubense* (Foc) (Stover, 1962; Ploetz, 2005a). New recent insights, however, have shown that it is caused by a suite of different *Fusarium* species with TR4 being represented by the new species *F. odoratissimum* (Maryani et al., 2019; Maryani, 2018). Physiologically, this species complex is divided in several races, according to their compatibility with "Gros Michel" (race 1) and "Bluggoe" (race 2) (Waite & Stover, 1960) and race 4, subdivided in subtropical race 4 affecting Cavendish

under abiotic stress (Pegg et al., 2019) and TR4 that kills "Cavendish" varieties and many local cultivars (Su et al., 1986; García, 2020). Race 3 was once classified based on its compatibility with *Heliconia* spp. but has now been abandoned from the race nomenclature. Vegetative compatibility groups (VCGs) were once important for race identification (Ploetz, 2005a; Zuo et al., 2018), but have now been mostly replaced by sequence based diagnosis (Ordóñez et al., 2015; Maryani et al., 2019).

The initial infection of *Fusarium* spp. causing FWB starts by penetration of banana roots, followed by colonization of the vascular system that eventually will be blocked by fungal biomass and plant structures, such as tyloses, to stop fungal proliferation. Older leaves will first show chlorosis and then turn brown, followed by the younger foliage which collectively tip at the petioles and hang down as a skirt around the pseudostem of the plant (Figure 3) (Stover, 1958). During this process these *Fusarium* spp. abundantly form inoculum in the form of macro- and microconidia as well as persistent chlamydospores that can survive for years in infested soil (de Ascensao & Dubery, 2000). Recent insights also have shown that the FWB causing species can also colonize weeds without any symptoms and effectively behave as endophytes until a new host is planted (Salacinas, 2019). In this way *Fusarium* spp. can survive over long periods without compatible hosts, which complicates disease management (Ploetz, 2005; Narayanasamy, 2011; Blomme et al., 2011; Salacinas, 2019).

The fact that banana was taken along with human activities, clearly contributed to the global dissemination of its pathogens (Marín et al., 2003), and recent dissemination of FWB caused by TR4 has clear anthropogenic links (Ordo Ordóñez et al., 2015). Nevertheless, awareness levels remained low until TR4 was detected in Jordan (García et al., 2013). Since then, much more attention developed which was again underscored by the incursion in Mozambique (IPPC, 2013; Viljoen et al., 2020) and more recently in Colombia (García et al., 2019), Turkey (Özarslandan & Akgül, 2019) and Mayotte (Aguayo et al., 2020). Since TR4 appeared in Jordan, subsequent studies revealed its presence in 11 additional countries. In comparison with the previous epidemic in "Gros Michel", the development of tissue culture significantly contributed to the dissemination of healthy plants and hence to the prevention of FWB. However, tissue culture practices also supported the transition of the global banana cultivation in major monocultures, ultimately contributing to the "Cavendish" banana commodity (Ploetz, 2005). Evidently, these are extraordinary vulnerable for diseases as exemplified by the excessive fungicide applications to control BLSD (Aguirre, 2016) as well as the aforementioned more recent introductions of FWB caused by TR4 (Buddenhagen, 1977; Ordóñez et al., 2015; Maryani et al., 2019).

TR4 has seriously affected banana production in Indonesia, the Philippines, the Indian subcontinent and is now also threatening Mediterranean, African and American (sub) tropical countries. The first outbreak of TR4 was in Southeast Asia, and since then it further spread from the center of origin (Maryani et al., 2019) with confirmed incursions in Africa, Western Asia, Australia (Northern territory and Queensland), China, Jordan, Lebanon, Mozambique, Oman and Pakistan, Israel, Turkey, Mayotte, Colombia (Ordóñez et al., 2015). Effective management options have clearly failed and there are currently no commercially available fungicides to reduce the impact of the disease (Salacinas, 2019). The success of host resistance is exemplified by Cavendish bananas that are grown on soils infested with Race 1 strains throughout the world, now already for nearly a century. Thus, breeding bananas with resistance to TR4 is a valid strategy (Damodaran et al., 2009; Amorim et al., 2013) and urgently required. Besides, genetic engineering technologies including CRISPR, should be considered to create resistant plants or increase the resistance level of existing banana varieties (Maxem, 2019; Dale 2017). Alternatively, mutation breeding has resulted in "Cavendish" clones with reduced susceptibility to TR4. These clones have been planted in affected areas and produce marketable fruits (Molina et al., 2016), but suffer from lower productivity as well as several post-harvest issues such as maturity stain that hamper massive adoption. Chen et al. (2013) used ethyl methane sulphonate (EMS) as mutation agent and obtained five putative FWB resistant lines. However, this approach needs a serious selection among several generation to find a stable genotype that meets the objectives. Breeders therefore focus on screening bananas in an effort to identify wild and cultivated bananas with adequate levels of resistance (Houbin et al., 2004; Hwang & Ko, 2004; Rashid et al., 2013; Zuo et al., 2018; García, 2019). The prime issue however, is that these varieties have a narrow focus on improved resistance to FWB but leave other burning issues such as BLSD untouched. In order to meet the plethora of producer and consumer demands banana breeding needs a disruptive change which fully underscores and captures the genetic gain from its ancestors.

In spite of the success of host resistance in "Cavendish" varieties to quench the FWB epidemic that was caused by Race 1 strains, research into the basis of that resistance is essentially lacking or inadequate. Ssali et al. (2013) screened progenies derived from susceptible cultivars pollinated by diploid resistant germplasm against Race 1 and suggested that resistance was inherited as a single recessive gene. Later, Fraser-Smith et al. (2016) used a similar approach to evaluate resistance and suggested that both resistance to TR4 and subtropical Race 4 were under the control of a single dominant allele. However, neither of these genes were mapped nor functionally analyzed. Despite reports on several candidate resistance genes (Fraser-Smith et al., 2016; Kayat, Javed, Wah, & Othman, 2004; Peraza-Echeverria et. al., 2008; Ssali et al., 2013; Sutanto, Sukma, Hermanto, & Sudarsono, 2014) the only hitherto identified, mapped and cloned

resistance gene is RGA2 which was derived from *M. acuminata* ssp. *malaccensis*. Its transfer to susceptible "Cavendish" resulted in resistance under field conditions for several years (Dale et al., 2017). Further unveiling of the genetic base of resistance to FWB in banana is an absolute necessity to further identify and map genes for resistance in order to develop efficient marker assisted breeding strategies.

Banana breeding: outlook and challenges

The developing FWB pandemic underscores the need for genetic improvement as a promising way to develop new resistant cultivars. This can be obtained by traditional breeding cycles *i.e.* selection of compatible parents, crossing and subsequent selection of preferred phenotypes. Examples of current breeding programs are those based at the Fundacion Hondureña de Investigacion Agricola (FHIA), the Brazilian Agricultural Research Corporation (Embrapa), the French agricultural research and international cooperation organization (CIRAD) and the International Institute for Tropical Agriculture (IITA) that focus on the improvement of local and export bananas, plantains and East African highland banana (EAHB). FHIA has released various new banana varieties with varying levels of resistance to e.g. BLSD that are commonly grown in Cuba (Hernández et al., 2007). Embrapa released new varieties for the Brazilian market, such as the latest cv. "Princesa" with resistance to Yellow Sigatoka and resistance to FWB (Amorim et al., 2013; Lédo & Alberto, 2008). CIRAD primarily focuses on new "Cavendish" varieties and plantains in collaboration with the African Center for Research on Bananas and Plantains (CARBAP) in Cameroon (Dépigny et al., 2016), whereas IITA is actively modernizing its EAHB program with support from the Bill and Melinda Gates Foundation (Wilberforce et al., 2014).

This classical way of introgression breeding is challenging, because cultivated bananas are sterile or low in fertility, which is a main limiting factor for seed set (Fortescue & Turner, 2004; Roux et al., 2004; Morán, 2013). For instance, pollination of 20,000 bunches of "Cavendish", resulted in a mere 200 seeds (Morán, 2013). Meanwhile alternative strategies are developed, including generating new triploids from the *M. acuminata* and *M. balbisiana* ancestors (Bakry et al., 2009, 2020) as well as private initiatives driven by the latest genetic tools for speeding up the selection process as well as to select the best parents (www.keygene.com).

As mentioned before, genetic modification is an alternative strategy to produce resistant seedless bananas and includes the development of cisgenic bananas, in addition to the most recent developments around genome editing at e.g. Tropic Biosciences in

the United Kingdom (Maxem, 2019). No doubt that the latter, such as CRISPR-driven strategies, are relatively fast and effective ways for disease improvement (Sander & Joung, 2014), but they still await public and political acceptance in greater parts of the world.

Cytogenetics in banana breeding

Breeders use cultivars and wild relatives as genetic resources for the introgression of desirable traits into their elite lines. However, as mentioned above, cultivated bananas are sterile and parthenocarpic, which makes classical breeding actually unfeasible. In contrast, wild bananas produce fertile gametes and their fruits are therefore seeded, although several hybrids between these diploids are sterile due to structural hybridity or heterozygosity for translocations and inversions, leading to unbalanced chromosome segregations during meiotic divisions, and so to aneuploidy and gamete lethality (Dodds, 1943; Wilson, 1946a; b; Shepherd, 1999). Recently, using mate-pair sequencing, fluorescent in-situ hybridization (FISH) with bacterial artificial chromosome (BAC) probes (BAC-FISH), targeted PCR and DArTseq, Martin et al. (2017) provided evidence for a large translocation in *M. acuminata* ssp. *malaccensis* between chromosomes 1 and 4 that may need to be considered in banana breeding. Occasionally, first or second division restitutions may occur leading to the formation of 2n or unreduced gametes, which in turn can give rise to (sterile) triploid or polyploid offspring (Dodds & Simmonds, 1946), and hence, may be useful for producing new edible bananas (Raboin et al., 2005).

Thus, careful observation of banana chromosomes during mitosis and meiosis is required to understand the cytogenetic cause of sterility. However, revealing details of the tiny banana chromosomes is a challenge. High-resolution chromosome microscopy and improved cytology are essential to obtain detailed images of the chromosomes. Adeleke et al. (2002) developed a silver nitrate staining protocol for spreading pollen mother cells (PMCs), which provided a detailed picture of chromosomes for karyotype and meiotic studies. Later, De Capdeville et al. (2009) developed cell spread slides with PMCs at pachytene that are appropriate for BAC-FISH. High resolution microscopy enables the observation of chromosome abnormalities in the pairing stage, such as asynaptic regions, inversion loops and translocation pairing switches (Adeleke et. al., 2002; De Capdeville et al., 2009), that could otherwise not be displayed by classical cytogenetic methods (e.g., Shepherd, 1999).

Diploid fertile bananas for breeding

As cultivars are sterile, diploid fertile bananas are needed as starting genitors in breeding programs to create hybrid progenies, that later can be used for developing triploid, alloploid or parthenocarpic sterile cultivars. Genetic analysis of fertile diploids has the additional advantage of mapping desirable traits on a linkage map. Such maps can also help to reveal structural chromosome variants (Hippolyte et al., 2010) and can be combined with genomic libraries to characterize and isolate selected gene clusters (Canto-Canché et al., 2007). In breeding practices, wild bananas and related *Musa* species can also be useful as donor species in pre-breeding programs. This germplasm can later be crossed with diploid or triploid cultivars to create new alloploid or triploid seedless cultivars (de Oliveira et al., 2001).

The most important diploid fertile accessions for banana improvement, are the ancestral *M. acuminata* and *M. balbisiana,* which occur all over Southeast Asia (Perrier et al., 2009; Volkaert 2018). Specific morphological traits that allow distinction between these two species are blotches on the pseudostem, petiole canal shape, hair on peduncle, pedicels-fruit length ratio, ovules rows configuration, bract habit and male flower color (Simmonds & Shepherd, 1955). The genetic diversity of *M. acuminata* is much higher than in *M. balbisiana* (Wong et al., 2002; Volkaert, 2011), resulting in a subdivision of *M. acuminata* into seven sub-species (Perrier et al., 2011; Simmonds & Shepherd, 1955), *i.e.* ssp. *acuminata,* ssp. *errans (*Blanco) RV Valmayor, ssp. *halabanensis* (Meijer) M Hotta, ssp. *malaccensis* (Ridl.) NW Simmonds, ssp. *microcarpa* (Becc.) NW Simmonds, ssp. *siamea* NW Simmonds and ssp. *truncata* (Ridl.). In addition, Nasution (1991) described fifteen additional varieties that were found in Indonesia (Figure 4). As to the smaller diversity of *M. balbisiana,* only four varieties were accepted in the Plant List (2013), *i.e.* var. *brachycarpa* (Backer) Häkkinen, var. *liukiuensis* (Matsum.) Häkkinen, var. *bakeri* (Hook.f.) Häkkinen and var. *dechangensis* (J.L.Liu & M.G.Liu) Häkkinen.

Part of the overall diversity is maintained in gene banks such as the International *Musa* Germplasm Collection (International Transit Centre, ITC), which is hosted by the Catholic University of Leuven (KUL) in Belgium. However, the larger part of diversity is maintained in *in-situ* collections such as the one at the Indonesian Fruits Research Institute (ITFRI, Indonesian Centre for Horticultural Research and Development (ICHORD)) in Solok, Sumatra, Indonesia. Besides, there is a plethora of undiscovered diverse germplasm in tropical forests awaiting deployment by breeders and researchers for banana improvement. The ITC has more than 1,500 accession in tissue culture, but only a small number of accessions of wild *M. acuminata* (N=91) and *M. balbisiana* (N=32) (http://www.promusa.org/ITC, Garming et al., 2010; Ruas et al.,

2017). Others research institutes, such as CARBAP and CIRAD also have limited wild banana collections (www.crop-diversity.org/mgis, Ruas et al., 2017). Nevertheless, these collections have been used as reference collections for breeding and many banana genetic studies (Volkaert, 2011). In Indonesia, wild banana collections are deposited at ITFRI-ICHORD (Ruas et al., 2017), Research Center for Biology (RCB-LIPI) (Poerba et. al., 2014) and the Bogor Botanical Gardens-Indonesian Institute of Sciences (Hapsari et al., 2015) with approximately 100 accessions in *in-situ* collections and hundreds of DNA collections. However, these collections are not well-studied and only partly accessible for public. So far, the ITC collection is the most accessed gene bank and is frequently used by banana researchers thereby significantly contributing to phylogeny studies (Perrier et al., 2009) and taxonomic studies at the species level (Gawel et al., 1992) or section level (Ude et al., 2002; Häkkinen, 2013).

Genetic patterns and phylogeny of wild bananas are of great interest for breeding, but unfortunately only a small number of wild accessions is maintained in the aforementioned banana collections and the greater diversity in local collections is not publicly accessible, thereby collectively presenting a serious bottleneck of detailed genetic studies. Volkaert (2011) also indicated that many haplotypes of cultivated bananas are not present in the above-mentioned collections. Besides, Li et al. (2013) indicated a high level of diversity across cultivated bananas in the ITC, suggesting the need for further exploration of wild bananas to trace the multiple alleles in cultivated bananas. Evidently, merging, exchanging or extending accessions from existing banana collections by systematically collecting more wild bananas from the Vavilov center of diversity in Southeast Asia are required for uplifting genetic and genomic studies and breeding programs (Sutanto et al., 2016; Sardos et al., 2018).

In short, modern banana improvement requires a thorough understanding of the genetic, cytogenetic and genomic context of banana germplasm which is inevitable for the understanding and exploration of diversity for the benefit of growers and consumers. This thesis is an attempt to contribute to this requirement.

Thesis outline

Chapter 1 introduces the aim of this thesis *i.e.* to obtain more insight into the rich genetic background of Indonesian banana germplasm. The genetic diversity of wild bananas in the center of diversity is huge and therefore, extended exploration to support breeding is urgently required to deal with contemporary threats and the lack of overall sustainability. The generated data will help to increase the genetic possibilities

and limitations for breeding and specifically to identify and map genes that control resistance to FWB.

Chapter 2 describes the genetic diversity of the Indonesian wild *M. acuminata* populations with a focus on the island of Sumatra. To this end, all provinces of Sumatra were explored to develop the most comprehensive collection of wild *M. acuminata*. All accessions were categorized by morphological descriptions and genetic diversity. The data deliver a highly diverse picture of *M. acuminata* on Sumatra comprising several haplotypes networking among wild *M. acuminata*.

Chapter 3 deals with the improvement of a chromosome preparation technique of PMCs in order to study meiotic stages of selected Indonesian banana germplasm. Bivalent configuration during diakinesis is common and was observed in all studied bananas. Bridge and lagging chromosomes were observed during anaphase I in some banana accessions, which indicated chromosome abnormalities in PMCs during meiosis. The chapter finishes with a comprehensive discussion on chromosome abnormalities and their impact on banana breeding.

Chapter 4 is an extensive analysis of a large reported translocation between chromosome 1 and chromosome 4 in *M. acuminata* ssp. *malaccensis* based on new cytogenetic and molecular data.

Chapter 5 focuses onthe generation of mapping populations and subsequent phenotyping and genotyping to map genes for resistance to FWB caused by Race 1 and TR4 strains. The developed *M. acuminata* ssp. *malaccensis* population was analyzed with a single nucleotide polymorphisms (SNPs) approach and genotyping-by-sequencing (DArT-seq) for high density maker generation. This enabled mapping quantitative trait loci (QTLs) harboring genes for resistance to Race 1 and TR4 on the distal part of chromosome 10.

Finally, **Chapter 6,** provides an overall evaluation of the results obtained in this thesis and places the data in the wider perspective of breeding and overall sustainability of banana production. In conclusion, this study underscores that Indonesia is a rich resource for banana diversity with very potential genes for resistance to FWB and likely also a plethora of other genes for managing many other biotic threats and abiotic stress factors. This is very important for future banana improvement.

Abbreviations
AFLP: Amplified Fragment Length Polymorphism
CRISPR/Cas: Clustered Regularly Interspaced Short Palindromic Repeat/CRISPR-Associated
PCR: Polymerase Chain Reaction
ITS: Internal Transcribed Spacer
RFLP: Restriction Fragment Length Polymorphism

References

Adeleke MT V, Pillay M & Okoli BE (2002) An improved method for examining meiotic chromosomes in *Musa* L. HortScience 37:959–961.

Aguayo J, Cerf I, Folscher AB, Fourrier-Jeandel C, Ioos R, Matthews MC, Mostert D, Renault C, Wilson V & Viljoen A (2020) First report of *Fusarium oxysporum* f. sp. *cubense* tropical race 4 (TR4) causing banana wilt in the Island of Mayotte. Plant Disease. https://doi.org/10.1094/PDIS-06-20-1196-PDN

Aguirre P (2016) The origin, versatility and distribution of azole fungicide resistance in the banana black Sigatoka pathogen *Pseudocercospora fijiensis*. Thesis. Wageningen University & Research

Allard R (1960) Principles of Plant Breeding. John Wiley & Sons, New York.

Amorim EP, Dos Santos-Serejo JA, Amorim VBO, Ferreira CF & Silva SO (2013) Banana breeding at embrapa cassava and fruits. Acta Horticulturae 986:171–176.

Arango REI, Diaz-trujillo C, Dhillon B & Aerts A (2016) Combating a global threat to a clonal crop : banana black Sigatoka pathogen *Pseudocercospora fijiensis* (Synonym *Mycosphaerella fijiensis*) genomes reveal clues for disease control. PLoS ONE Genet 12:1–36.

De Ascensao ARDCF & Dubery IA (2000) Biochemistry and cell biology panama disease: cell wall reinforcement in banana roots in response to elicitors from *Fusarium oxysporum* f. sp. *cubense* race four. Phytopathology 90:1173–1180.

Baker JG (1893) A synopsis of the genera and species of Museae. Annals of Botany os-7:189–222.

Bakry F, Carreel F, Jenny C & Horry J-P (2009) Genetic improvement of banana. breeding plantation tree crops: tropical species. (ed by SM Jain & PM Priyadarshan) 2nd edn. Springer, New York, NY, pp 3–50.

Bakry F, Horry J-P & Jenny C (2020) Making banana breeding more effective. Achieving sustainable cultivation of bananas Volume 2: Germplasm and genetic improvement. (ed by G Kema & A Drenth) Burleigh Dodds Science Publishing, Cambridge, UK,.

Barraza D, Jansen K, van Wendel de Joode B & Wesseling C (2011) Pesticide use in banana and plantain production and risk perception among local actors in Talamanca, Costa Rica. Environmental Research 111:708–717.

Blomme G, Dita M, Jacobsen KS, Pérez Vicente L, Molina A, Ocimati W, Poussier S & Prior P (2017) Bacterial diseases of bananas and enset: current state of knowledge and integrated approaches toward sustainable management. Frontiers in Plant Science 8:1290.

Blomme G, Eden-Green S, Mustaffa M, Nwauzoma B & Thangavelu R (2011) Major Diseases of Banana. Banana

Breeding. (ed by M Pillay & A Tenkuano) CRC Press, pp 85–120.

Buddenhagen IW (1977) Resistance and vulnerability of tropical crops in relation to their evolution and breeding. the genetic basis of epidemics in agriculture. (ed by PR Day) New York Acad. Sci., New York, pp 309–326.

Buddenhagen I (2009) Understanding strain diversity in *Fusarium oxysporum* f. sp. *cubense* and history of introduction of "tropical race 4" to better manage banana production. Acta Horticulturae 828:193–204.

Canto-Canché B, Guillén-Maldonado DK, Peraza-Echeverría L, Conde-Ferráez L & James-Kay A (2007) Construction and characterization of a bacterial artificial chromosome library of the causal agent of black Sigatoka fungal leaf spot disease of banana and plantain, *Mycosphaerella fijiensis*. Molecular Biotechnology 36:64–70.

De Capdeville G, Souza Júnior MT, Szinay D, Diniz LEC, Wijnker E, Swennen R, Kema GHJ & De Jong H (2009) The potential of high-resolution BAC-FISH in banana breeding. Euphytica 166:431–443.

Cheesman EE (1947) Classification of the Bananas : The genus *Musa* L . Kew Bulletin 2:106–117.

Chen YF, Chen W, Huang X, Hu X, Zhao JT, Gong Q, Li XJ & Huang XL (2013) Fusarium wilt-resistant lines of Brazil banana (*Musa* spp ., AAA) obtained by EMS-induced mutation in a micro-cross- section cultural system. Plant Pathology 62:112–119.

Chen J, Xie J, Duan Y, Hu H, Hu Y & Li W (2016) Genome-wide identification and expression profiling reveal tissue-specific expression and differentially-regulated genes involved in gibberellin metabolism between Williams banana and its dwarf mutant. BMC Plant Biology:1–18.

Christelová P, Valárik M, Hřibová E, De Langhe E & Doležel J (2011) A multi gene sequence-based phylogeny of the Musaceae (banana) family. BMC Evol Biol 11:103 (2011)

Conde-Ferráez L, Waalwijk C, Canto-Canché BB, Kema GHJ, Crous PW, James AC & Abeln ECA (2007) Isolation and characterization of the mating type locus of *Mycosphaerella fijiensis*, the causal agent of black leaf streak. Molecular Plant Pathology 8:111–120.

Conedera M, Krebs · P, Tinner · W, Pradella · M & Torriani · D (2004) The cultivation of *Castanea sativa* (Mill.) in Europe, from its origin to its diffusion on a continental scale. Veget Hist Archaeobot 13:161–179.

Crichton RR, Vezina A & Van den Bergh I (2016) An online checklist of banana cultivars. Acta Horticulturae 2:13–18.

Dale JL (1987) Banana bunchy top: an economically important tropical plant virus disease. Advances in Virus Research 33:301–325.

Dale JL, James A, Paul J-Y, Khanna H, Smith M, Peraza-Echeverria S, Gar-

cia-Bastidas F, Kema G, Waterhouse P, Mengersen K & Harding R (2017) Transgenic Cavendish bananas with resistance to Fusarium wilt tropical race 4. Nature Communications 8:1496.

Damodaran T, Kumar N, Kavino M & Damodaran T (2009) Breeding and evaluation of Musa hybrids resistant to *Fusarium oxysporum* f. sp. *cubense* race 1. Fruits 64:3–12.

Dépigny S, Noupadja P, Tomekpe K, Bonneviot P, Lescot T & Bakry F (2016) 'CARBAP K74': A triploid plantain-like hybrid designed to promote sustainable plantain-based cropping systems. Acta Horticulturae 1196:63–70.

Dita M, Barquero M, Heck D, Mizubuti ESG & Staver CP (2018) Fusarium wilt of banana: Current knowledge on epidemiology and research needs toward sustainable disease management. Frontiers in Plant Science 871:1–21.

Dodds K (1943) Genetical and cytological studies of *Musa*. Certain edible diploids. J Genet 45:113–138.

Dodds KS & Simmonds NW (1946) Genetical and cytological study of *Musa*; The formation of polyploid spores. Journal of Genetics 47:223–242.

FAO (2003) The world banana economy, 1985-2002. Rome.

FAO (2014) FAO Statistical Yearbook 2014 - Near East and North Africa Food and Agriculture. FAO, Cairo.

FAO (2020) FAOSTAT. http://fao.org/faostat.

Fortescue JA & Turner DW (2004) Pollen fertility in *Musa*: Viability in cultivars grown in Southern Australia. Australian Journal of Agricultural Research 55:1085–1091.

Fraser-Smith S, Czislowski E, Daly A, Meldrum R, Hamill S, Smith M & Aitken EAB (2016) Single gene resistance to *Fusarium oxysporum* f. sp. *cubense* Race 4 in the wild banana *Musa acuminata* subsp. *malaccensis*. Acta Horticulturae. 1114th edn. International Society for Horticultural Science (ISHS), Leuven, Belgium, pp 95–100.

Gaidashova S V, Uwimpuhwe B & Karamura EB (2008) Identification of banana varieties with resistance to nematodes in rwanda. African Crop Science Journal 16:27–33.

García-Bastidas F (2019) Panama Disease in Banana: Spread, Screen and Genes.Thesis. Wageningen University & Research

García-Bastidas F, Ordóñez N, Konkol J, Al-Qasim M, Naser Z, Abdelwali M, Salem N, Waalwijk C, Ploetz RC & Kema GHJ (2013) First Report of *Fusarium oxysporum* f. sp. *cubense* Tropical Race 4 Associated with Panama Disease of Banana outside Southeast Asia. Plant Disease 98:694.

García-Bastidas FA, Quintero-Vargas JC, Ayala-Vasquez M, Schermer T, Seidl MF, Santos-Paiva M, Noguera AM, Aguilera-Galvez C, Wittenberg A, Hofstede R, Sørensen A & Kema

GHJ (2019) First Report of Fusarium Wilt Tropical Race 4 in Cavendish Bananas Caused by *Fusarium odoratissimum* in Colombia. Plant Disease:PDIS-09-19-1922-PDN.

García-Bastidas FA, Van der Veen AJT, Nakasato-Tagami G, Meijer HJG, Arango-Isaza RE & Kema GHJ (2019) An improved phenotyping protocol for panama disease in banana. Frontiers in Plant Science 10:1–12.

Garming H, Roux N & Den I Van (2010) The impact of the *Musa* International Transit Centre-Review of its services and cost effectiveness and recommendations for rationalization of its operations. Bioversity International, Montpellier, France.

Gawel NJ & Jarret RL (1991) Chloroplast DNA restriction fragment length polymorphisms (RFLPs) in *Musa* species. Theor Appl Genet 81:783–786.

Gawel NJ, Jarret RL & Whittemore AP (1992) Restriction fragment length polymorphism (RFLP)-based phylogenetic analysis of *Musa*. Theoretical and Applied Genetics 84–84:286–290.

Gubbuk H, Pekmezci M, Onus AN & Erkan M (2004) Identification and selection of superior banana phenotypes in the cultivar dwarf cavendish using agronomic characteristics and RAPD markers. Pakistan Journal of Botany 36:331–342.

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

Hapsari LIA, Kennedy J, Lestari DAYU, Masrum A & Lestarini W (2017) Ethnobotanical survey of bananas (Musaceae) in six districts of East Java , Indonesia. 18:160–174.

Hapsari L, Lestari DA & Masrum A (2015) Album Koleksi Pisang Kebun Raya Purwodadi, Seri 1: 2010-2015. UPT Kebun Raya Purwodadi, Pasuruan, Indonesia.

Hernández R, Ramírez T, Noa-carrazana JC, Rodríguez-fernández R, Cañal MJ, Flores-estévez N, Corujo M, Noceda C, Ventura JC, Cienfuegos D, Caminos C, Carolinas T, Domingo S, Clara V, Veracruzanas C, Col N & Cp EZ (2007) Generation of five new *Musa* hybrids with resistance to black sigatoka and high yield Centro para la Transformación Agrícola Sostenible (CETAS), Universidad Instituto de Investigaciones en Viandas Tropicales. (INIVIT). Finca Laboratorio de Biotecnolog. American Journal of Agricultural and Biological Sciences 2:43–48.

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

Hippolyte I, Bakry F, Seguin M, Gardes L, Rivallan R, Risterucci A-M, Jenny C, Perrier X, Carreel F, Argout X, Piffanelli P, Khan IA, Miller RN, Pappas GJ, Mbéguié-A-Mbéguié D, Matsumoto T, De Bernardinis V, Huttner E, Kilian A, Baurens F-C, D'Hont A, Cote F, Courtois B & Glaszmann J-C (2010) A saturated SSR/DArT

linkage map of *Musa acuminata* addressing genome rearrangements among bananas. BMC Plant Biology 10:65.

Hohn T, Richert-Poggeler K, Staginnus C, Harper G, Schwarzacher T, Teo C, Teycheney P, Iskra-Caruana M & Hull R (2008) Evolution of integrated plant viruses. Plant virus evolution. (ed by M Roossinck) Springer, Heidelberg, pp 58–81.

Hooks CRR, Wright MG, Kabasawa DS, Manandhar R & Almeida RPP (2008) Effect of banana bunchy top virus infection on morphology and growth characteristics of banana. Annals of Appllied Biology 153:1–9.

Horry JP & Jay M (1988) Distribution of anthocyanins in wild and cultivated banana varieties. Phytochemistry 27:2661–2672.

Houbin C, Chunxiang X, Qirui F, Guibing H, Jianguo L, Zehuai W & Molina AB (2004) Screening of banana clones for resistance to fusarium wilt in China. Advancing banana and plantain R&D in Asia and the Pacific. pp 165–174.

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.

Hwang S-C & Ko W-H (2004) Cavendish Banana Cultivars Resistant to Fusarium Wilt Acquired through Somaclonal Variation in Taiwan.

IITA (2000) Improving Plantain and Banana Based. International Institute of Tropical Agriculture, Project 2, Annual Report. Ibadan.

IPGRI (1996) Descriptors for Banana (*Musa* spp.). IPGRI. Rome.

IPPC (2013) New banana disease found in Mozambique (*Fusarium oxysporum* f.sp. *cubense* Tropical Race 4). htpps://www.ippc.int. accessed on 9 December 2020

Israeli Y, Ben-Bassat D & Reuveni O (1996) Selection of stable banana clones which do not produce dwarf somaclonal variants during in vitro culture. Scientia Horticlturae 67:197-205

Jarret RL & Litz RE (1986) Isozymes as genetic markers in bananas and plantains. Euphytica 35:539-549

John Kress W & Specht CD (2006) The evolutionary and biogeographic origin and diversification of the tropical monocot order Zingiberales. Aliso 22: 621-632

Kayat F, Javed MA, Wah HY & Othman RY (2004) Identification of molecular markers for disese resistace genes to FOC in *Musa acuminata* ssp *malaccensis* for marker assisted selection. The 4th Annual Seminar of National Science Fellowship:40–44.

Kennedy J (2009) Bananas and people in the homeland of genus *Musa*: Not just pretty fruit. 7:19.

Khayat E (2020) Targeted improvement of Cavendish clones. Achieving sustainable cultivation of bananas Volume 2: Germplasm and genetic improvement. (ed by G Kema &

A Drenth) Burleigh Dodds Science Publishing, Cambridge, UK.

Khayat E & Ortiz R (2011) Genetics of important traits in *Musa*. Banana Breeding: Progress and Challenges. (ed by M Pillay & A Tenkouano) CRC Press, Boca Raton, FL, pp 71–83.

Kimunye JN, Muzhinji N, Mostert D, Viljoen A, van der Merwe AEB & Mahuku G (2020) Genetic diversity and mating type distribution of *Pseudocercospora fijiensis* on banana in Uganda and Tanzania. Phytopathology. https://doi.org/10.1094/PHYTO-04-20-0138-R

Koeppel D (2008) Banana: the fate of fruit that chasnged the world. Penguin Putman Inc, New York, NY.

Kumar P, Gupta VK, Misra AK, Modi DR & Pandey BK (2009) Potential of molecular markers in plant biotechnology. Plant Omics Journal Southern Cross Journals 2:141–162.

Kurz S (1865) Note on plantains of the Indian Archipelago. J Agric Hort Soc India 14:295–301.

Lédo A, Junior J, Silva S & Alberto C (2008) Banana Princesa. Embrapa.

Lentfer CJ (2009) Tracing domestication and cultivation of bananas from phytoliths : an update from Papua New Guinea. Ethnobotany Research and Applications 7:247–270.

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.

Linnaeus C (1753) Species plantarum. Impensis Laurentii Salvii, Stockholm.

Linnaeus C (1759) Systema Naturae. Impensis Laurentii Salvii, Stockholm.

Liu A-Z, Kress WJ & Li D-Z (2010) Phylogenetic Analyses of the Banana Family (Musaceae) Based on Nuclear Ribosomal (ITS) and Chloroplast (trnL-F) Evidence. Taxon 59:20-28

Magee C (1927) Investigation on the bunchy top disease of the banana. Bulletin of the Council for Scientific and Industrial Research. Vol. 30. HJ Green. Melbourne

Marín DH, Monte D, Produce F, Romero RA, Brands C, Guzmán M & Sutton TB (2003) Black Sigatoka: an increasing threat to banana cultivation. The American Phytopathological Society 87:208–222.

Martin G, Carreel F, Coriton O, Hervouet C, Cardi C, Derouault P, Roques D, Salmon F, Rouard M, Sardos J, Labadie K, Baurens F-C & D 'hont A (2017) Evolution of the Banana Genome (*Musa acuminata*) Is Impacted by Large Chromosomal Translocations. Molecular Biology and Evolution 34:2140–2152.

Maryani N, Lombard L, Poerba YS, Subandiyah S, Crous PW & Kema GHJ (2019) Phylogeny and genetic diversity of the banana Fusarium wilt pathogen *Fusarium oxysporum* f. sp. *cubense* in the Indonesian centre of origin. Studies in Mycology 92:155–194.

Maxem A (2019) CRISPR could save ba-

nanas from fungus. Nature 574:9.

Molina AB, Sinohin VO, Fabregar EG, Ramillete EB, Loayan MM & Chao CP (2016) Field resistance of Cavendish somaclonal variants and local banana cultivars to tropical race 4 of Fusarium wilt in the Philippines. Acta Horticulturae 1114:227–230.

Morán AJF (2013) Improvement of Cavendish Banana Cultivars through Conventional Breeding. Acta Horticulturae 986:205–208.

Narayanasamy P (2011) Assessment of variability in fungal plant pathogens. Microbial Plant Pathogens-Detection and Disease Diagnosis Volume 1: Fungal Pathogene. Springer, Doerdrect. pp 245–272.

Nasution RE (1991) A taxonomy study of the species *Musa acuminata* Colla with its intraspecific taxa in Indonesia. Memoir of Tokyo University of Agriculture Volume XXXII.

Nwakanma DC, Pillay M, Okoli BE & Tenkouano A (2003) Sectional relationships in the genus *Musa* L. inferred from the PCR-RFLP of organelle DNA sequences. Theoretical and Applied Genetics 107:850–856.

de Oliveira S, Manoel M, Alves E, Silveira J & Lima M (2001) Banana breeding program at Embrapa. Crop Breeding and Applied Biotechnology 1:399–436.

Ordonez N, Seidl MF, Waalwijk C, Drenth A, Kilian A, Thomma BPHJ, Ploetz RC & Kema GHJ (2015) Worse comes to worst: bananas and panama disease—When plant and pathogen clones meet bananas: their origin and global rollout. PLoS Pathogens 11(11): e1005197.

Ortiz R & Vuylsteke D (1994) Inheritance of black sigatoka disease resistance in plantain-banana (Musa spp.) hybrids. Theoretical and Applied Genetics 89:146–152.

Ortiz R & Vuylsteke D (1995) Inheritance of dwarfism in plantain (*Musa* spp., AAB group). Plant Breeding 114:466–468.

Özarslandan M & Akgül DS (2019) First Report of *Fusarium oxysporum* f. sp. *cubense* Race 4 Causing Fusarium Wilt Disease of Banana in Turkey. Plant Disease:PDIS-09-19-1881-PDN.

Pegg KG, Coates LM, Neill WTO, Turner DW, Dita M & Carpentier SC (2019) The Epidemiology of Fusarium Wilt of Banana. Frontiers in Plant Science 10:1–19.

Peraza-Echeverria S, Dale JL, Harding RM, Smith MK & Collet C (2008) Characterization of disease resistance gene candidates of the nucleotide binding site (NBS) type from banana and correlation of a transcriptional polymorphism with resistance to *Fusarium oxysporum* f.sp. *cubense* race 4. Molecular Breeding 22:565–579.

Perrier X, Bakry F, Carreel F, Jenny C, Horry JP, Lebot V & Hippolyte I (2009) Combining biological approaches to shed light on the evolution of edible bananas. Ethnobotany Research and Applications 7:199–216.

Perrier X, De Langhe E, Donohue M, Lentfer C, Vrydaghs L, Bakry F, Carreel F, Hippolyte I, Horry J-P, Jenny C, Lebot V, Risterucci A-M, Tomekpe K, Doutrelepont H, Ball T, Manwaring J, de Maret P & Denham T (2011a) Multidisciplinary perspectives on banana (*Musa* spp.) domestication. Proc Natl Acad Sci USA 108:11311–8.

Ploetz RC (2005) Panama disease: an old nemesis rears its ugly head Part 1: The beginnings of the banana export trades. The American Phytopathological Society 6.

Ploetz RC (2015) Management of Fusarium wilt of banana: A review with special reference to tropical race 4. Crop Protection 73:7–15.

Price NS (1994) Field trial evaluation of nematode susceptibility within *Musa*. Fundam. appl. Nematol 17:391–396.

Raboin LM, Carreel F, Noyer JL, Baurens FC, Horry JP, Bakry F, Montcel HT Du, Ganry J, Lanaud C & Lagoda PJL (2005) Diploid ancestors of triploid export banana cultivars: Molecular identification of 2n restitution gamete Donors and n gamete donors. Molecular Breeding 16:333–341.

Rashid K, Mamat M, Bakar A, Daran M, Nezhadahmadi A, Ruslan F & Kayat F (2013) Seed Progeny Population of Wild Banana *Musa acuminata* ssp. *malaccensis* for Fusarium Screening. Life Science Journal 10:671-680.

Retnoningsih A, Megia R & Hartana A (2011) Microsatellite Markers for Classifying and Analysing Relationship between Banana Cultivars in Indonesia Genetic. :153–160.

Rhodes P (1964) A new banana disease in Fiji. Commonwealth Phytopathological News 10:38–41.

Roux NS, Toloza A, Dolezel J & Panis B (2004) Usefulness of embryogenic cell suspension cultures for the induction and selection of mutants in *Musa* spp. Banana improvement: Cellular, Molecular Biology and Induced Mutations. (ed by S Jain & R Swennen) Science Publisher, New Hempshire, USA, pp 33–43.

Ruas M, Guignon V, Sempere G, Sardos J, Hueber Y, Duvergey H, Andrieu A, Chase R, Jenny C, Hazekamp T, Irish B, Jelali K, Adeka J, Ayala-Silva T, Chao CP, Daniells J, Dowiya B, Effa Effa B, Gueco L, Herradura L, Ibobondji L, Kempenaers E, Kilangi J, Muhangi S, Ngo Xuan P, Paofa J, Pavis C, Thiemele D, Tossou C, Sandoval J, Sutanto A, Paka GV, Yi G, Van Den Houwe 13 I, Roux N & Rouard M (2017) MGIS: managing banana (*Musa* spp.) genetic resources information and high-throughput genotyping data. Database: the journal of biological database and curation 2017: bax046. doi:10.1093/database/bax046

Sagot MP (1887) Sur Le Genre Bananier. Bulletin de la Société Botanique de France 34:328–330.

Salacinas M (2019) Spot on: Managing Panama Disease of Bananas in the Philippines. Thesis. Wageningen

University & Research

Sander JD & Joung JK (2014) CRISPR-Cas systems for genome editing, regulation and targeting. Nat Biotechnol 32:347–355.

Saraswathi MS, Uma S, Vadivel E, Durai P, Siva SA, Rajagopal G & Sathiamoorthy S (2011) Diversity analysis in Indian cooking bananas (*Musa*, ABB) through morphotaxonomic and molecular characterisation. Acta Horticulturae 897:123–132.

Sardos J, Christelova P, Cızkova J, Paofa J, Sachter-Smith GL, Janssens SB, Rauka G, Ruas M, Daniells JW, Dolezel J & Roux N (2018) Collection of new diversity of wild and cultivated bananas (*Musa* spp .) in the Autonomous Region of Bougainville , Papua New Guinea. Genetic Resources and Crop Evolution.

Scholthof KG (2003) One Foot in the furrow: linkages between agriculture , plant pathology , and public health. Annu. Rev. Public Health 24:153–174.

Shepherd K (1999) Cytogenetics of the genus *Musa*. International Network for the Improvement of Banana and Plantain, Montpellier, France.

Simmonds NW (1953) Segregation in some diploid bananas. Journal of Genetics 51:458–469.

Simmonds NW (1962) The evolution of the bananas. Longman, London, UK.

Simmonds NW & Shepherd K (1955) The taxonomy and origins of the cultivated bananas. Journal of the Linnean Society of London, Botany

55:302–312.

Ssali TR, Kiggundu A, Lorenzen J, Karamura E, Tushemereirwe W & Viljoen A (2013) Inheritance of resistance to *Fusarium oxysporum* f. sp. *cubense* race 1 in bananas. Euphytica 194:425–430.

Stover R (1958) Studies on Fusarium wilt ofbananas. II. Some factors influencing survival and saprophytic multiplication of F. oxysporum f. cubense in soil. Canadian Journal of Botany 36:311–324.

Stover RH (1962) Fusarial wilt (panama disease) of bananas and other *Musa* species. Fusarial wilt (panama disease) of bananas and other *Musa* species. Commonwealth Mycological Institute. Kew, UK

Stover R & Simmonds N (1987) Bananas: Tropical Agriculture Series. Inc., New York.

Su H, Hwang S & Ko H (1986) Fusarium wilt of Cavendish bananas in Taiwan. Plant Disease 70:814–818.

Sulistyaningsih LD & Wawo AH (2011) Kajian Etnobotani Pisang-pisang Liar (*Musa* spp.) Di Malinau, Kalimantan Timur. Biosfera 28:43–47.

Sutanto A, Edison HS, Riska, Nasution F, Hermanto C, Cizkova J, Hribova E, Dolezel J, Roux N, Horry JP, Daniells JW & De Langhe E (2016) Collecting banana diversity in eastern Indonesia. Acta Horticulturae 1114:19–25.

Sutanto A, Sukma D, Hermanto C & Sudarsono S (2014) Isolation and characterization of Resistance Gene Analogue (RGA) from Fusarium re-

sistant banana cultivars. Emirates Journal of Food and Agriculture 26:508–518.

Swennen R & Vuylsteke D (1990) Aspects of plantain breeding at IITA. Sigatoka leaf spot diseases of bananas. Proceedings of an international workshop. San Jose, 28 Mar-1 Apr 1989. (ed by R Fullerton & R Stova) INIBAP, Montpellier, pp 252–266.

Tenkouano A, Pillay M & Ortiz R (2011) Breeding Techniques. Banana Breeding: Progress and Challenges. (ed by M Pillay & A Tenkouano) CRC Press, Boca Raton, FL, pp 181–202.

Ude G, Pillay M, Nwakanma D & Tenkouano A (2002) Genetic Diversity in *Musa acuminata* Colla and *Musa balbisiana* Colla and some of their natural hybrids using AFLP Markers. Theoretical and Applied Genetics 104:1246–1252.

Valmayor R V., Jamaluddin SH, Silayoi B, Kusumo S, Danh LD, Pascua O & Espino RR (2000) Banana Cultivar Names and Synonyms in Southeast Asia. INIBAP-Asia and the Pacific Office. Los Banos

Viljoen A, Mostert D, Chiconela T, Beukes I, Fraser C, Dwyer J, Murray H, Amisse J, Matabuana EL, Tazan G, Amugoli OM, Mondjana A, Vaz A, Pretorius A, Bothma S, Rose LJ, Beed F, Dusunceli F, Chao C-P & Molina AB (2020) Occurrence and spread of the banana fungus *Fusarium oxysporum* f . sp . *cubense* TR4

in. S Afr J Sci 116:#8608

Volkaert H (2011) Molecular analysis reveals multiple domestications of edible bananas. Acta Horticulturae 897:143–152.

Waite BH & Stover RH (1960) Studies on fusarium wilt of bananas: vi. variability and the cultivar concept in *Fusarium oxysporum* f. sp. *cubense*. Canadian Journal of Botany 38:985–994.

Wilberforce T, Michael B, Moses N, Robooni T, Alex B, Ssali T, Jerome K, Jum L & Rony S (2014) Performance of Narita banana hybrids in the preliminary yield trial Uganda. IITA

Wilson GB (1946a) Cytological studies in the Musae. i. Meiosis in some triploid clones. Genetics 31:241–258.

Wilson GB (1946b) Cytological studies in the Musae. iii. Meiosis in some seedling clones. Genetics 31:483–493.

Wong C, Kiew R, Argent G, Set O, Lee SK & Gan YY (2002) Assessment of the validity of the sections in *Musa* (Musaceae) using AFLP. Annals of Botany 90:231–238.

Zohary D & Spiegel-Roy P (1975) Beginnings of Fruit Growing in the.

Zuo C, Deng G, Li B, Huo H, Li C, Hu C, Kuang R, Yang Q, Dong T, Sheng O & Yi G (2018) Germplasm screening of Musa spp. for resistance to F*usarium oxysporum* f. sp. *cubense* tropical race 4 (Foc TR4). European Journal of Plant Pathology 4:1–12.

Chapter 2.

Genetic diversity of wild *Musa acuminata* on Sumatra

Ahmad F, Poerba YS, Sulistyaningsih LD, Subandiyah S, Kema GHJ, de Jong H & Volkaert HA

Abstract

Indonesia is a major centre of origin of bananas. Therefore, the entire archipelago is of great interest and value for banana genetic diversity and consequently is a significant region to study genetic variation of wild relatives and the evolution of edible bananas. In this study, we report an extensive exploration of one of the major constituents of cultivated bananas, the wild *Musa acuminata* Colla across Sumatra. In total we studied the genetic diversity of 164 accessions at 164 locations in eight provinces of Sumatra and of 20 wild banana accessions in the collection of the Research Center for Biology (LIPI), originating from other regions of Indonesia using morphological characteristics and multilocus genotyping. By combining the geographical coordinates of the collected material with the morphological variation - separately used in a principal components analysis (PCA) - we developed a spatial multivariate analysis. Based on the morphological data we identified five types of wild *M. acuminata* on Sumatra that based on PCA were distributed over three major clusters. One of these cluster (partially) overlapped two other smaller clusters of diversity. Based on the collected information, we propose to merge the previously described var. *halabanesis* (Meijer) Nasution and var. *alasensis* Nasution back as one subspecies *halabanesis* (Meijer) Hotta. Moreover, we argue that the var. *sumatrana* (Becc.) Nasution should be considered as a subspecies taking into account the geographical distribution. The multilocus genotyping - using the taxonomically relevant genes alcohol dehydrogenase (ADH1), catalase (CAT2), granule-bound starch synthase (GBSS1) and isocitrate dehydrogenase (IDH1) - showed a high genetic diversity of *M. acuminata* often in admixed stands of different haplotypes across Sumatra. We discovered a number of new haplotypes and observed that the majority of captured haplotypes clustered mostly according to presumed subspecies. Based on the allele frequency, there was a significant deviation from a Hardy-Weinberg equilibrium with a high level of homozygosity, suggesting genetically isolated subpopulations in each subspecies. The discovered genetic diversity in the wild *M. acuminata* population across Sumatra expands the current knowledge and awaits is deployment in ongoing breeding efforts for this important crop.

Introduction

Bananas belong to the genus *Musa*, the largest in the family Musaceae in the order Zingiberales (Kress & Specht, 2006). The genus comprises 69 species that have been classified into two sections, Musa (2x=2n=22) and Callimusa (2x=2n=14/18/20) (Häkkinen, 2013). The centre of diversity of *Musa* is Southeast Asia, particularly on the islands of the Indonesian archipelago (Perrier et al., 2011). Two species in the section Musa, *Musa acuminata* Colla (A genome) and *M. balbisiana* Colla (B genome) are the key species in the origin of cultivated bananas (Stover & Simmonds, 1987). Intra- and interspecific hybridization lead to a multitude of edible bananas at diploid (AA, AB), triploid (AAA, AAB, ABB) and tetraploid (ABBB) level. The edible bananas were brought into cultivation and carried across the tropics, starting thousands of years ago likely along with major migration events   (Perrier et al., 2009; Pagani et al., 2016). In contrast to the cultivated bananas, wild banana fruits are filled with seeds and only have little pulp. They grow in the forests, along riversides, while some of them survive among plantations and abandoned fields. Musa acuminata is a morphologically highly diverse species. Currently, eight sub-species are recognised within *M. acuminata*, *i.e.,* ssp. *burmannica* NW Simmonds, ssp. *errans* (Blanco) RV Valmayor, ssp. *halabanensis* (Meijer) M Hotta, ssp. *malaccensis* (Ridl.) NW Simmonds, ssp. *microcarpa* (Becc.) NW Simmonds, ssp. *siamea* NW Simmonds, ssp. *truncata* (Ridl.) NW Simmonds and ssp. *banksii* (F. Muell.) NW Simmonds (Häkkinen & Väre, 2008; Hotta, 1989; N W Simmonds, 1956; Valmayor, 1998; The Plant List, 2010). In addition, Nasution (1991) described 15 varieties of *M. acuminata* in Indonesia alone, though some of them are indicated as synonymous with described subspecies (Häkkinen & Väre, 2008; (The Plant List, 2010) (the diversity of *M. acuminata* in infographic see Supplementary Figure 1). The other wild relative, *M. balbisiana* is less diverse with four described varieties, *i.e.*, var. *brachycarpa* (Backer) Häkkinen, var. *liukiuensis* (Matsum.) Häkkinen, var. *bakeri* (Hook.f.) Häkkinen and var. *dechangensis* (J.L.Liu & M.G.Liu) Häkkinen (The Plant List, 2010).

Presently, approximately 1,900 named varieties (cultivars) of banana are grown by farmers in backyard gardens, small fields or large scale plantations (Crichton et al., 2016, http://www.promusa.org/Banana+cultivar+checklist). Global banana production is estimated at 155 million tons/year, thereby ranking fourth in production after rice, wheat and corn (FAO, 2020). In spite of its importance for food security, banana breeding, research and development programs are very limited in private and public sectors. However, some international institutes and consortia have initiated research for banana improvement (Ruas et al., 2017; https://www.cropscience.bayer.com/innovations/seeds-traits/a/history-modern-banana).

The study of genetic diversity of cultivated bananas and their wild ancestors is important for conservation and breeding. Swangpol et al. (2007) used four chloroplast regions and Volkaert (2011) used four chromosomal gene regions to study the contribution of various wild bananas to the cultivars. Later, Li et al. (2013) reported haplotypes of cultivated bananas associated with wild bananas from Southeast Asia. However, most of these studies used only a limited number of samples from the centre of origin. Volkaert (in prep) traced haplotypes found in cultivated bananas back to their presence in wild germplasm using the alcohol dehydrogenase (ADH1), auxin response factor (ARF17), catalase (CAT2), granule-bound starch synthase (GBSS1) and isocitrate dehydrogenase (IDH1) genes in a multilocus genotyping approach. However, haplotypes present in several cultivars were not found among any of the wild banana accessions available in the International Transit Center (ITC) collection at the Katholieke Universiteit Leuven in Belgium. We, therefore, considered that an in-depth study of wild bananas from the centre of origin, particularly the western part of Indonesia would potentially add to the understanding of their diversity and phylogeny and their exchange with cultivated banana populations. Nasution (1991) reported the occurrence of five varieties of wild *M. acuminata* on Sumatra, which is one of the main islands in western Indonesia. To extend the exploration and collection of genetic diversity of wild bananas, we sampled *M. acuminata* at 164 collection sites across Sumatra and studied the diversity using morphological observations and multilocus genotyping. As a result of the extensive haplotype analysis we propose to merge two *M. acuminata* varieties into an existing subspecies and elevate another variety to the level of subspecies.

Materials and methods
Plant and data collection
Plant materials were collected during five week-long expeditions criss-crossing Sumatra island from North to South between March and June 2017. In wild populations with comparable morphology, samples were collected every 5-10 km, but we also included plants with different morphologies once these were identified. The leaf base, bud, fruits with seed and other tissues were collected and processed for preservation in the herbarium. Cigar leaves were collected and stored on ice for DNA isolation in the laboratory. For some accessions also the corm was dug out for planting in the live collection at the Cibinong Botanical Gardens. GPS locations were recorded for each sample and plant morphology was documented according to the "Descriptors for Banana" (IPGRI, 1996). Data for 20 other wild banana accessions originating from other regions of Indonesia that were maintained in the collection of the Research Center for Biology (LIPI) (supplementary Table 1) were added to capture the overall genetic and morphological variation of Indonesian bananas.

Data analysis
In order to determine and visualize the clustering of the collected accessions based on the multivariate morphological characters, a principal component analysis (PCA) was performed using ClustVis (Metsalu & Vilo, 2015). Accessions and characters with more than 25% missing data were excluded. Eventually, 88 accessions (54% of total) and 65 morphological characters were retained (Supplementary Table 3). To determine potential spatial (sub)structure among the accessions a principal coordinates of neighbour matrices (PCNM) (Borcard & Legendre, 2002) was conducted using the vegan library (Oksanen et al., 2020) in R (R Core Team, 2020) on the same data matrix along with the GPS coordinates.

DNA was isolated following the CTAB method (Syamkumar et. al., 2003) with modification, *i.e.*, an addition of 0.1 g Polyvidone 25 (PVP) (Merck, Cat no. 1.07443) in 800 µL extraction buffer containing 4% Cetyltrimethylammonium bromide **(**CTAB) (Merck, Cat no. 219374). DNA was also extracted from freeze-dried leaves in the LIPI collection using the Promega Wizard® Magnetic DNA Purification System (Promega, Madison, USA) or Geneaid™ DNA Isolation Kit (New Taipei City, Taiwan), following the manufacturer's instructions.

For the population genetic study DNA fragments of four nuclear protein coding regions (ADH1, CAT2, GBSS1 and IDH1) were PCR amplified followed by sequencing of the products using internal primers (Table 1). The PCRs were performed in a 15 µL reaction volume using 0.075 unit/µL GoTaq® G2 DNA Polymerase (Promega, Madison, USA) with 32 µM dNTPs, 0.53 µM each primers, 1x GoTaq® Flexy 2 (Promega, Madison, USA) buffer and 1.5 – 2.0 mM $MgCl_2$ following the GoTaq® G2 DNA Polymerase manual. The pre-denaturation for 4 minutes at 94°C was followed by 40 cycles of 45 seconds at 94°C, 45 seconds at 53°C, 90 seconds at 72°C and a final elongation at 72°C and the resulting amplicons were checked by agarose gel electrophoresis. Based on the geographical and morphological diversity we selected 75 representative samples of wild *M. acuminata* accessions together with 20 wild *M. acuminata* accessions and 2 diploid AA cultivars from the LIPI collection (Supplementary Table 1), which were sent for sequencing (Macrogen Europe B. V., Amsterdam, The Netherlands).

**Table 1** Primers for PCR and DNA sequencing in this study

Gene	PCR primer	Sequencing primer	Functional association
ADH1	F: 5-TTTTGGGAAGCCAAGG-TAGGTG -3	F: 5-CCATGAGGCAGCAGGG-TA -3	Alcohol dehydrogenase
	R: 5-AACGCAGATATCATGG-CGTCGAT -3	R: 5-TTACCTTCTTCAAATCTC -3	
CAT2	F: 5- ACCAGGAGTA-CAAACCCCTA -3	F: 5- ACTGTTATTCATGAGCG -3	Catalase
	R: 5-CGGAATAAGAGAAAAT-TCTGGT -3	R: 5-CCACTCCGGGTAGT-TTCC -3	
GBSS1	F: 5-AGCGTGCAGGTTGAGG-TATTGC -3	F: 5-GTTCGCTTCTTC-CACTGCTA -3	Granule-bound starch synthase
	R: 5-GTCGTACCTGCATGGAA-CACATC -3	R: 5-AAGAAGACCAGTGTGC-CA -3	
IDH1	F: 5-GATTTCTTAGCTCAAG-GTGAG -3	F: 5-AATTCTATGACTTATGG-CA -3	Isocitrate dehydrogenase
	R: 5-GAACTCTTCGGTGT-TCAGATAC -3	R: 5-ATCTCGAGTAACCCT-GAGT -3	

Sequencing electrophoretograms were displayed in ChromasLite (Technelysium, http://technelysium.com.au/wp/chromas/) and manual rescoring of the polymorphic base calls was done when necessary. Haplotypes were reconstructed from diploid sequences in DnaSP version 6.12.01 (Librado & Rozas, 2009) using the PHASE algorithm (Stephens et al., 2001). Diploid sequences with insertion-deletion (indel) polymorphisms were separated using Indelligent v.1.2 (Dmitriev & Rakitov, 2008, http://dmitriev.speciesfile.org/indel.asp). Subsequently the haplotype networks were constructed using the statistical parsimony approach in TCS v2.1 (Clement et al. 2000). Indels with more than one nucleotide were reduced to single positions. Haplotypes from previous research comprising wild and cultivated bananas were included as references (Volkaert, in prep). The expected heterozygosity, fixation index and test for deviations from Hardy-Weinberg (HWE) were calculated using GenAlEx 6.5 (Peakall and Smouse 2006, 2012).

Results

A total of 164 wild *M. acuminata* accessions was collected (supplementary Table 2) across seven provinces of Sumatra (Figure 1). During the expedition it became already clear that most of the wild bananas could be tentatively grouped in three types (I, II and III) based on morphological characters. Type I was found below 750 m in most areas of the island, had a red inflorescence bud and angular seed. Type II was found in lowland

areas and mountain valleys below 1,000 m, mostly on the western half of Sumatra, had a purplish inflorescence bud and small globular seed with a smooth surface. Meanwhile type III grew in mountainous areas at an altitude of 750 m up to 1,400 m, had also purple inflorescence bracts and generally longer fruits containing flat seed with rough edges. We also identified intermediate phenotypes deviating from the morphology of the widely distributed Types I, II and III. At a few locations on Southern Sumatra, we identified plants with a morphology generally similar to Type I except a markedly longer petiole (usually up to half of the leaf blade). Henceforth, we refer to these plants as Type IV. Finally, we identified a population of wild bananas with several distinct characteristics of the male bud and fruit at very high elevations (1,200 – 1,500 m) in the centre of Aceh province, which we characterize as Type V (Supplementary Figure 2).
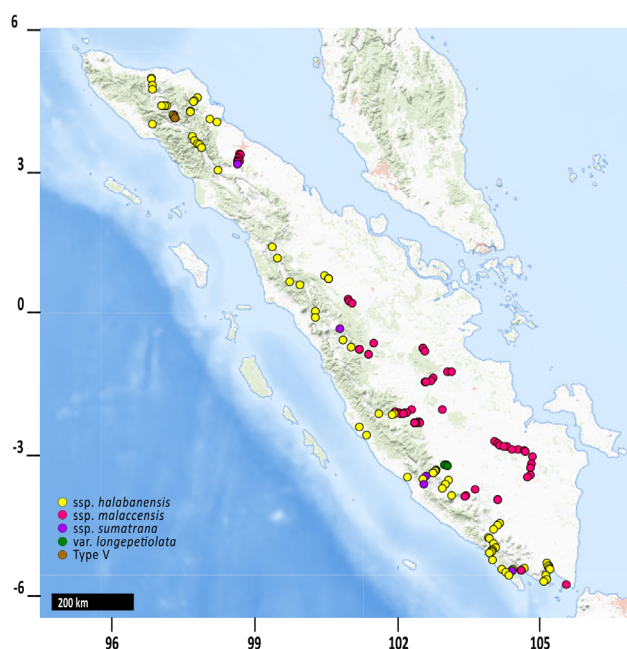
Figure 1 Collecting sites of 164 wild *Musa acuminata* accessions on Sumatra. Dots represent the collection sites and colours indicate subspecies/varieties found at the site (legend inside Figure).

General morphology of Musa acuminata on Sumatra
The sampled *M. acuminata* accessions conform to the following general description. Plants suckering freely; mature pseudostem height 2–5 m (rarely >5 m), slender to normal, coloured green yellow to green or green-brown to black purple or dark brown; with or without pigmentation, shiny or dull; sap watery or milky. Petiole with small to large blotches at base; petiole canal can be closed, overlapping, open with margin spreading, wide with erect margins or straight with erect margins; lamina truncated at top. Male bud colour red to purple with very little to abundant wax, shaped like a

top to rounded, male flower colour white to creamy, with abundant pollen. Fruit bunch horizontal, at an angle or vertical. Fruit with numerous seeds, flat or angular or globular in shape; smooth or wrinkled. *M. acuminata* grows in open areas along rivers, forest, palm plantations, along roadsides, rubber plantations or close to villages. Below the key identification and diagnostic description to describe *M. acuminata* on Sumatra:

Key to the infraspecific taxa of Musa acuminata on Sumatra

1 a.   Length of petiole less than half of the lamina length 2
   b.   Length of petiole half or more of the lamina length ..... var. *longepetiolata*
2 a.   Male bud ovoid, like a top or lanceolate, tip not imbricate 3
   b.   Male bud rounded, tip imbricate ....type V
3 a.   Seed angular 4
   b.   Seed globular, smooth ....ssp. *halabanensis*
4 a.   Seed wrinkled ....ssp. *malaccensis*
   b.   Seed flat with rough edge ....ssp. *sumatrana*

Diagnostic description of the infraspecific taxa of Musa acuminata on Sumatra
Musa acuminata subsp. *malaccensis* (Ridl.) N.W. Simmonds (Type I)
Musa malaccensis Ridl., Transactions of the Linnean Society of London, Botany ser. II. 3: 385 (1893). — *Musa acuminata* Colla ssp. *malaccensis* (Ridl.) N.W. Simmonds, Kew Bulletin 11: 466 (1957, "1956").

Musa acuminata Colla var. *malaccensis* (Ridl.) Nasution, Memoirs of the Tokyo University of Agriculture 32: 75 (1991). — Type: Malaya, Pahang, Tanjong Gajah Mati, 1891, Ridley s.n. (lecto-, SING 062891).

Diagnostic characters. Male bud like a top, bract shape medium to large shoulder, apex slightly pointed or pointed, slightly waxy, outer bract overlapping younger ones at apex of bud, external bract red to red purple sometimes with yellow streaks, internal bract whitish to red, apex acute, sometimes tinted with yellow, lifting one or two or more bracts at a time, revolute before falling. Seeds angular and wrinkled. Details of the morphotaxonomy described in Supplementary descriptions.

Distribution. This subspecies was found in North Sumatra, Jambi, Riau, Bengkulu, South Sumatra and Lampung province in habitats at forest borders at open areas, open canyons, river beds, villages, palm oil and rubber plantation borders and along roadsides.

Specimens examined. G. Koerintji, Bunnemeyer 10346 (3 sheets) (BO); Malaccensis: Tanjong Gajah Mati, 1891, H.N. Ridley s.n. (Holo - SING); Palembang, 2 November 1929, C.G.J. van Steenis 3950 (4 sheets) (BO). Asahan, Sumatra, H.S. Yates 1690 (2 sheets) (BO); Mount Kerintji, 4 April 1954, A.H.G. Alston 14236 (1 sheet) (BO); Jl. H. Makam, Kota Bongko, Jambi, 12 April 2017, F. Ahmad FA-J 12 (4 sheets) (BO); Ds. Muara Kilis, Tebo Hilir, Muara Tebo, Jambi, 14 April 2017, F. Ahmad FA-J 29 (3 sheets) (BO); Jejawi, OKI, South Sumatra, 25 April 2017, F. Ahmad F-SSB 10 (4 sheets) (BO); Baru Raja, Prabumulih, 18 May 2017, F. Ahmad & Y.S. Poerba AP-SSL 3 (4 sheets) (BO); Ds. Siabu, Salo, Bangkiang, Kampar, Riau, 8 May 2017, F. Ahmad & Y.S. Poerba AP-RSB 01 (4 sheets) (BO); Kuantan Mudik, Riau, 12 May 2017, F. Ahmad & Y.S. Poerba AP-RSB 21 (4 sheets) (BO); Ds. Seberida, Batang Dangkal, Inhu, Riau, 13 May 2017, F. Ahmad 7 Y.S. Poerba AP-RSB 22 (5 sheets) (BO); Jl. Raya Pesisir, Desa Banding, Raja Basa, Lampung Selatan, 22 May 2017, F. Ahmad & Y.S. Poerba AP-SSL 30 (4 sheets) (BO).

Musa acuminata ssp. *halabanensis* (Meijer) Hotta (Type II)

Musa halabanensis Meijer, Acta Botanica Neerlandica 10: 250 (1961). — *Musa acuminata* Colla ssp. *halabanensis* (Meijer) M. Hotta, Occasional Papers, Kagoshima University Research Center for the South Pacific 16: 68 (1989). — *Musa acuminata* Colla var. *halabanensis* (Meijer) Nasution, Memoirs of the Tokyo University of Agriculture 32: 51-53 (1991). — Type: Indonesia, North Sumatra, Pematang Siantar, 9.III.1958, W. Meijer 7264 (holo-, L).

Musa acuminata Colla var. *alasensis* Nasution, Memoirs of the Tokyo University of Agriculture 32: 48-51, Fig.19 (1991). — Type: Indonesia, Aceh, Kotacane, Ketambe, Rusdy 1638 (holo-, BO).

Diagnostic characters. Male bud ovoid; bract shape medium shoulder, apex mostly intermediate or obtuse, moderately waxy, outermost bract overlaps at apex of bud or sometimes slightly imbricate, external bract purple–brown, internal bract whitish, apex obtuse, not tinted with yellow colour, lifting one or two or more bracts at a time, revolute before falling. Initial hand or hands of the bunch generally partly or not developed. Seeds globular, mostly smooth. Details of the morphotaxonomy described in Supplementary descriptions.

Distribution. This subspecies was found in most provinces of Sumatra in habitats at forest borders, open areas, open canyons, river beds, villages, palm oil and rubber plantation borders and along roadsides.

Specimens examined. Around Anai waterfall, west of Padangpanjang, West Sumatra,

11 January 1981, M. Hotta 26153 (4 sheets) (BO); Bt. Gajabuih, Ulu Gadut, about 15 km east Padang City, West Sumatra, 11 February 1981, M. Hotta & H. Okada 60 (2 sheets) (BO); Along valley of Ulu Gadut, about 10 km east Padang City, West Sumatra, 24 January 1981, M. Hotta 26128 (1 sheet) (BO); Eastern foot of Gunung Sago, south of Payakumbuh, West Sumatra, 28 August 1983, M. Hotta 26599 (1 sheet) (BO); Gunung Talang, Laras Talang, Sumatra, 1919, Bunnemeyer 5559 (1 sheet) (BO); Bt. Gajabuih, Ulu Gadut, about 15 km east Padang City, West Sumatra, 6 January 1981, M. Hotta 25853 (2 sheets) (BO); Ds. Lawe Mangkudu, Ketambe, Southeast Aceh, 26 February 2013, Y.S. Poerba & K.U. Nugraheni PN 17 (1 sheet) (BO); Desa Sukamaju, Pesisir Barat, Kab. Pesisir Barat, Lampung, 20 May 2017, F. Ahmad & Y.S. Poerba AP-SSL 17 (3 sheets) (BO);

Musa acuminata Colla ssp. *sumatrana* (Becc.) Ahmad, Volkaert, Sulist. & Poerba (Type III)

Musa sumatrana Becc., in André, L'Illustration Horticole 27: 37 (1880). — *Musa acuminata* Colla var. (Becc.) Nasution, Memoirs of the Tokyo University of Agriculture 32: 80 (1991). — Type: Indonesia, Sumatra, Padang, Aug. 1878, O. Beccari 489 (lecto-, K000292216, K000292218).

Homonym. *Musa sumatrana* Ridl., Bulletin of Miscellaneous Information Kew 1926: 90 (1926).

Diagnostic characters. Male bud like a top or lanceolate; bract shape small or medium to large shoulder, apex slightly pointed or intermediate, moderately waxy, convolute, old bract overlap at apex of bud, external bract purple, internal bract red with yellow pigmentation, apex slightly pointed or intermediate, not tinted with yellow colour, lifting one or two or more bracts at a time, revolute before falling. Seeds angular flat with rough edge.

Description. Plant suckering freely. Mature pseudostem 3–6 m tall, 15 up to 30 cm in diam. at base, covered with dead remains of older sheaths. Sheaths purplish or blackish, underlying colour cream with pink-purple pigmentation, very little to moderately waxy; sap milky. Petiole 40-50 cm long, to 2-3 cm in diam., petiole canal leaf straight with erect margins or margins curved inward to margins overlapping, petiole margin colour purple to blue. Leaves: lamina lanceolate, 150-250 × 40-60 cm, base slightly asymmetric to symmetric, both sides rounded or one side rounded or one side pointed, apex truncate, abaxial medium green with waxy, adaxial green not waxy; colour of abaxial midrib yellow green with pink pigmentation, colour of adaxial midrib light green. Inflorescence hanging vertically, bunch mostly very compact; peduncle 30-

40 cm long, to 8-10 cm in diam., green to dark green, pubescent. Rachis mostly with a curve, flowers hermaphrodite and followed with male flowers. Inflorescence bud like a top or lanceolate, to 4-5 × 12-15 cm. Bract shape small or medium to large shoulder, apex slightly pointed or intermediate, moderate waxy, 6-7 × 10-11 cm, convolute, outer bract overlap at apex of bud, external bract dark purple, internal bract red with yellow pigmentation, apex slightly pointed or intermediate, not tinted with yellow colour, lifting one or two or more bracts at a time, revolute before falling. Ovary cream. Ovules in 4 rows per locule. Male flower; compound tepals whitish, with 3 prominently thickened keels, light yellow. Free tepal translucent white, ovate, apex acute with developed sign. Fruit bunch with 20-25 hands, fruits in 2 row, 20-25 fruits. Individual fruit straight, up to 15 cm long (including pedicel), 1-1.5 cm in diam., apex bottle necked. Pedicel ca. 1 cm long, straight or straight in the distal part. Immature fruit peel green, not waxy, fruit pulp white, seed up 200 to per fruit, angular flat with rough edge, 7- 8.3 mm long, 5-6.5 mm wide 2-3 mm depth.

Distribution. This subspecies was found in Aceh, North Sumatra, West Sumatra, Bengkulu, and Lampung provinces at elevations of 750-1500 m asl at forest borders, open areas or at some shaded river beds, villages and along roadsides.

Specimens examined. Sungei Kering, Kerintji, 3 Maret 1954, A.H.G. Alston 14120 (2 sheets) (BO); Ds. Lae Pondom, Sumbul Pegagan, Dairi, 28 February 2013, Y.S. Poerba & K.U. Nugraheni PN 44c (6 sheets) (BO); Ds. Tiyang Layar, Pancur Batu, Deli Serdang, 1 May 2013, Y.S. Poerba & K.U. Nugraheni PN 73 (3 sheets) (BO); Ds. Ketangkuhen, Sibolangit, Deli Serdang, 1 May 2013, Y.S. Poerba & K.U. Nugrahen PN 67 (3 sheets) (BO), PN 68 (4 sheets) (BO); Desa Sikeben, Sibolangit, Deli Serdang, 1 May 2013, Y.S. Poerba & K.U. Nugrahen PN 65 (4 sheets) (BO).

Musa acuminata Colla var. *longepetiolata* Nasution (type IV)
Memoirs of the Tokyo University of Agriculture 32: 63, fig. 24 (1991). — Type: Indonesia, Palembang, Musi Rawas, June 1982, Rusdy 1630 (holo-, BO)
Diagnostic characters. In general, the habitus, bunch, male bud and flower were similar to subsp. *malaccensis*, but the petiole length at least half of its lamina. Details of the morphotaxonomy described in Supplementary descriptions.

Distribution. This subspecies was only collected in Bengkulu and South Sumatra in habitats at forest borders, open areas, villages and along roadsides.

Specimen Examines. Nasution's herbarium sheets in Herbarium Bogoriense are missing.
Accessions ANS 15-17 (type V)

Diagnostic characters. The recognizable characters of this type are the imbricate and rounded male bud, and the bract was not revolute before falling. The pedicel of the fruit is relatively long alike *M. balbisiana*. Details of the morphotaxonomy described in Supplementary descriptions.

Distribution. We found this variety at only three locations a few kilometres apart at high elevation (1500 m) in Aceh growing at forest borders and open areas and along roadsides.

Specimen examined. Pantau Cuaca, Kab. Gayo Luwes, Aceh, 22 March 2017, F. Ahmad, Y.S. Poerba & H.A. Volkaert ANS-17 (4 sheets) (BO).

Phenotypic analysis

The PCA of morphological data clearly separated the Types I, II, and V bananas. However, the cluster that contains Type I completely overlapped the cluster containing Type IV bananas and partially overlapped the cluster with Type III accessions (Figure 2). As the PCA cannot test the apparent association of genetic diversity with geographical origin, we conducted a spatial multivariate analysis, PCNM, which provided variation of Eigen values that represent the difference of morphological characters among accessions. Based on this, we identified sub-populations of ssp. *halabanensis* and ssp. *malaccensis* in West, Central and Eastern Sumatra (indicated by the differences of circles size that represent the Eigen value in Figure 3).

Molecular analysis

To evaluate the genetic diversity of wild *M. acuminata* on Sumatra, we obtained high quality sequence data from 69 accessions and added 20 accessions from the LIPI collection. The ADH locus included a highly polymorphic microsatellite repeat which had to be excluded from further analysis because the correct number of repeat units was impossible to determine in heterozygous individuals. Eventually, we obtained 910, 1156, 949 and 1105 nucleotide positions for ADH1, CAT2, GBSS1 and IDH1, respectively (Table 2). All sequences will be deposited at GenBank. We observed 83 substitutions (one per 24 bases) and four indels (1 per 489 position) in the coding regions and 252 substitutions (one per 9 bases) and 19 indels in the non-coding regions (one per 114 position) (Table 2). The indel sizes ranged from single nucleotide to 276 bp long.
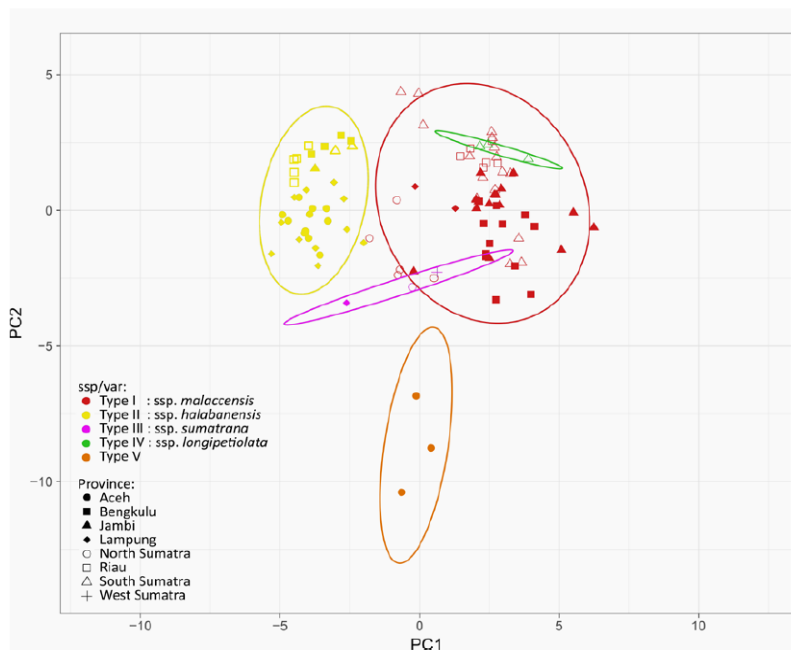
Figure 2 Principal component analysis of the morphological characters of 88 wild *Musa acuminata* accessions from Sumatra with less than 25% missing data.
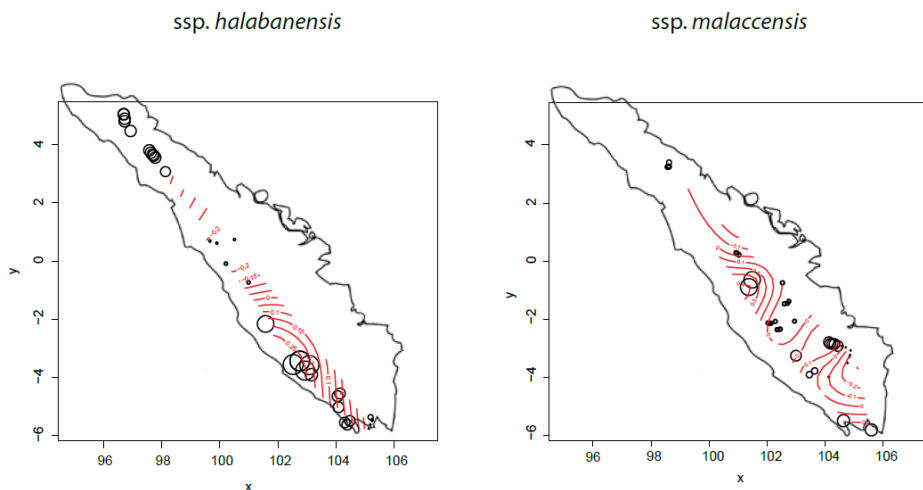
Figure 3 A spatial multivariate analysis (eigen vector calculated using principal coordinate analysis of neighbour matrices, PCNM) of two wild *Musa acuminata* subspecies on Sumatra according to morphological characters. Dot size is proportionate to the Eigen values that represent the difference of morphological characters among accessions. x: latitude, y: longitude.

Table 2 Frequency of substitution and insertion/deletion in exons and introns in sequencing reads of four loci.

	Substitution		Insertion/deletion		Nucleotides		
	Exon	Intron	Exon	Intron	Exon	Intron	Total
ADH1	25	36	1	1	494	416	910
CAT2	23	76	2	3	666	490	1156
GBSS1	18	58	1	7	494	501	995
IDH1	17	82	0	8	301	804	1105
Total	83	252	4	19	1955	2211	4166

Overall, the sequenced accessions, combined with the references, resulted in 53, 71, 98 and 62 haplotypes for ADH1, CAT2, GBSS1 and IDH1, respectively (Table 3). In total we identified 141 haplotypes among the *M. acuminata* accessions sampled from Sumatra, of which 103 were newly detected haplotypes, 15, 26, 39 and 23 for ADH1, CAT2, GBSS1 and IDH1, respectively. The number of singleton haplotypes (found only once in the whole dataset) was large, 5, 9, 19 and 8 for ADH1, CAT2, GBSS1 and IDH1, respectively, or almost one third of all haplotypes.

Even though the number of individuals sampled per subspecies was relatively small, certainly when compared to the large number of haplotypes detected, population genetic analysis in GenAlEx revealed some interesting trends. The three main subspecies on Sumatra, viz. ssp. *halabanensis*, ssp. *malaccensis* and spp. *sumatrana*, contained almost discrete haplotype sets for the four loci as indicated by the large number of private alleles, indicating a high level of population differentiation. Surprisingly though, in spite of the large number of alleles there is a high level of homozygosity, hence the observed heterozygosity is lower than expected (Table 4) for each locus resulting in highly significant deviations from a Hardy-Weinberg equilibrium (Table 5).

Table 3 Number of accessions and number of new haplotypes and similar haplotypes to references detected at four loci.

	Accessions		Haplotypes			
Locus	# wild Sumatra accessions	# LIPI accessions	new haplotypes	Haplotypes from Sumatra similar to references	References	Total
ADH1	57	12	15	8	38	53
CAT2	62	19	26	14	45	71
GBSS1	65	16	39	7	59	98
IDH1	67	14	23	9	39	62
		Total	103	38	181	284

Table 4 General population genetic information of ssp. *malaccensis*, ssp. *sumatrana* and ssp. *halabanensis* from Sumatra including the number of populations (N), number of different alleles (Na), number of effective alleles (Ne), observed heterozygosity (Ho), expected heterozygosity (He), unbiased expected heterozygosity (uHe) and fixation index (F).

Pop		ADH1	CAT2	GBSS1	IDH1
ssp. *halabanensis*	N	24	29	31	31
	Na	5	15	21	14
	Ne	1.49	4.847	11.509	6.45
	I	0.721	2.116	2.741	2.128
	Ho	0.125	0.517	0.645	0.323
	He	0.329	0.794	0.913	0.845
	uHe	0.336	0.808	0.928	0.859
	F	0.62	0.348	0.293	0.618
ssp. *malaccensis*	N	22	24	25	26
	Na	10	14	19	8
	Ne	7.224	9.681	13.441	4.829
	I	2.092	2.447	2.746	1.755
	Ho	0.409	0.708	0.72	0.5
	He	0.862	0.897	0.926	0.793
	uHe	0.882	0.916	0.944	0.808
	F	0.525	0.21	0.222	0.369
ssp. *sumatrana*	N	8	6	6	7
	Na	3	5	5	6
	Ne	1.662	3.13	3.13	5.444
	I	0.703	1.358	1.358	1.748
	Ho	0.25	0.167	0.167	0
	He	0.398	0.681	0.681	0.816
	uHe	0.425	0.742	0.742	0.879
	F	0.373	0.755	0.755	1

Haplotype networks generated from the four genes clustered the accessions more or less according to presumed subspecies (Figure 4). Haplotypes of ssp. *halabanensis*, ssp. *truncata*, and ssp. *banksii* grouped fairly well for all genes. In contrast, at all loci, the haplotypes of ssp. *malaccensis* and ssp. *zebrina* populated a large central area of the networks. Haplotypes found in the ssp. *sumatrana* accessions were scattered among those of ssp. *halabanensis* and ssp. *malaccensis* aside from some subspecies-specific haplotypes. The accession ANS-15, which was morphologically described as Type V, has unique haplotypes for all tested loci. However, the haplotype (A68) of another Type V accession, ANS-17, was shared with ssp. *halabanensis* and ssp. *sumatrana* based on the sequence of CAT2, as the other sequences failed. The haplotypes of all accessions

are shown in Table 6. In total, 55 of 69 wild *M. acuminata* accessions from Sumatra were heterozygous for at least one locus. Some accessions had haplotypes from two very distinct clusters (Figure 4) which might indicate admixtures among wild *M. acuminata* through cross pollination between subspecies. Among the accessions from Sumatra, the observed homozygosity at ADH1, CAT2, GBSS1 and IDH1 was 72%, 45%, 38% and 64%, respectively.

Table 5 The level of significance from deviations of Hardy-Weinberg (HWE) of ssp. *malaccensis*, ssp. *sumatrana* and ssp. *halabanensis* populations on Sumatra. Significance level indicated by asterisk * for P<0.05, ** for P<0.01, *** for P<0.001 and ns is not significant.

Populations	Loci	DF	ChiSq	P	Significance
ssp. *halabanensis*	ADH1	10	48.142	0.000	***
ssp. *halabanensis*	CAT2	105	184.308	0.000	***
ssp. *halabanensis*	GBSS1	210	297.585	0.000	***
ssp. *halabanensis*	IDH1	91	178.767	0.000	***
ssp. *malaccensis*	ADH1	45	127.373	0.000	***
ssp. *malaccensis*	CAT2	91	160.467	0.000	***
ssp. *malaccensis*	GBSS1	171	196.486	0.088	ns
ssp. *malaccensis*	IDH1	28	52.644	0.003	**
ssp. *sumatrana*	ADH1	3	2.889	0.409	ns
ssp. *sumatrana*	CAT2	10	24	0.008	**
ssp. *sumatrana*	GBSS1	10	24	0.008	**
ssp. *sumatrana*	IDH1	15	35	0.002	**

Figure 4 Network and distribution of studied haplotypes of the ADH1, CAT2, GBSS1 and ADH1 loci. The colours represent subspecies/varieties/cultivars, the size of the circles corresponds with haplotype frequency. Lines in the graph represent (a series of) mutational changes with open dots (o) indicating presumed intermediate haplotypes not detected in the studied accessions.
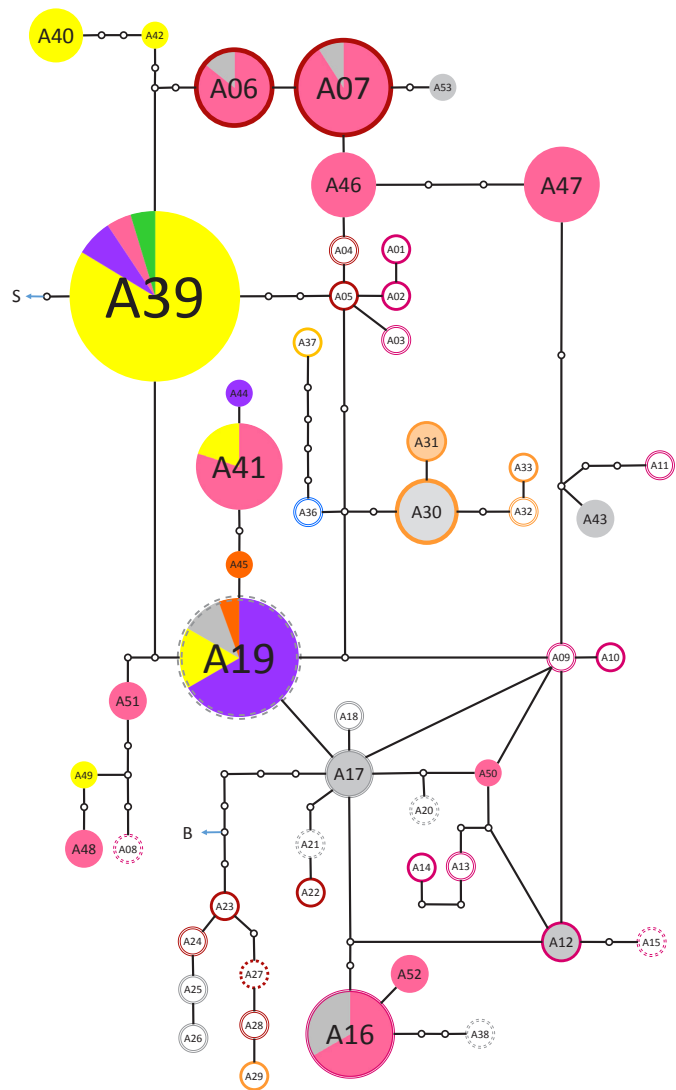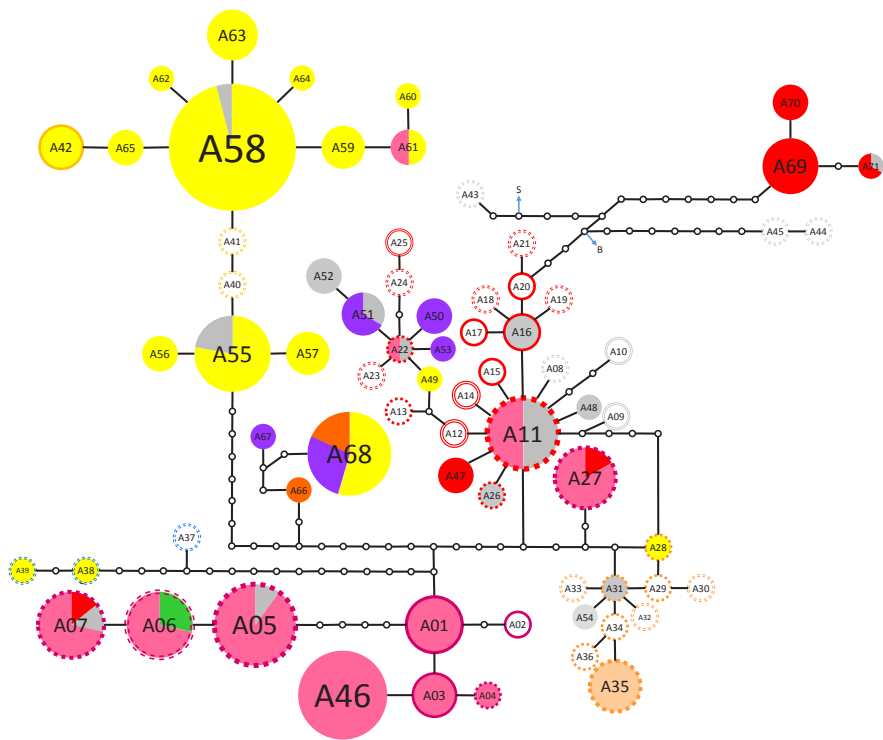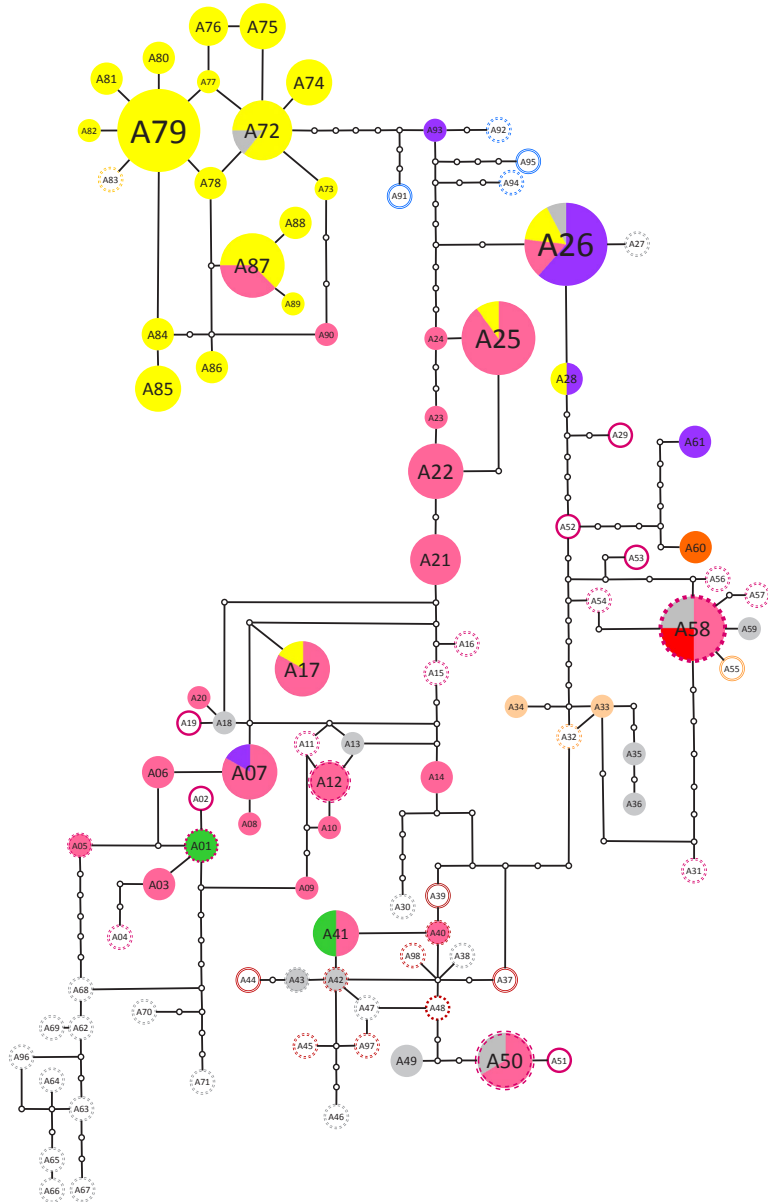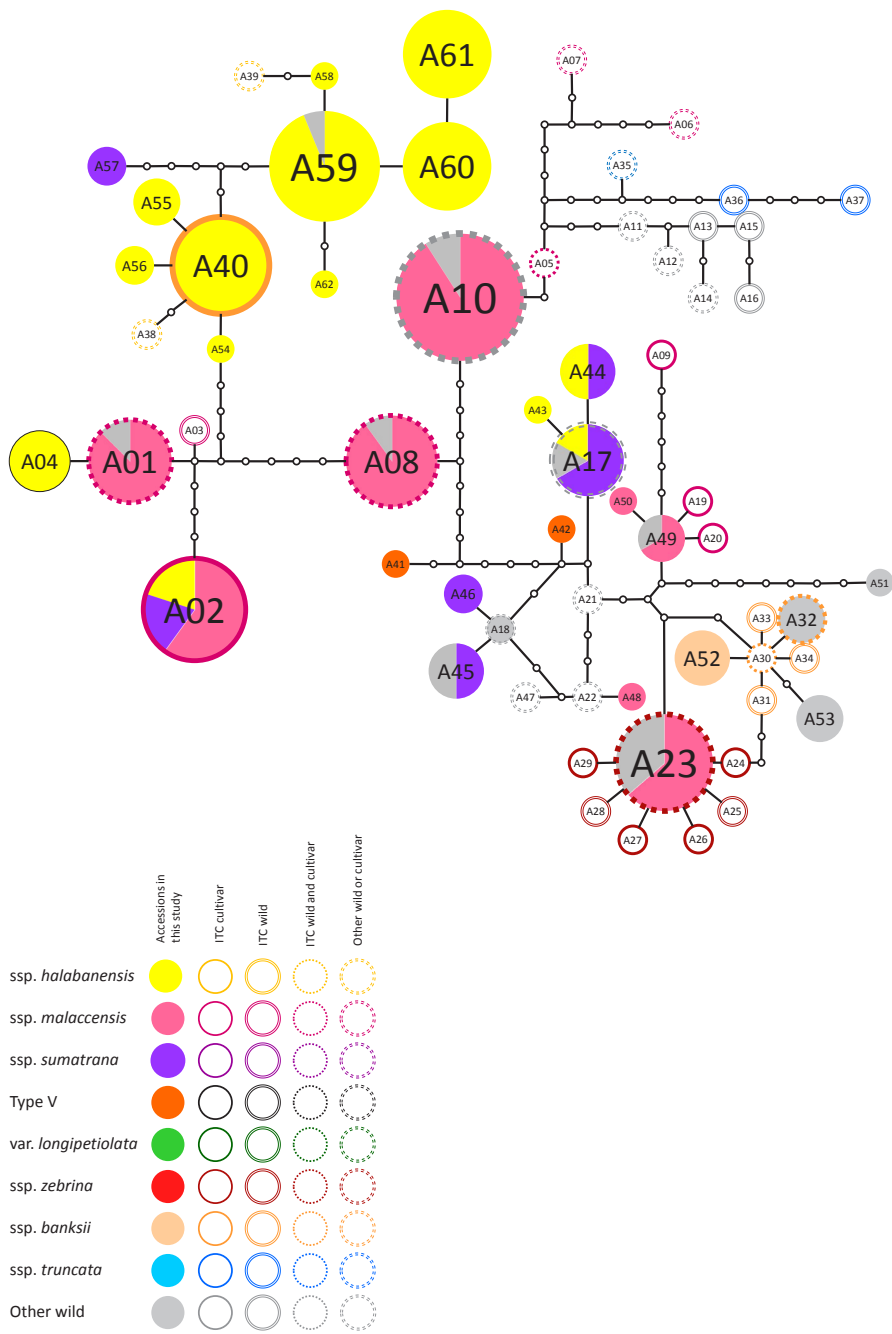
ADH

Figure 4 continues

CAT

GBSS

IDH

Table 6 List of accessions that were sent for sequencing and haplotypes based on four loci presented for each accession. Failed sequencing reactions were found in some accessions (na). Haplotypes identical to the references are indicated by an asterisk.

Accession	ssp/var/cv/type	ADH1		CAT2		GBSS1		IDH1	
ANS-1	*malaccensis*	A46	A47	A01*	A27*	A20	A25*	A10*	A10*
ANS-4	*sumatrana*	A19*	A44	na	na	A7	A93	A02*	A02*
ANS-5	*sumatrana*	A19*	A19*	A68	A68	A28	A28	A46	A46
ANS-8	*halabanensis*	A39	A39	A58	A59	A72	A73	A58	A58
ANS-9	*halabanensis*	na	na	A60	A61	A75	A75	A59	A59
ANS-15	Type V	A19*	A45	A22*	A66	A60	A60	A41	A42
ANS-17	Type V	na	na	A68	A68	na	na	na	na
ANS-20	*halabanensis*	A39	A39	na	na	A76	A88	A53	A59
ANS-22	*halabanensis*	A39	A39	A55	A58	A74	A88	A59	A59
ANS-23	*sumatrana*	A39	A19*	A53	A67	A61	A61	A56	A56
ANS-25	*halabanensis*	A40	A40	A58	A58	A74	A74	A54	A54
ANS-29	*halabanensis*	A40	A40	A58	A58	A74	A87	A54	A58
ANS-31	*halabanensis*	A41	A41	na	na	na	na	A02*	A02*
ANS-32	*halabanensis*	A39	A42	A55	A59	A75	A77	A40	A40
ANS-34	*halabanensis*	A39	A19*	A59	A68	A28	A75	A18*	A58
ANS-35	*halabanensis*	A39	A39	A58	A58	A72	A76	A40	A40
ANS-37	*malaccensis*	A41	A41	A69	A70	A05*	A06	A02*	A02*
ANS-38	*malaccensis*	na	na	na	na	na	na	na	na
ANS-39	*malaccensis*	A41	A41	A69	A70	A07	A07	A02*	A02*
ANS-41	*malaccensis*	A41	A41	A69	A69	A07	A08	A02*	A02*
J-1	*malaccensis*	A41	A41	A69	A71	A07	A07	A02*	A02*
J-4	*malaccensis*	A46	A47	A27*	A46	A03	A10	A10*	A10*
J-6	*malaccensis*	na	na	A6*	A46	A17	A22	A10*	A23*
J-8	*malaccensis*	na	na	na	na	na	na	A10*	A10*
J-12	*malaccensis*	A06*	A06*	A46	A46	A06	A87	A01*	A10*
J-14	*malaccensis*	na	na	A46	A46	A23	A87	A08*	A10*
J-16	AA wild	A19*	A19*	A22*	A22*	A26	A26	A45	A45
J-19	*halabanensis*	na	na	A57	A57	A86	A86	A40	A40
J-21	*halabanensis*	A39	A39	A42*	A57	A26	A87	A40	A40
J-23	AA wild	A43	A43	na	na	A18	A25*	A10*	A10*
J-26	*malaccensis*	A07*	A47	A46	A46	A25*	A87	A10*	A10*
J-27	*malaccensis*	A07*	A47	A05*	A27*	A22	A22	A08*	A10*
J-29	*malaccensis*	A06*	A50	A01*	A05*	A17	A17	A10*	A10*
J-33	*malaccensis*	A07*	A47	A03*	A03*	A03	A17	A10*	A48
RSB-1	*malaccensis*	A06*	A06*	A01*	A07*	A12*	A14	A02*	A10*
RSB-4	*halabanensis*	A39	A39	A58	A58	A87	A87	A58	A59
RSB-6	*halabanensis*	na	na	na	na	A76	A89	na	na
RSB-9	*halabanensis*	A39	A39	A58	A62	A72	A84	A58	A58

Table continues

Accession	ssp/var/cv/type	ADH1		CAT2		GBSS1		IDH1	
RSB-11	*halabanensis*	A39	A19*	na	na	A26	A87	A17*	A58
RSB-13	*sumatrana*	A19*	A19*	A68	A68	A26	A26	A44	A44
RSB-15	*halabanensis*	na	na	A65	A65	na	na	na	na
RSB-18	*malaccensis*	na	na	A3*	A61	A24	A90	A58	A58
RSB-19	*malaccensis*	na	na	A46	A46	A17	A25*	A08*	A08*
RSB-20	*malaccensis*	A07*	A47	A01*	A07*	A12*	A17	A08*	A10*
SSB-1	*malaccensis*	A46	A47	A04*	A27*	A22	A50	A01*	A01*
SSB-3	*malaccensis*	A51	A51	A05*	A06*	A21	A25*	A01*	A08*
SSB-6	*malaccensis*	A07*	A7*	A05*	A06*	A25*	A25*	A01*	A08*
SSB-10	*malaccensis*	A52	A52	A22*	A22*	A41*	A50	A23*	A50
SSB-13	*malaccensis*	A16*	A16*	na	na	A41*	A50	A23*	A23*
SSB-19	*halabanensis*	na	na	A55	A55	A79	A79	A40	A40
SSB-22	*longipetiolata*	A07*	A07*	A06*	A06*	A01*	A01*	A04*	A04*
SSB-25	*halabanensis*	A39	A39	A56	A56	A85	A85	A59	A61
SSB-29	*halabanensis*	A39	A39	A55	A58	A85	A85	A58	A58
SSB-30	*sumatrana*	A19*	A19*	A51	A51	A26	A26	A17*	A17*
SSB-33	*halabanensis-sumatrana*	A39	A19*	A51	A55	A26	A72	A17*	A58
SSB-34	*halabanensis*	A39	A39	A42*	A42*	A79	A84	A40	A40
SSL-1	*malaccensis*	A07*	A07*	A05*	A7*	A21	A21	A01*	A01*
SSL-4	*halabanensis*	A39	A39	A58	A58	A72	A79	A55	A55
SSL-10	*halabanensis*	A39	A39	A49	A63	A79	A81	A43	A60
SSL-11	*halabanensis*	A39	A39	A58	A58	A80	A80	A58	A60
SSL-14	*halabanensis*	na	na	A55	A58	A78	A81	A58	A58
SSL-15	*halabanensis*	A39	A39	A58	A58	A78	A79	A57	A60
SSL-18	*halabanensis*	na	na	A55	A63	A72	A79	A60	A60
SSL-19	*sumatrana*	A19*	A19*	A50	A50	na	na	A17*	A17*
SSL-22	*halabanensis*	A39	A39	A58	A63	A72	A79	A59	A59
SSL-23	*malaccensis*	A16*	A16*	A11*	A47	A41*	A41*	A23*	A49
SSL-27	*halabanensis*	A39	A39	A58	A64	A79	A82	A59	A60
SSL-30	*malaccensis*	A16*	A16*	A11*	A47	A40*	A50	A23*	A49
SSL-33	*halabanensis*	A39	A39	A58	A63	A79	A79	A60	A60
SSL-34	*halabanensis*	A39	A39	A58	A58	A79	A79	A60	A60
CBN-1	*malaccensis*	A46	A46	A05*	A07*	A25*	A25*	A08*	A23*
CBN-3	*malaccensis*	A06*	A46	A05*	A07*	A12*	A21	A08*	A10*
CBN-4	*malaccensis*	A48	A48	A05*	A06*	A21	A22	A04*	A10*
CBN-9	*malaccensis*	A47	A49	A06*	A46*	A22	A25*	A04*	A10*
PAN-01	*banksii*	A31*	A31*	A35*	A35*	A58*	A58*	A51	A51
PA-19	*banksii*	na	na	A35*	A35*	A58*	A58*	A51	A51
PA-79	*acuminata*	na	na	A11*	A16*	A58*	A58*	A32*	A32*
APH-192	*bantamensis*	A16*	A16*	A11*	A48	A50	A50	A23*	A49
LIPI-218	*breviformis*	A17*	A16*	A11*	A16*	A49	A49	A23*	A23*
PNK-24	*flava*	na	na	A55	A58	na	na	na	na

Accession	ssp/var/cv/type	ADH1		CAT2		GBSS1		IDH1	
PNK-26	*flava*	na	na	A71	A71	A33	A34	na	na
PAA-97	*halabanensis*	na	na	A68	A68	A26	A26	A44	A44
PAA-114	*halabanensis*	A39	A39	na	na	na	na	na	na
APH-258	*zebrina*	na	na	A11*	A11*	na	na	na	na
PAA-110	*sumatrana*	A19*	A19*	A68	A68	A26	A26	A45	A45
PH-01	*rutilifes*	A17*	A17*	A11*	A26*	A42*	A43*	A23*	A62
PAN-26	*tomentosa*	A30*	A30*	A52	A52	A58*	A59	A47	A47
PAR-100	*tomentosa*	A30*	A30*	A31*	A54	A35	A36	A52	A52
APH-385	*zebrina*	na	na	A7*	A27*	na	na	na	na
LIPI-010	*malaccensis*	A06*	A12*	A3*	A05*	A09	A14	A01*	A10*
ANS-26	"Pisang DinginLidi"	A07*	A53	A05*	A07*	A13	A58*	A01*	A08*
LIPI-232	"Pisang Rejang"	A07*	A53	A05*	A07*	A01*	A01*	A01*	A08*

Discussion

In conservation biology the identification of genetic diversity is crucial (Chung et al., 2020). Also for crop improvement knowledge about the genetic diversity is essential to ensure adequate diversity in a breeding programme (Louwaars, 2018). The alternative is monoculture cropping that in addition to its vulnerability to disease development also affects subterranean life such as in coffee (Zhao et al., 2018) and banana (Shen et al., 2017) as well as aboveground life strategies of insects (Varah et al., 2020). Banana cropping is one of the largest global monocultures in contemporary agriculture with Cavendish varieties globally dominating production (~50%) (FAO, 2016) and the export trade (≥95%) (Voora et al., 2020). Therefore, we studied the genetic diversity of wild *M. acuminata,* one of the progenitors of modern edible bananas, of Sumatra, Indonesia. These data represent an unparalleled exploration of genetic diversity in this species, which provides a high-resolution picture on the spatial diversity across the island that will benefit future breeding programs. A previous study elucidated the haplotype composition of wild and edible bananas from the ITC collection and hypothesized that western Indonesia is important for the origin of cultivars (Volkaert, 2011). Our study substantiates that hypothesis by investigating wild germplasm from Sumatra, which is underrepresented in (*ex situ*) genebanks as exemplified by just a single accession of *M. acuminata* ssp. *sumatrana* that was only recently introduced in the ITC (Ruas et al., 2017). Hence, we undertook an effort to explore the diversity of wild *M. acuminata* on Sumatra and describe its distribution, morphological and genetic diversity as a first step of a conservation and deployment strategy.

Classification of the wild Musa acuminata of Sumatra

Based on plant habit, male floral bud, bunch, fruit and seed characters (Simmonds & Weatherup, 1990; Simmonds & Shepherd, 1955), we recognized five types of wild *M. acuminata* on Sumatra. These types were generally easily recognizable even in their natural environments. Nasution (1991) described two varieties from Sumatra with a top-shaped male bud of red colour, viz. var. *malaccensis* and var. *longepetiolata*, the main difference between the two being the very long petiole in the latter variety. We observed var. *longepetiolata* (Type IV) only at a few locations in a rather limited area of Southern Sumatra. Though plants belonging to the var. *longepetiolata* can be easily distinguished by their remarkably long petioles, neither the morphological (PCA) nor the haplotype data indicate that this variety would constitute a distinct genepool. Nevertheless, concluding that var. *longepetiolata* should be revised and merged altogether into ssp. *malaccensis* is premature, since only very few accessions of var. *longepetiolata* were assessed in this study.

One group of wild bananas could be easily distinguished from all others on Sumatra by its purple male bud colour and strikingly globular seed. Nasution (1991) described two varieties with globular seed but differing in male bud shape. One, var. *halabanensis*, had a rounded male bud, while the other, var. *alasensis*, had a lanceolate male bud. The var. *alasensis* was described by Nasution (1991) from samples collected in the Alas River valley in Aceh (Nasution, 1991). We did not find any wild *M. acuminata* with globular seed and lanceolate male bud during our resampling of that area. However, we observed that at an early stage the male bud of the wild bananas in the area is lanceolate which turns into a more rounded shape later on (Supplementary Figure 3). According to the guidelines in the "Descriptors for banana (*Musa*)" (IPGRI 1996) the description of the male bud should be done at a later stage. Hence we speculate that Nasution (1991) described var. *alasensis* from an early stage of the male bud. Our genotyping data also supports the conclusion that Nasution's designation of var. *alasensis* (accession maintained in the LIPI collection) was premature and should be corrected and placed in ssp. *halabanensis*.

We observed undeveloped fruits in nearly all bunches in one or more of the early hands of 38 accessions of ssp. *halabanensis* plants in their natural environment (supplementary Figure 4). This was also reported by Itino et al. (1991), but not substantiated by photography. Their drawing seems to indicate that only the fruits of the hands at the underside of the bunch remained undeveloped. In our exploration, we found that a variable number of the early hands in bunches did not develop but this was independent of the orientation of the hand on the rachis, while later hands were fully developed. The incomplete development of fruits could result from a lack of pollination or from

(partial) sterility without parthenocarpic fruit development. A lack of pollination is not very likely since the populations of wild bananas were generally quite large with several individuals flowering simultaneously with either male or female flowers. Sterility is common in cultivated or hybrid bananas, because of the presence of chromosomal translocations and / or inversion in heterozygous bananas leading to unbalanced chromosome segregations at meiotic division (Ahmad et al., 2020), resulting in aneuploidy and gamete lethality (Dodds, 1943; Shepherd, 1999; Wilson, 1946b, 1946a). Cultivated bananas in general are sterile, but the presence of parthenocarpy drives the fruit development (Simmonds, 1953). Hence, cytogenetic observation of the meiotic stages of pollen mother cells and pollination experiments could shed light on the cause of underdeveloped fruits. Since sterility due to cytogenetic irregularities would occur in all developing flowers in a bunch, as in the cultivated bananas, the underdeveloped fruits in early hands observed here is most likely due to unknown developmental or physiological factors that need further study.

We also observed two accessions of ssp. *halabanensis,* one in Aceh (ANS-32) and the other in Bengkulu province, (SSB-29), where the fruit peel split open at maturity (Supplementary Figure 5). The cracked peel in ssp. *halabanensis* probably results from a similar mechanism as the fruit peel splitting in *M. schizocarpa*. As the observation of the fruit splitting depends on the right stage of fruit development and may also be influenced by the environment it is presently unclear whether this phenomenon would be characteristic for the subspecies as a whole or is limited to only some subpopulations as a result of genetic differentiation or environmental factors.

We identified eight plants that matched Nasution's (1991) description of var. *sumatrana*. They were easily recognisable by the vertical bunch with long fruits, dark colour of the male floral bud and flat seed with rough edges. Nasution (1991) classified this type as a variety, but as we found this banana across all seven provinces of Sumatra at high elevation and with a constant morphology, we consider it as a subspecies. This is supported by both the PCA and the haplotype analysis. However the ssp. *sumatrana* is genetically not isolated from the surrounding plants belonging to other subspecies. Aside from a few unique haplotypes, the ssp. *sumatrana* individuals contained haplotypes that belong to the large cluster of ssp. *malaccensis,* ssp. *halabanensis* and ssp. *zebrina* haplotypes.

At three locations close to each other but all at high elevations at the southern side of Mount Mugajah, Central Aceh, we found an unusual wild banana, Type V, *i.e.,* accessions ANS-15, ANS-16 and ANS-17. These plants were distinct by means of a rounded male bud with clearly imbricated bract tips, the bracts lifting rather than rolling back and the

long pedicel of the fruit. However, it is premature to describe it as a new subspecies or variety due to the limited sample size and very limited molecular data. It requires more in-depth sampling and analyses to determine whether this is an expression of environmental plasticity or a true genetic variation.

Among the 65 morphological traits that were recorded, seed morphology seemed to be a very robust character for the identification of ssp. *sumatrana* which has flattened seed with rough edges and ssp. *halabanensis*, which has smaller globular seed, differentiating them from all other wild *M. acuminata* whose seeds are irregularly angulate and slightly depressed with a smooth or minutely tuberculate surface. According to Nasution (1991) an angular shape of wrinkled seed is common among wild *M. acuminata*, which he described in 13 *M. acuminata* varieties underscoring the importance of generative organs in traditional plant taxonomy (Ornduff, 1978).

Molecular data and networks

The fact that so many unique haplotypes have been determined in this study may raise the question whether spurious sequencing errors have been interpreted as mutations. However, most of the loci have been covered by sequencing in both directions and technical errors have been meticulously checked in both sequencing reactions. Actually, there might be even more polymorphic positions than we report as ambiguous positions that would create singleton haplotypes were disregarded. Every polymorphic position that has been retained was either unambiguous on its own or confirmed by its presence in more than one genotype. Such large numbers of alleles in a species are not exceptional for natural populations of tropical plants (Wattanakulpakin et al., 2015). Even if a few polymorphic positions have been erroneously included or wrongly assigned to a haplotype by the phasing algorithm, they will not significantly impact the network structure underlying the grouping into subspecies clusters.

The haplotype networks contain some level of homoplasy, especially for GBSS and ADH. Aside from the haplotype ascertainment difficulties because of the presence of polymorphisms that were observed only once, homoplasy may be due to historical recombination between divergent alleles or parallel mutations and reverse mutations.

Haplotype analysis can support the subspecies grouping

Nasution (1991) classified Indonesian wild *M. acuminata* according to morphological characters based on his exploration and existing specimens in the Herbarium Bogoriense. Bananas are known to have highly variable phenotypes depending on environmental and seasonal conditions. Thus, characterization based on morphology alone may lead to spurious naming of forms as exemplified by the banana cultivars (Hakkinen 2013).

For instance, ssp. *sumatrana* accessions grown at low elevations in the LIPI gardens are morphologically very similar to ssp. *malaccensis* / var. *zebrina* (Supplementary Figure 5). Thus, molecular studies are necessary to supplement or correct the classification based on morphology (Rana et al., 2014; Soltis & Soltis, 1998; 2000). To support the morphological observations, we used multilocus DNA sequencing to determine the haplotypes and genotypes of wild banana accessions across Sumatra. In the parsimony networks the haplotypes clustered largely according to the morphological subspecies assignments (Figure 4) independent of their geographical origin.

In this study, we propose the merger of two varieties, var. *halabanensis* and var. *alasensis* (Nasution, 1991), into one subspecies, ssp. *halabanensis*. We justify the merging not only based on our resampling in locations where var. *alasensis* was described by Nasution (1991), the Alas River valley, but additionally on the haplotype analysis. The haplotypes detected among all wild *M. acuminata* with globular seed and rounded male bud, i. e. ssp. *halabanensis* mostly group in a single cluster for each of the four loci (Figure 4) indicating that this population is genetically distinct. Hence, across Sumatra these plants are more closely related to each other than to other subspecies that may be growing in nearby locations. However, haplotypes that typically occur in other subspecies can occasionally also be found in plants that otherwise conform to the ssp. *halabanesis* morphology.

In addition to some unique haplotypes, the ssp. *sumatrana* accessions shared haplotypes with ssp. *halabanensis*, ssp. *malaccensis* and ssp. *zebrina* (Table 6). The ssp. *sumatrana* generally grows at higher elevations whereas within the same area the other subspecies grow mostly at lower elevations. The occurrence of these subspecies in close proximity can explain the sharing of haplotypes through admixture one or several generations earlier. Though the barriers to geneflow are not absolute, the distinct morphology and ecology together with the presence of several private haplotypes justifies our decision to consider these bananas a distinct subspecies

For the distinct Type V plants the limited DNA sequence data that were obtained indicated the presence of unique haplotypes at all four loci, *i.e.*, ADH-A45, CAT-A66, GBSS-A60, IDH-A41 and IDH-A42 (Table 6) suggesting that it would be genetically distinct to other *M. acuminata* subspecies or varieties. More accession of this type will have to be analysed to understand the genetic relationship of this banana to the other subspecies. It remains to be seen whether additional unique haplotypes can be found and whether they would cluster separately from other populations, before we can describe this type as a new infraspecific group be it as subspecies or variety.

The haplotypes of the ssp. *malaccensis* and var. *longepetiolata* accessions of Sumatra are distributed in the centre of the networks among those of ssp. *malaccensis* and ssp. *zebrina* from other origins. This result accords with Volkaert (2011), who observed that ssp. *malaccensis* has a high genetic diversity and does not form a specific cluster in the network, which might indicate a polyphyletic origin for this subspecies as currently described based on morphology.

The observed admixtures should not be surprising given the sometimes very patchy but intermingled occurrence of the different morphological types of wild bananas. It might even be visible in the morphology as in accession SSB-33, which had a fruit that was intermediate between ssp. *sumatrana* and ssp. *halabanensis* (Supplementary Figure 6) and has haplotypes of ssp. *halabanensis* admixed with others that cluster with ssp. *sumatrana* or ssp. *malaccensis* across the four loci studied (Table 6). The occurrence of such intermediate morphologies was already reported by Nasution (1991).

Taken together, the haplotype composition of accessions can support conservation strategies by selecting appropriate accessions for model morphology characterization. If in a highly diverse species like *M. acuminata* distinct molecular clusters can be identified, they can guide the selection of type specimens that contain haplotypes exclusive to a single cluster.

Population genetics of wild M. acuminata of Sumatra
Highly significant deviations from Hardy-Weinberg equilibrium were observed for all four loci whether the analysis was done for all accessions, only accessions from Sumatra or for the individual subspecies with on Sumatra (Table 5). The high number of identified haplotypes and the relatively modest number of analysed accessions, resulted in a very large number of diploid genotypes occurring at very low frequencies and many potential genotypes completely missing, making the significance testing for Hardy-Weinberg equilibrium less reliable. The clear deficiencies of heterozygotes may indicate that there are genetically isolated subpopulations between but also within each subspecies. It is assumed that the different islands of the Indonesian archipelago have been isolated and joined during glacial cycles over the past 400,000 years (Husson et al., 2020) and thus it might be expected that many islands have distinct subpopulations. Our finding that there are highly significant deviations of a Hardy-Weinberg equilibrium even within each subspecies on a single island indicates the presence of more fine-scaled genetically distinct subpopulations.

Since we found a significant deviation from Hardy-Weinberg equilibrium for each locus, we investigated whether Principal Component Analysis of Neighbourhood Matrices,

also known as Moran's Eigenvector maps, which analyses diversity data in combination with (geographical) distance data, could give some insights on the scale of population structure. However, the sampling of accessions was not designed with this purpose in mind and thus the dataset is suboptimal for such an analysis. Because of the large number of haplotypes and genotypes and only four loci investigated, the molecular data could not be used. Analysing the individual SNP positions is not a solution either as the SNPs are tightly linked and thus not independent data. Therefore, only the phenotypic data could be analysed. We chose to use Euclidian distance data as they can be easily calculated from the GPS data but that may not be the most appropriate distance measure. The mountainous topography of Western Sumatra and the occurrence of the ssp *malaccensis* and ssp. *halabanensis* being restricted to lower elevations would indicate that the distances along river valleys might be more appropriate. Despite these caveats, some preliminary analyses such as the PCNM graph indicate the occurrence of subpopulations that would accord with the complex topography of Sumatra. Hence, each watershed, each valley, may have its unique subpopulation of bananas for each subspecies, locally interbreeding but somewhat isolated from the populations of the other parts of the island. Consequently, further studies of the wild banana population structure of Sumatra should adopt a sampling strategy across the various topographies of the island to include even more accessions.

Origins of edible bananas

Although heterozygosity of particular loci can be tested by various co-dominant markers such as SSR (Creste et al., 2004; Retnoningsih et al., 2010) or RFLPs (Carreel et al., 2002), they cannot readily identify which wild banana population or subspecies contributed to a particular cultivar. SSRs have been used to identify the diploid progenitors of the triploid Cavendish and Gros Michel bananas (Raboin et al., 2005). Still it has proven to be very difficult to assign genetic origins for most other banana cultivars (Hippolyte et al., 2012). The advantage of DNA sequence analysis is that not only exactly matching patterns are informative, but because the evolutionary relationships among sequences can be determined, even imperfect matches can be partially informative. In a previous study, 181 haplotypes for ADH1, CAT2, GBSS1 and IDH1 in wild and cultivated bananas have been reported (Volkaert 2011) which suggested a high genetic diversity in wild *M. acuminata*. Our analysis supports this observation as 103 haplotypes were characterised within the Sumatran *M. acuminata* (Table 3).

The large number of haplotypes identified in wild *M. acuminata* of Sumatra underscores that the island is indeed an important part of the centre of diversity of banana. In a previous study, Volkaert (2011) described a unique haplotype detected in the cultivar

"Pisang Jari Buaya" which was not found in any wild accession in the international germplasm collection (ITC). The present analysis provides additional data that indeed supports a contribution from a wild banana from Sumatra to this group of cultivars. This underscores that the deficiency of genetic information of wild *M. acuminata* in existing collections limits our taxonomical insights. Hence, a deep study of wild *M. acuminata* as well as *M. balbisiana* in other parts of Indonesia is necessary to unveil a higher taxonomical resolution.

In this study we found several haplotypes shared among wild and cultivated bananas (see Table 6). The similarity of the haplotypes can be used to discern the relationship among bananas. For instance, haplotype A07 of ADH1 (Figure 4), can be an indication that ssp. *malaccensis* is one of the progenitors of "Pisang Rejang" and "Pisang Dingin Lidi". Some haplotypes such as A11 of CAT2 that are located in the centre of the network (Figure 4) were also identified in five different subspecies/varieties from diverse locations, *i.e.*, ssp. *zebrina* from Java, ssp. *malaccensis* from Sumatra, ssp. *burmanica* from mainland Asia and var. *acuminata* from Papua. This is confirmed by Volkaert (in prep), who observed that haplotype A11 of CAT2 occurs at different islands, albeit in dissimilar frequencies. The most likely explanation is that islands in South East Asia were previously connected due to low sea level (Husson et al., 2020; Voris, 2000), thereby preventing isolation and facilitating natural genetic admixtures.

The analysis of DNA sequences is necessary to describe the haplotype diversity and distributions in wild or cultivated bananas (Swangpol et al., 2007; Volkaert, 2011; Volkaert in prep). In a previous study, Volkaert (2011; in prep) showed that 10, 7, 8, 13 and 7 haplotypes of ADH1, ARF17, CAT2, GBSS1 and IDH1, respectively, detected in cultivated bananas could not be traced to the studied wild bananas obtained from ITC. One of the cultivars with such unidentified haplotypes was "Pisang Jari Buaya". In the present study, the "Pisang Jari Buaya" haplotype for the CAT2 (A42) and IDH1 (A40) loci was identified in the accessions SSB-34 and J-21, both ssp. *halabanensis*. Albeit that for GBSS no haplotypes identical to "Pisang Jari Buaya" (A83) were identified in the wild *M. acuminata* accessions, other very similar (A79, A84 and A87) haplotypes were observed (Figure 3), all belonging to the *ssp. halabanensis* cluster. In addition, the accession J-21 has a GBSS haplotype belonging to the ssp. *zebrina* cluster (A26) which is just one single mutation different from the GBSS haplotype found in "Pisang Jari Buaya" (A27). This is a strong indication that ssp. *halabanensis* has contributed as a parent of "Pisang Jari Buaya" and indicates that the evolution of this cultivar may have happened in ssp. *halabanensis* populations in Southern Sumatra such as the Jambi, Bengkulu and Lampung provinces. Unfortunately, the haplotypes of ADH are distinct between "Pisang Jari Buaya" and any ssp. *halabanensis* accessions (Figure 4)

suggesting that ssp. *halabanensis* has contributed ADH alleles to "Pisang Jari Buaya" or that the sampling depth was too shallow.

Taken together, comparing the haplotype diversity on Sumatra with the diversity in the most widespread cultivars, then Sumatran wild *M. acuminata* did not contribute much to the edible bananas (Volkaert, in prep). The ssp. *malaccensis* haplotypes found on Sumatra are closely related but not identical to the haplotypes in cultivars. However, the Sumatran wild bananas are probably at the origin of the cultivars belonging to the "Pisang Jari Buaya" group. Moreover, we can conclude with very high confidence that the ssp. *halabanensis* contributed to these cultivars and that most likely an ancient population related to contemporary populations in southern Sumatra are the progenitor population.

Previous studies (Volkaert 2011; Li et al. 2013, Martin et al. 2020) have indicated that most cultivated bananas have originated from intraspecific hybridisations mostly involving ssp. *malaccensis,* ssp. *zebrina* and ssp. *banksii*. The current analysis of PJB confirms this. But in our exploration of the diversity in the Sumatran wild bananas we discovered many accessions that have such a pattern of admixture between various subspecies, *i.e.* ssp *malaccensis* with ssp. *halabanensis* or ssp. *sumatrana* with ssp *halabanensis*. Even more complex patterns where haplotypes from three subspecies can be found in a single accession were recorded such as in accession ANS-15 (type V) that has haplotypes that are identical to ssp. sumatrana for locus ADH (A19), ssp. zebrina for locus CAT2 (A22) in addition to haplotypes belonging to its unique haplotypes (Table 6). This should not surprise us given the close intermixing of morphologically distinct subspecies at several locations. But this also implies that the mountainous topography of Sumatra island lends itself very well to natural admixture. Thus the admixed genotypes observed in several of the banana cultivars could well have happened naturally at various places where the individual types occur in close proximity. There would not be any need for humans facilitating the admixture by carrying semi-domesticated cultivars from one location to another during their migrations. The reason that human transportation has been invoked to explain the complex genotypes (Perrier et al. 2011) is more likely the inadequate exploration of the genetic diversity in wild populations and the reliance on the limited number of genotypes represented in the international germplasm collections.

None of the other main islands in Indonesia, *i.e.*, Java, Kalimantan, Sulawesi and Papua, have been explored in much detail leaving the possibility open for other major taxonomical discoveries that would further our understanding of the banana diversity and potential for deployment. Therefore, exploring wild *M. acuminata* and other

ancestors of contemporary edible bananas on other Indonesian islands as well as the greater Southeast Asian region is important from both a conservation and breeding perspective.

Conclusions

The present study contributes to the knowledge of the diversity of wild *M. acuminata* of Sumatra including morphological variations at the infraspecific level and based on multilocus genetic data reports on a fine tuned diversity with consequences for taxonomy. This study revealed the diversity of wild *M. acuminata* of Sumatra consists of five types *i.e.*, ssp. *malaccensis* (Type I), ssp. *halabanensis* (Type II), ssp. *sumatrana* (Type III), var. *longepetiolata* (Type IV) and one type that has never been described and will need further study to classify it into the *Musa acuminata* group (Type V). Moreover, The data show that many new haplotypes are not present in the international collections (ITC) and hence are inaccessible for diversifying breeding programs. Yet, the haplotypes of several wild accessions are similar to the haplotypes identified in contemporary banana cultivars, suggesting that indeed, some edible bananas may have originated from Sumatra or nearby regions. This finding indicated that the genetic knowledge of Indonesian banana is incomplete. Further exploration of wild ancestors of banana cultivars across diverse topographies of the Indonesian archipelago is important from a conservation standpoint, particularly with an eye on competing claims for agricultural or urban land use.

Author contribution statement

FA and HAV designed this study and carried out the experiments and data analysis. YSP and SS helped with the exploration and selection of banana materials for this study. LDS helped with the taxonomy study and herbarium observation. FA, HAV, GHJK and HdJ wrote the manuscript. HAV, GHJK and HdJ supervised the project.

2

References

Ahmad F, Bourke P, Schouten H, Volkaert H, de Jong H. (2020) Cytogenetics of Structural Rearrangement in *Musa* Hybrids and Cultivars. In: Kema G, Drenth A, editors. Achieving sustainable cultivation of bananas Volume 2: Germplasm and genetic improvement. Cambridge, UK: Burleigh Dodds Science Publishing

Borcard, D., & Legendre, P. (2002). All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecological Modelling, 153(153), 51–68.

Carreel, F., Gonzalez de Leon, D., Lagoda, P., Lanaud, C., Jenny, C., Horry, J. P., & Tezenas du Montcel, H. (2002). Ascertaining maternal and paternal lineage within *Musa* by chloroplast and mitochondrial DNA RFLP analyses. Genome, 45(4), 679–692. https://doi.org/10.1139/g02-033

Chung, M. Y., Son, S., Herrando-moraira, S., Tang, C. Q., Maki, M., Kim, Y., Lopez –Pujol J., Hamrick J. L., Chung, M. G. (2020). Incorporating differences between genetic diversity of trees and herbaceous plants in conservation strategies. Conservation Biology, 34:1142-1151

Creste, S., Tulmann Neto, A., Vencovsky, R., de Oliveira Silva, A., & Figueira, A. (2004). Genetic diversity of Musa diploid and triploid accessions from the Brazilian banana breeding program estimated by microsatellite markers. *Genetic Resources and Crop Evolution*, 51, 723–733.

Crichton, R. R., Vezina, A., & Van den Bergh, I. (2016). An online checklist of banana cultivars. Acta Horticulturae 2(1114):13–18.

Dmitriev, D. A., & Rakitov, R. A. (2008). Decoding of superimposed traces produced by direct sequencing of heterozygous indels. PLoS Computational Biology, 4(7). https://doi.org/10.1371/journal.pcbi.1000113

Dodds, K. (1943). Genetical and cytological studies of *Musa*. Certain edible diploids. J Genet 45(2), 113–138

FAO. (2016). Banana facts and Figures. FAO, United Nations. Retrieved from http://www.fao.org/economic/est/est-commodities/bananas/bananafacts/en/

FAO. (2020). Fruit and vegetables – your dietary essentials. The International Year of Fruits and Vegetables, 2021, background paper. Rome. https://doi.org/10.4060/cb2395en

Häkkinen, M. (2013). Reappraisal of sectional taxonomy in *Musa* (Musaceae). Taxon 62(4), 809–813. https://doi.org/10.12705/624.3

Häkkinen, M., & Väre, H. (2008). Typification and check-list of *Musa* L. names (Musaceae) with nomenclatural notes. Adansonia 30(1), 63–112.

Hamilton, C. W., & Reichard, S. H.

(1992). International Association for Plant Taxonomy (IAPT). Taxon 41(3), 485–498.

Hippolyte, I. Jenny, C., Gardes, L., Bakry, F., Rivallan, R., Pomies, V., Cubry, P., Tomekpe, K., Risterucci, A. M., Roux, N. Rouard, M., Arnaud, E., Kolesnikova-Allen, M. & Perrier, X. (2012). Foundation characteristics of edible *Musa* triploids revealed from allelic distribution of SSR markers. Annals of Botany 109, 937-951

Hotta, M. (1989). Identification list of *Ensete* and *Musa* (Musaceae) in SE Asia and West Malesia. South Pacific Ocean Research Report: Nankai Lab Occasional papers (Vol. 16). Kagoshima.

Husson, L., Boucher, F., Sar, A.-C., Sepulchre, P., & Cahyarini, S. (2020). Evidence of Sundaland's subsidence requires revisiting its biogeography. Journal of Biogeography 47, 843–853. https://doi.org/10.1111/jbi.13762

IPGRI (1996). Descriptors for Banana (*Musa* spp.). IPGRI, Rome, Italy.

Itino, T., Kato, M., & Hotta, M. (1991). Pollination Ecology of the Two Wild Bananas, *Musa acuminata*, subsp. *halabanensis* and *M. salaccensis*. Biotropica, *23*(2) 151–158.

John Kress, W., & Specht, C. D. (2006). The Evolutionary and Biogeographic Origin and Diversification of the Tropical Monocot Order Zingiberales. Aliso: A Journal of Systematic and Evolutionary Botany 22(1), 49.

Li, L. F., Wang, H. Y., Zhang, C., Wang, X. F., Shi, F. X., Chen, W. N., & Ge, X. J. (2013). Origins and domestication of cultivated banana inferred from chloroplast and nuclear genes. PLoS ONE. https://doi.org/10.1371/journal.pone.0080502

Librado, L., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25(11), 1450–1452.

Louwaars, N. P. (2018). Plant breeding and diversity : A troubled relationship? Euphytica, *214*(7) 1–9. https://doi.org/10.1007/s10681-018-2192-5

Metsalu, T., & Vilo, J. (2015). ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. Web Server Issue Published Online, 43. https://doi.org/10.1093/nar/gkv468

Oksanen, A. J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin P. R., O'Hara R. B., Simpson G. L., Solymos P, Stevens M. H. H., Szoecs E., Wagner, H. (2020). Package 'vegan': Community ecology package version 2.5-7. https://cran.r-project.org/web/packages/vegan/vegan.pdf

Ornduff, R. (1978). Reproductive Characters and Taxonomy. Systematic Botany, *3*(4), 420–427.

Perrier, X., Bakry, F., Carreel, F., Jenny, C., Horry, J. P., Lebot, V., & Hippolyte, I. (2009). Combining biological approaches to shed light

on the evolution of edible bananas. Ethnobotany Research and Applications 7:199–216.

Perrier, X., De Langhe, E., Donohue, M., Lentfer, C., Vrydaghs, L., Bakry, F., Careel F., Hippolyte I., Horry J-P., Jenny C., Lebot V., Risteruci A.-M., Tomekpe., Doutrelepont H., Ball T., Manwaring J., de Mret P., Denham, T. (2011). Multidisciplinary perspectives on banana (*Musa* spp.) domestication. Proceedings of the National Academy of Sciences, 108(28):11311–11318. https://doi.org/10.1073/pnas.1102001108

R Core Team (2020). R: A Language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Raboin, L.-M., Carreel, F., Noyer, J.-L. (2005). Diploid ancestors of triploid export banana cultivars: molecular identification of 2n restitution gamete donors and n gamete donors. Molecular Breeding 16:333-341

Rana, T. S., Narzary, D., Mahar, K. S., & Meena, B. (2014). Methods and Approaches in Plant Molecular Systematics. In T. Rana, K. Nair, & D. Upreti (Eds.), Plant Taxonomy and Biosystematics: Classical & Modern Methods. New Delhi: New India Publishing Agency pp. 383–416

Retnoningsih, A., Megia, R., & Hartana, A. (2010). Molecular Verification and Diversity Analysis of Indonesian

BB, AAB and ABB Banana Cultivars. Tree and Forestry Science and Biotechnology 4(1): 69–76.

Ruas, M., Guignon, V., Sempere, G., Sardos, J., Hueber, Y., Duvergey, H., Adeka A., Chase., Jenny C., Hazekanp., Irish., Jaleli., Ayala-Silva T., Chao C. P., Daniells J., Dowiya B., Effa B. E., Gueco L., Herradura L., Ibobondji L., Kempenaers E., KilangiJ., Musangi S., Xuan P. N., Paofa J., Pavis C., Thiemele D., Tossou C., Sandoval J., Sutanto A., Paka G. V., Yi G., Van de Houwe I., Roux., Rouard, M. (2017). MGIS: managing banana (*Musa* spp.) genetic resources information and high-throughput genotyping data. Database : The Journal of Biological Databases and Curation 2017:1–12. https://doi.org/10.1093/database/bax046

Shen, Z., Penton, C. R., Lv, N., Xue, C., Yuan, X., Li, R., & Shen, Q. (2017). Banana Fusarium Wilt Disease Incidence Is Influenced by Shifts of Soil Microbial Communities Under Different Monoculture Spans. Microbial Ecology 75:739–750. https://doi.org/10.1007/s00248-017-1052-5

Shepherd, K. (1999). *Cytogenetics of the genus Musa*. Montpellier, France: International Network for the Improvement of Banana and Plantain

Simmonds, N W. (1953). Segregation in some diploid bananas. Journal of Genetics 51:458–469

Simmonds, N W. (1956). Botanical Results of the Banana Collecting Expedition, 1954-5. *Kew Bulletin* 11(3):463–489

Simmonds, N W, & C Weatherup, S. T. (1990). Numerical Taxonomy of the Wild Bananas (*Musa*). The New Phytologist 115(3):567-571

Simmonds, Norman Willison, & Shepherd, K. (1955). The taxonomy and origins of the cultivated bananas. Botanical Journal of the Linnean Society, 55(359):302–312.

Soltis, D. E., & Soltis, P. S. (1998). Choosing an Approach and an Appropriate Gene for Phylogenetic Analysis. In Molecular Systematic of Plants II pp. 1–42. https://doi.org/10.1007/978-1-4615-5419-6

Soltis, E. D., & Soltis, P. S. (2000). Contributions of plant molecular systematics to studies of molecular evolution. Plant Molecular Biology, *42*(1), 45–75. https://doi.org/10.1023/A:1006371803911

Stephens, M., Nicholas, J. S., & Donelly, P. (2001). A New Statistical Method for Haplotype Reconstruction from Population Data. Amer. J. Human Genet 68, 978–989

Stover, R., & Simmonds, N. (1987). *Bananas: Tropical Agriculture Series* (3rd ed.). New York: Inc.

Swangpol, S., Volkaert, H., Sotto, R. C., & Seelanan, T. (2007). Utility of Selected Non-coding Chloroplast DNA Sequences for Lineage Assessment of Musa Interspecific Hybrids. BMB Reports, 40(4):577–587. https://doi.org/10.5483/BMBRep.2007.40.4.577

Syamkumar, S., Lowarence, B., & Sasikumar, B. (2003). Isolation and Amplification of DNA From Rhizomes of Turmeric and Ginger. Plant Molecular Biology Reporter 21: 171a–171e

Valmayor, R. V. (1998). Classification et caractérisation de *Musa exotica*, *M. alinsanaya* et *M. acuminata* ssp. *errans*. Infomusa 10, 35–39.

van Steenis, C. (1957). Specific and infraspecific delimitation. Flora Malesiana Serie I, 5(3):CLXVII–CCXXXIV.

Varah, A., Jones, H., Smith, J., & Potts, S. G. (2020). Agriculture , Ecosystems and Environment Temperate agroforestry systems provide greater pollination service than monoculture. Agriculture, Ecosystems and Environment, 301: 107031. https://doi.org/10.1016/j.agee.2020.107031

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

Voora, V., Larrea, C., & Bermudez, S. (2020). *Global Market Report : Bananas*. Retrieved from https://www.iisd.org/system/files/publications/ssi-global-market-report-banana.pdf

Voris, H. K. (2000). Maps of Pleistocene Sea Levels in Southeast Asia:

2

Shorelines, River Systems and Time Durations. Journal of Biogeography 27(5):1153–1167.

Wattanakulpakin, T., Iamtham, S., Grubbs, K. C., & Volkaert, H. A. (2015). Population genetic analysis of *Xylia xylocarpa* (Fabaceae — Mimosoideae) in Thailand. Tree Genetics and Genomes 11:7. https://doi.org/10.1007/s11295-014-0825-y

Wilson, G. B. (1946a). Cytological studies in the Musae. III. Meiosis in some seedling clones. Genetics 31:483–493.

Wilson, G. B. (1946b). Cytological studies in the Musae. I. Meiosis in some triploid clones. Genetics 31:241–258.

Zhao, Q., Xiong, W., Xing, Y., Sun, Y., Lin, X., & Dong, Y. (2018). Long-Term Coffee Monoculture Alters Soil Chemical Properties and Microbial Communities. *Scientific Reports* 8:6116. https://doi.org/10.1038/s41598-018-24537-2
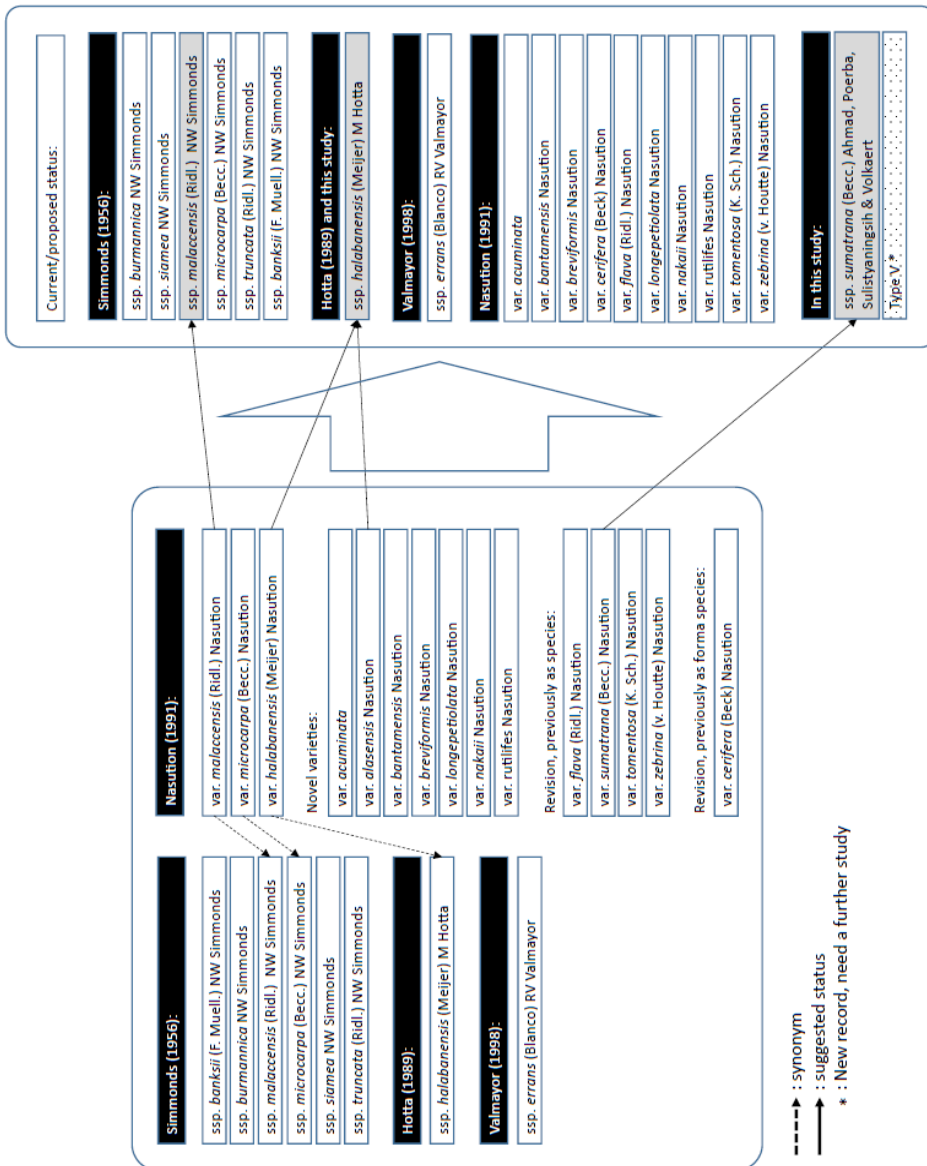
Supplementary Figure 1 Infraspecific classification of *Musa acuminata* over time. Nasution (1991) identified 15 varieties from Indonesia, three varieties indicated as a synonym of two subspecies proposed by Simmonds (1956) and a subspecies proposed by Hotta (1989) (dashed arrow). In the current study, we proposed two varieties ranked as subspecies and two other species to be merged as subspecies (arrow).

A

Supplementary Figure 2 continues

B

C

D

Supplementary Figure 2 continues

E

Supplementary Figure 2 Photographs of wild *Musa acuminata* accessions on Sumatra. Based on morphological characters of habitus, leaf, bunch, flower, fruit and seed, five types of banana were described, *i.e.*; A) ssp. *malaccensis* (Type II), B) ssp. *halabanensis* (Type II), C) ssp. *sumatrana* (Type III), D) var. *longepetiolata* (type IV), E) Type V.

Supplementary Figure 3 Characteristics of male bud development. At early stage or just after the last female flower emerged, the male bud is rather lanceolate (A, B, D, E). In contrast, at later stage it becomes shorter and rounded (C, F). Pictures A, B, C are the male buds of ssp. *halabanensis*; Pictures D, E, F are the male buds of ssp. *malaccensis*.

Supplementary Figure 4 In some cases, fruits of ssp. *halabanensis* were undeveloped with no or few seeds (arrow).

Supplementary Figure 5 In two cases we found bunches of ssp. *halabanensis* with cracked peel, similar to this phenomenon in *M. schizocarpa*.

Supplementary Figure 6 In one case, we found an accession with intermediate fruit characteristics of ssp. *sumatrana* and ssp. *halabanesis* (SSB-33), indicated by a horizontal bunch (ssp. *halabanensis*) and long fruit (ssp. *sumatrana)* with angular seed.

Supplementary Table 1 Banana accessions of the Research Center for Biology (LIPI) collection included in this study.

Accession	ssp/var	Remark	Origin	GPS latitude	longitude
Cbn-1	*malaccensis*	wild	West Java	-6.49195	106.847
Cbn-3	*malaccensis*	wild	West Java	-6.48967	106.8449
Cbn-4	*malaccensis*	wild	West Java	-6.49175	106.8458
Cbn-9	*malaccensis*	wild	West Java	-6.495	106.8459
PAN-01	*banksii*	wild	Nort Sulawesi	1.39436	124.8267
PA-19	*banksii*	wild	Papua -4.52861	137.1178	
PA-79	*acuminata*	wild	West Papua	-1.28194	134.1903
APH-192	*bantamensis*	wild	West Java	-6.80667	107.3278
LIPI-218	*breviformis*	wild	West Java	na	na
PNK-24	*flava*	wild	South Kalimantan	-1.8605	115.6317
PNK-26	*flava*	wild	South Kalimantan	-2.76111	114.8308
PAA-97	*halabanensis*	wild	West Sumatra	-0.32917	100.8711
PAA-114	*halabanensis*	wild	West Sumatra	-0.98667	100.5775
APH-258	*malaccensis*	wild	West Java	-7.3875	107.2278
PAA-110	*sumatrana*	wild	West Sumatra	na	na
PH-01	*rutilifes*	wild	East Java	-7.80333	112.7636
PAN-26	*tomentosa*	wild	Nort Sulawesi	1.27528	124.9375
PAR-100	*tomentosa*	wild	South Sulawesi	-4.49333	120.0011
APH-385	*zebrina*	wild	West Java	-6.83333	107.2628
LIPI-010	*malaccensis*	wild	Sumatra na	na	
ANS-26	*Pisang Dingin Lidi*	AA	cultivar Aceh	4.82958	96.74552
LIPI-232	*Pisang Rejang*	AA	cultivar Bengkulu	na	na

Supplementary Table 2 Accessions of wild *Musa acuminata* collected during the Sumatra exploration.

Accession	ssp/var	Province	GPS	
			latitude	longitude
ANS-1	malaccensis	North Sumatra	3.41622	98.60067
ANS-2	malaccensis	North Sumatra	3.40378	98.60232
ANS-3	malaccensis	North Sumatra	3.33193	98.57978
ANS-4	sumatrana	North Sumatra	3.20865	98.5387
ANS-5	sumatrana	North Sumatra	3.20865	98.5387
ANS-6	ND	North Sumatra	3.08075	98.17793
ANS-7	ND	North Sumatra	3.06643	98.12935
ANS-8	halabanensis	Aceh	3.06833	98.12953
ANS-9	halabanensis	Aceh	3.54927	97.77737
ANS-10	halabanensis	Aceh	3.61793	97.72543
ANS-11	halabanensis	Aceh	3.64823	97.68988
ANS-12	halabanensis	Aceh	3.71778	97.62923
ANS-13	ND	Aceh	3.76823	97.58808
ANS-14	halabanensis	Aceh	3.80127	97.57412
ANS-15	New	type Aceh	4.19052	97.20778
ANS-16	New	type Aceh	4.19507	97.21275
ANS-17	New	type Aceh	4.23878	97.18855
ANS-18	sumatrana	Aceh	4.24877	97.18332
ANS-19	halabanensis	Aceh	4.2675	97.16787
ANS-20	halabanensis	Aceh	4.4664	97.02988
ANS-21	halabanensis	Aceh	4.465	96.97932
ANS-22	halabanensis	Aceh	4.46347	96.93043
ANS-23	sumatrana	Aceh	4.6955	96.8379
ANS-24	halabanensis	Aceh	4.79655	96.7232
ANS-25	halabanensis	Aceh	4.88555	96.72828
ANS-27	halabanensis	Aceh	4.05222	96.72828
ANS-28	halabanensis	Aceh	5.0314	96.6923
ANS-29	halabanensis	Aceh	5.04588	96.69822
ANS-30	halabanensis	Aceh	4.64145	97.68957
ANS-31	halabanensis	Aceh	4.57008	97.61585
ANS-32	halabanensis	Aceh	4.5425	97.59477
ANS-33	halabanensis	Aceh	4.33798	97.52783
ANS-34	halabanensis	Aceh	4.33035	97.5289
ANS-35	halabanensis	Aceh	4.16237	97.9605
ANS-36	halabanensis	Aceh	4.09963	99.10205
ANS-37	malaccensis	North Sumatra	3.2601	98.55543
ANS-38	malaccensis	North Sumatra	3.26122	98.55607
ANS-39	malaccensis	North Sumatra	3.2679	98.57822
ANS-40	malaccensis	North Sumatra	3.2743	98.59202
ANS-41	malaccensis	North Sumatra	3.27337	98.59612
J-1	malaccensis	Jambi	-2.06725	102.9399

Accession	ssp/var	Province	GPS	
			latitude	longitude
J-2	malaccensis	Jambi	-2.34387	102.4491
J-3	malaccensis	Jambi	-2.33238	102.4407
J-4	malaccensis	Jambi	-2.33703	102.4358
J-5	malaccensis	Jambi	-2.33853	102.4344
J-6	malaccensis	Jambi	-2.33753	102.4102
J-7	malaccensis	Jambi	-2.34403	102.3692
J-8	malaccensis	Jambi	-2.34772	102.3599
J-9	malaccensis	Jambi	-2.35443	102.3518
J-10	malaccensis	Jambi	-2.35897	102.3377
J-11	malaccensis	Jambi	-2.3585	102.3414
J-12	malaccensis	Jambi	-2.07377	102.2755
J-13	malaccensis	Jambi	-2.13778	102.1234
J-14	malaccensis	Jambi	-2.14263	102.0462
J-15	malaccensis	Jambi	-2.11935	101.9411
J-16	ND	Jambi	-2.04483	101.3463
J-17	ND	Jambi	-2.0434	101.3114
J-18	ND	Jambi	-2.11857	101.2321
J-19	halabanensis	Jambi	-2.43182	101.1558
J-20	halabanensis	Jambi	-2.618	101.3107
J-21	halabanensis	Jambi	-2.16155	101.5743
J-22	halabanensis	Jambi	-2.17797	101.865
J-23	halabanensis	Jambi	-2.14572	101.9032
J-24	malaccensis	Jambi	-2.12463	102.0247
J-25	malaccensis	Jambi	-2.14362	102.1045
J-26	malaccensis	Jambi	-2.13247	102.1772
J-27	malaccensis	Jambi	-2.13247	102.1772
J-28	malaccensis	Jambi	-1.46962	102.5715
J-29	malaccensis	Jambi	-1.46673	102.5865
J-30	malaccensis	Jambi	-1.44782	102.6944
J-31	malaccensis	Jambi	-1.38127	102.7424
J-32	ND	Jambi	-1.28352	102.9012
J-33	malaccensis	Jambi	-1.26253	103.0452
J-34	malaccensis	Jambi	-1.26122	103.1448
SSB-1	malaccensis	South Sumatra	-2.93503	104.6979
SSB-2	malaccensis	South Sumatra	-2.92523	104.5701
SSB-3	malaccensis	South Sumatra	-2.92435	104.4375
SSB-4	malaccensis	South Sumatra	-2.85658	104.3217
SSB-5	malaccensis	South Sumatra	-2.85088	104.2554
SSB-6	malaccensis	South Sumatra	-2.8432	104.1779
SSB-7	malaccensis	South Sumatra	-2.75197	104.0707
SSB-8	malaccensis	South Sumatra	-2.7881	104.1288

Supplementary Table 2 continues

Accession	ssp/var	Province	GPS	
			latitude	longitude
SSB-9	malaccensis	South Sumatra	-3.078	104.8809
SSB-10	malaccensis	South Sumatra	-3.23375	104.859
SSB-12	malaccensis	South Sumatra	-3.30633	104.8341
SSB-13	malaccensis	South Sumatra	-3.46883	104.8216
SSB-14	malaccensis	South Sumatra	-3.51163	104.77
SSB-16	malaccensis	South Sumatra	-3.78368	103.6369
SSB-17	malaccensis	South Sumatra	-3.91395	103.454
SSB-18	malaccensis	South Sumatra	-3.93955	103.4348
SSB-19	halabanensis	South Sumatra	-3.76502	102.9342
SSB-20	halabanensis	South Sumatra	-3.67098	103.0164
SSB-21	halabanensis	South Sumatra	-3.57488	103.0761
SSB-22	malaccensis	Bengkulu	-3.2824	103.0562
SSB-23	malaccensis	Bengkulu	-3.2425	103.0278
SSB-24	malaccensis	Bengkulu	-3.25858	102.9896
SSB-25	halabanensis	Bengkulu	-3.90425	103.1476
SSB-26	malaccensis	Bengkulu	-3.36867	102.8047
SSB-27	halabanensis	Bengkulu	-3.3871	102.7847
SSB-28	halabanensis	Bengkulu	-3.43235	102.7509
SSB-29	halabanensis	Bengkulu	-3.43235	102.7509
SSB-30	sumatrana	Bengkulu	-3.48088	102.5878
SSB-31	halabanensis	Bengkulu	-3.55593	102.5206
SSB-32	sumatrana	Bengkulu	-3.65885	102.5506
SSB-33	Intermediate between halabanensis-sumatrana	Bengkulu	-3.66682	102.5426
SSB-34	halabanensis	Bengkulu	-3.51468	102.1878
RSB-1	malaccensis	Riau	0.22672	101.0106
RSB-2	malaccensis	Riau	0.26953	100.9413
RSB-3	malaccensis	Riau	0.30028	100.9165
RSB-4	malaccensis	Riau	0.73653	100.5027
RSB-5	halabanensis	Riau	0.7371	100.5027
RSB-6	halabanensis	Riau	0.81343	100.4202
RSB-7	halabanensis	Riau	1.42395	99.29143
RSB-8	halabanensis	Riau	1.17612	99.39957
RSB-9	halabanensis	Riau	0.68777	99.65783
RSB-10	halabanensis	Riau	0.61123	99.89555
RSB-11	halabanensis	North Sumatra	0.04403	100.2157
RSB-12	halabanensis	North Sumatra	-0.08805	100.2082
RSB-13	sumatrana	North Sumatra	-0.33828	100.7432
RSB-14	halabanensis	North Sumatra	-0.56482	100.8199
RSB-15	halabanensis	North Sumatra	-0.73478	100.9866
RSB-16	halabanensis	West Sumatra	-0.7661	101.1628
RSB-17	malaccensis	West Sumatra	-0.76985	101.1692

Accession	ssp/var	Province	GPS	
			latitude	longitude
RSB-18	malaccensis	West Sumatra	-0.76963	101.1682
RSB-19	malaccensis	West Sumatra	-0.88502	101.3503
RSB-20	malaccensis	West Sumatra	-0.88735	101.3609
RSB-21	malaccensis	West Sumatra	-0.63645	101.4744
RSB-22	malaccensis	Riau	-0.74117	102.5228
RSB-23	malaccensis	Riau	-0.74537	102.5299
RSB-24	malaccensis	Riau	-0.81205	102.5771
SSL-1	malaccensis	South Sumatra	-2.9705	104.7128
SSL-2	malaccensis	South Sumatra	-3.99207	104.1221
SSL-3	malaccensis	South Sumatra	-3.99658	104.1284
SSL-4	halabanensis	Lampung	-4.5058	104.1764
SSL-5	halabanensis	Lampung	-4.5481	104.13
SSL-6	halabanensis	Lampung	-4.64188	104.0482
SSL-7	halabanensis	Lampung	-4.79658	103.9317
SSL-8	halabanensis	Lampung	-4.81848	103.9495
SSL-9	halabanensis	Lampung	-4.9464	104.0373
SSL-10	halabanensis	Lampung	-5.01863	104.0809
SSL-11	halabanensis	Lampung	-5.14373	103.9595
SSL-12	halabanensis	Lampung	-5.10292	104.0014
SSL-13	halabanensis	Lampung	-5.10113	104.0006
SSL-14	halabanensis	Lampung	-5.0798	104.0255
SSL-15	halabanensis	Lampung	-5.27928	104.0119
SSL-16	halabanensis	Lampung	-5.48922	104.2199
SSL-17	halabanensis	Lampung	-5.54458	104.2964
SSL-18	halabanensis	Lampung	-5.62443	104.3768
SSL-19	sumatrana	Lampung	-5.52518	104.426
SSL-20	halabanensis	Lampung	-5.51043	104.4355
SSL-21	halabanensis	Lampung	-5.48498	104.4676
SSL-22	halabanensis	Lampung	-5.5026	104.6344
SSL-23	malaccensis	Lampung	-5.5025	104.6344
SSL-24	halabanensis	Lampung	-5.46512	104.6965
SSL-25	halabanensis	Lampung	-5.3636	105.1845
SSL-27	halabanensis	Lampung	-5.4238	105.2356
SSL-28	halabanensis	Lampung	-5.4207	105.2226
SSL-29	halabanensis	Lampung	-5.43002	105.2025
SSL-30	malaccensis	Lampung	-5.82062	105.5979
SSL-31	halabanensis	Lampung	-5.46332	105.2368
SSL-32	halabanensis	Lampung	-5.48492	105.2401
SSL-33	halabanensis	Lampung	-5.61308	105.1707
SSL-34	halabanensis	Lampung	-5.7139	105.1787
SSL-35	halabanensis	Lampung	-5.75313	105.1263

2

Supplementary Table 3 Morphological characteristics for the PCA.

Accession	ANS-01	ANS-37	ANS-38	ANS-39	ANS-40	ANS-41	J-01	J-02	J-05	J-09	J-10	J-11	J-12	J-13	J-14	J-24	J-25	J-29	J-30	J-31	RSB-01	RSB-02	RSB-03	RSB-20	RSB-21	RSB-22	SSB-02	SSB-03	SSB-04	SSB-05	SSB-06	SSB-08	SSB-09	SSB-10	SSB-12	SSB-14	SSB-16	SSB-17
ssp/var:	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis	Type I ssp. malaccensis
Province:	North Sumatra	North Sumatra	North Sumatra	North Sumatra	North Sumatra	North Sumatra	Jambi	Jambi	Jambi	Jambi	Jambi	Jambi	Jambi	Jambi	Jambi	Jambi	Jambi	Jambi	Jambi	Jambi	Riau	Riau	Riau	Riau	Riau	Riau	South Sumatra	South Sumatra	South Sumatra	South Sumatra	South Sumatra	South Sumatra	South Sumatra	South Sumatra	South Sumatra	South Sumatra	South Sumatra	South Sumatra
6.2.1 Pseudostem height (m)	2	1	3	3	3	3	2	2	1	2	1	2	3	1	1	2	1	2	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3
6.2.3 Pseudostem colour	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	2	2	2
6.2.8 Wax on leaf sheaths	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	3	3
6.2.8.1 Wax on leaf sheaths	1	1	1	1	1	1	3	3	3	3	3	3	3	2	3	3	3	3	3	3	3	3	3	3	4	3	3	3	2	3	3	1	1	3	3		3	3
6.3.1 Blotches at the petiole base	3	3	2	3	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	2	3	3	3	3	3	3	3	3	2	3	3	3	3	3	3	3	3	3
6.3.2 Blotches colour	1	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	3	3	3	3	3	4	4	3	3	3	4	4	4
6.3.3 Petiole canal leaf III	3	3	3	3	3	3	3	3	3	3	4	3	3	3	3	3	3	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	2	3	1
6.3.6 Petiole margin colour (chart A)	1	2	2	3	2	3	2	3	3	3	3	3	3	3	3	3	2	1	3	3	3	3	1	2	3	3	3	3	3	3	3	1	1	1	3	1		
6.3.9 Leaf blade length (cm)	2	1	2	2	2	1	2	2	1	2	1	1	1	1	3	1	1	2	1	1	1	1	1	1	2	1	1	2	2	2	2	1	1	1	1	1	1	1
6.3.10 Leaf blade width (cm)	2	1	2	2	2	1	2	1	2	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	2	1	1	2	2	2	1	1	1	1	1	1	1	1
6.3.10.11 Leaf ratio	3	2	2	2	2	-	2	2	3	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	-	3	3	3	3	3	3	3	3	3	2	3
6.3.11 Petiole length (cm)	2	1	2	2	2	1	1	1	1	2	2	2	1	2	1	1	2	2	2	2	1	1	1	2	2	1	1	2	2	2	1	1	1	1	1	1	1	1
6.3.15 Appearance of leaf lower surface	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	1	1
6.3.17 Insertion point of leaf blade of petiole	2	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	1	1	2
6.3.18 Shape of leaf blade base	2	2	2	2	2	1	2	2	2	2	3	3	3	3	3	3	2	1	1	1	3	2	1	1	3	2	1	1	3	1	2	1	3	3	3	3	3	2
6.3.20 Colour of midrib dorsal surface	1	1	1	1	1	1	2	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6.3.20.1 Pigmentation on midrib dorsal	1	1	1	1	1	1	2	2	1	2	2	1	2	2	2	1	2	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
6.3.20.2 Color of pigmentation on midrib dorsal	1	1	1	1	1	1	4	4	4	4	4	1	4	4	4	1	4	4	1	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
6.3.21 Colour of midrib ventral surface	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
6.4.1 Peduncle length (cm)	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6.4.3 Peduncle width (cm)	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6.4.4 Peduncle colour	3	3	4	4	4	4	2	2	4	4	3	3	3	2	4	2	2	2	2	2	2	3	3	3	3	3	2	3	2	2	3	1	1	2	2	2	2	3
6.4.5 Peduncle hairness	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	-	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
6.4.5 Peduncle hairness	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	-	3	3	3	1	3	3	3	3	3	1	3	3	3	3	3	3	3	3	3
6.4.6 Bunch position	2	2	1	1	1	2	1	1	2	2	2	2	2	2	2	2	2	3	2	2	3	1	1	1	1	2	1	2	1	2	2	1	2	1	2	1	1	2
6.4.8 Bunch appearance	2	1	2	-	1	1	1	1	1	1	1	2	2	2	2	2	2	3	2	2	1	1	1	2	2	1	2	1	2	2	1	2	1	1	1	1	1	2
6.4.12 Rachis position	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
6.4.13 Rachis appearance	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6.4.15 Male bud shape	2	1	1	1	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6.7.1 Fruit position	1	4	1	-	1	1	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
6.7.2 Number of fruits	2	1	2	-	2	2	1	3	2	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	2	3	3	1	3	1	1	3
6.7.3 Fruit length [cm]	3	1	1	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6.7.4 Fruit shape (longitudinal curvature)	1	1	1	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6.7.5 Transverse section of fruit	-	2	2	-	3	3	1	-	1	-	3	3	3	2	3	2	3	3	2	2	2	3	3	3	3	3	-	3	3	3	3	3	3	3	3	3	3	3
6.7.6 Fruit apex	-	-	-	-	4	4	-	4	4	-	1	4	4	4	4	4	4	4	4	2	4	4	4	2	2	4	4	2	4	4	4	4	4	4	4	4	4	4
6.7.7 Remains of flower relicts at fruit apex	-	1	1	-	1	1	1	3	3	3	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6.7.16 Cracks in fruit peel	-	1	1	-	1	1	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1
6.7.17 Pulp in fruit	-	1	2	-	2	2	-	1	1	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	2	2	1	2	1	2	2	1	1	1	1	1
6.7.24 Seed surface	-	1	1	-	2	2	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-	2	1	1	1	2	1	1	1	1	2	1	1	1	1	1	1
6.7.25 Seed shape	-	2	2	-	2	2	-	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
7 Fruit development	1	2	1	-	2	2	-	-	-	2	2	2	2	2	2	2	2	2	2	2	2	2	-	2	3	2	3	2	3	1	1	2	1	1	1	1	1	1
6.5.1 Bract base shape	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	2	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
6.5.2 Bract apex shape	2	2	2	2	2	2	1	1	1	1	1	2	1	1	2	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
6.5.3 Bract imbrication	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6.5.31 Bract imbrication level	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1
6.5.4 Colour of the bract external face	3	3	3	3	3	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3
6.5.5 Colour of the bract internal face	2	2	2	2	2	2	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
6.5.6 Colour on the bract apex	2	2	2	2	2	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1
6.5.7 Colour stripes on bract	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2
6.5.9 Fading of colour on bract base	1	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	2	1	1	1	2	2	1	1	1	1	1	1
6.5.11 Male bract lifting	2	2	2	2	2	2	2	2	2	3	3	2	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	2	2	2	2	2	2
6.5.13 Wax on the bract	3	2	2	2	2	2	2	1	2	2	2	2	2	2	3	2	2	3	2	2	2	1	1	1	2	1	1	1	1	1	1	1	1	1	1	2	2	3
6.6.2 Compound tepal basic colour	2	1	2	2	2	2	3	3	1	1	2	2	2	2	2	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6.6.3 Compound tepal pigmentation	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6.6.4 Lobe colour of compound tepal	1	2	2	2	2	1	1	2	1	1	1	1	2	1	2	1	1	1	1	1	1	1	1	1	2	1	1	2	2	1	1	1	1	1	2	1	1	1
6.6.8 Free tepal appearance	1	3	3	3	3	3	2	2	2	3	3	3	3	3	-	3	1	-	-	-	3	3	3	3	3	3	-	-	3	3	3	3	3	3	3	3	3	3
6.6.9 Free tepal apex development	2	2	2	2	2	1	1	1	2	2	2	2	2	2	-	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
6.6.13 Anther colour	-	1	1	1	1	1	5	-	-	-	1	-	-	-	5	1	1	-	-	-	5	-	-	-	-	-	5	5	5	5	-	5	2	1	-	1	5	1
6.6.16 Style basic colour	-	1	1	1	1	1	4	4	4	1	1	-	1	-	-	1	1	-	-	-	1	-	-	-	-	-	1	-	-	-	1	1	1	1	1	1	1	1
6.6.17 Pigmentation on style	-	1	1	1	1	2	2	2	1	1	1	-	1	-	-	1	1	-	-	-	1	-	-	-	-	-	1	-	-	-	1	1	1	1	1	1	1	1
6.6.19 Style shape	1	1	1	1	1	1	1	1	1	1	1	-	1	-	-	1	1	-	-	-	1	-	-	-	-	-	1	-	-	-	1	1	1	1	1	1	1	1
6.6.20 Stigma colour	2	2	2	2	1	1	1	1	2	2	1	-	1	5	5	2	1	-	-	-	5	2	5	2	-	2	1	-	-	2	1	2	1	1	1	2	2	2
6.6.21 Ovary shape	2	1	2	2	2	1	1	1	1	1	1	-	1	-	-	2	1	1	-	-	1	1	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6.6.22 Ovary basic colour	1	1	2	2	2	2	1	1	3	2	2	-	-	2	2	2	2	-	-	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1
6.6.23 Ovary pigmentation	1	1	1	1	1	1	-	1	1	1	1	-	-	1	1	1	1	-	-	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

2

Code	Type	Locality
SSB-2-4	Type I ssp. malaccensis	South Sumatra
SSL-30	Type I ssp. malaccensis	Lampung
SSL-01	Type I ssp. malaccensis	South Sumatra
SSL-02	Type I ssp. malaccensis	South Sumatra
SSL-03	Type I ssp. malaccensis	South Sumatra
SSL-22	Type I ssp. malaccensis	Lampung
SSB-22	Type I ssp. malaccensis	South Sumatra
SSB-22	Type IV var. longipetiolata	South Sumatra
SSB-23	Type IV var. longipetiolata	South Sumatra
SSB-26	Type IV var. longipetiolata	South Sumatra
ANS-08	Type II ssp. halabanensis	Aceh
ANS-09	Type II ssp. halabanensis	Aceh
ANS-10	Type II ssp. halabanensis	Aceh
ANS-11	Type II ssp. halabanensis	Aceh
ANS-12	Type II ssp. halabanensis	Aceh
ANS-14	Type II ssp. halabanensis	Aceh
ANS-22	Type II ssp. halabanensis	Aceh
ANS-24	Type II ssp. halabanensis	Aceh
ANS-25	Type II ssp. halabanensis	Aceh
ANS-28	Type II ssp. halabanensis	Aceh
ANS-29	Type II ssp. halabanensis	Aceh
J-21	Type II ssp. halabanensis	Jambi
RSB-05	Type II ssp. halabanensis	Riau
RSB-09	Type II ssp. halabanensis	Riau
RSB-10	Type II ssp. halabanensis	Riau
RSB-12	Type II ssp. halabanensis	Riau
RSB-15	Type II ssp. halabanensis	South Sumatra
SSB-19	Type II ssp. halabanensis	South Sumatra
SSB-21	Type II ssp. halabanensis	Bengkulu
SSB-25	Type II ssp. halabanensis	Bengkulu
SSB-28	Type II ssp. halabanensis	Bengkulu
SSB-29	Type II ssp. halabanensis	Bengkulu
SSB-31	Type II ssp. halabanensis	Lampung
SSL-05	Type II ssp. halabanensis	Lampung
SSL-06	Type II ssp. halabanensis	Lampung
SSL-10	Type II ssp. halabanensis	Lampung
SSL-17	Type II ssp. halabanensis	Lampung
SSL-18	Type II ssp. halabanensis	Lampung
SSL-21	Type II ssp. halabanensis	Lampung
SSL-25	Type II ssp. halabanensis	Lampung
SSL-27	Type II ssp. halabanensis	Lampung
SSL-31	Type II ssp. halabanensis	Lampung
SSL-32	Type II ssp. halabanensis	Lampung
SSL-33	Type II ssp. halabanensis	Lampung
SSL-35	Type II ssp. halabanensis	Lampung
ANS-05	Type III ssp. sumatrana	North Sumatra
RSB-13	Type III ssp. sumatrana	West Sumatra
SSL-19	Type III ssp. sumatrana	Lampung
ANS-15	Type V	Aceh
ANS-16	Type V	Aceh
ANS-17	Type V	Aceh

(The page continues with a dense morphological character-state matrix of single-digit values, one column per specimen above and one row per character; individual cell values are not reproduced here with certainty.)

85

Supplementary descriptions

Musa acuminata subsp. *malaccensis* (Ridl.) N.W.Simmonds (Type I)

Plant suckering freely, vertical. Mature pseudostem 2–4.5 m tall, to 15-30 cm in diam. at base, covered with dead remains of older sheaths. Sheaths light green-yellow with reddish, brown, purplish, or blackish blotches, underlying colour cream with very little to moderately waxy, sap milky or watery. Petiole 40-60 cm long, petiole canal straight with erect margins or margins curved inward, petiole margin colour green, pink-purple to red or purple to blue. Lamina lanceolate to oblong, 150-250 × 35-70 cm, base asymmetric or symmetric, both sides rounded or one side rounded, one side pointed or both sides pointed, apex truncate, abaxial light green with waxy or nonwaxy, adaxial green, colour of abaxial midrib mostly yellow or green, colour of adaxial midrib light green. Inflorescence peduncle 20-60 cm long, 3-12 cm in diam., green to dark green with brown blotches, pubescent. Rachis bare, mostly falling vertically or sub-horizontal. Male bud like a top, 5-9×10-18 cm, bract shape medium to large shoulder, apex slightly pointed or pointed , slightly waxy, 5-9×10-18 cm, outer bract overlapping younger ones at apex of bud, external bract red to red purple sometimes with yellow streaks, internal bract whitish to red, apex acute, sometimes tinted with yellow, lifting one or two or more bracts at a time, revolute before falling. Initial flowers carpellate with angular ovary, yellowish green. Ovules in 4 rows per locule. Stigma flat, 6 × 6 mm, greyish. Style 3.2 cm long, creamy. Compound tepals white or cream, with 2 prominently thickened keels, light yellow, apex deep yellow and curved outward. Free tepal ovate, translucent white, apex acute, little or no visible sign of tip development. Anther in the same level of the lobes on the compound tepal. Staminodes 5, filaments creamy. Later flowers pistillate with anther lobes white, cream, rusty brown or pinkish. Ovary creamy, without additional pigmentation. Stigma creamy. Fruit bunch horizontal, compact, with 5-9 hands. Fruits in 2 row, appearing as 2 rows when immature. Individual fruit straight or straight in the distal part, up to 10 cm long (including pedicel), 1-1.5 cm in diameter, prominently ridged, apex obtuse. Pedicel up to 1 cm long, straight. Immature fruit peel green, not waxy. Fruit pulp white, up to 180 seed per fruit, angular, wrinkled, 3.1-6.6 mm long, 3.7-5.4 mm wide, 3.6 mm depth.

Musa acuminata subsp. *halabanensis* (Meijer) Hotta (Type II)

Plant suckering freely. Mature pseudostem 2-4 m tall, to 15-30 cm in diam. at base, covered with dead remains of older sheaths. Sheaths mostly brown-black, black-purple, green-yellow with brown-purplish or black-purplish blotches, underlying colour cream, very little or no visible sign of wax, sap mostly milky or watery. Petiole 40-60 cm long, petiole canal leaf straight with erect margins, petiole margin colour pink-purple to red or purple to blue. Leaves lamina lanceolate to oblong, 150-250 × 35-70

cm, base asymmetric or symmetric, both sides rounded or one side rounded, one side pointed or both sides pointed, apex truncated, abaxial light green and shiny, adaxial green and shiny, colour of abaxial midrib mostly yellow or light green with pink-purple or purple to blue, blotches colour of adaxial midrib light green. Inflorescence initially horizontal rather abruptly turning down at male bud stage. Peduncle 20-40 cm long, 3-7 cm in diam., green to dark green or with brown blotches, slightly pubescent to very pubescent. Rachis bare, mostly with a curve or falling vertically. Male bud ovoid, 7-9 × 9-15 cm. Bract shape medium shoulder, apex mostly intermediate or obtuse, moderately waxy, ca. 15 × 7 cm, mostly not imbricate, old bract overlap at apex of bud or slightly imbricate, external bract purple-brown, internal bract whitish, apex obtuse, not tinted with yellow colour, lifting one or two or more bracts at a time, revolute before falling. Basal flowers carpellate, prominently ridged, creamy or whitish. Ovules in 4 rows per locule. Stigma flat, 6 × 6 mm, greyish. Style 3.2 cm long, creamy;. Compound tepals white or cream, with 3 prominently yellow lobes, apex deep yellow and curved outward. Free tepal ovate, translucent white, apex acute with developed sign of development. Anther in the same level of the lobes on the compound tepal. Staminodes 5, filaments creamy, anther lobes white, cream. Ovary creamy, without additional pigmentation. Stigma yellow. Fruit bunch horizontal, mostly lax if the fruit is not developed to very compact if the fruit is developed and bearing seed, with 7-10 hands, fully developed with seed but sometimes initial hands or even most hands not developing into fruit. Fruits in 2 rows, 14-18 individual fruits per hand. Individual fruit straight, up to 10 cm long (including pedicel), 1-1.5 cm in diam., prominently ridged, apex bottle-necked, slightly sessile. Immature fruit peel green, not waxy, fruit pulp white, up to 300 seed to per fruit, globular, mostly smooth, 3.1-6.6 mm long, 3.7-5.4 mm width, 3.6-5 mm depth.

Musa acuminata Colla var. *longepetiolata* Nasution (Type IV)

Plant suckering freely; vertical. Mature pseudostem 2-3 m tall, to 15-25 cm in diam. at base, covered with dead remains of older sheaths. Sheaths light green with reddish, brown or purplish or blackish blotches, underlying colour cream, very little to moderately waxy; sap milky or watery. Petiole 60-100 cm long, petiole canal leaf straight with erect margins, petiole margin colour green, pink-purple to red or purple to blue. Leaves: lamina lanceolate, 120-250 × 35-50 cm, base asymmetric or symmetric, one side pointed or both sides pointed, apex truncate; abaxial light green, waxy, adaxialgreen; colour of abaxial midrib mostly green, colour of adaxial midrib light green. Inflorescence horizontal, bunch mostly moderately compact; peduncle 20-60 cm long, to 4-6 cm in diam., green to dark green with brown blotches, pubescent; rachis bare or neutral flowers, mostly falling vertically or horizontal; Male bud like a top, to 5-9×10-15 cm; bract shape medium to large shoulder, apex slightly pointed or pointed , slightly waxy, to 5-9×10-18 cm, not imbricate, old bract overlap at apex of bud or young bract

slightly overlap, external bract red to red purple, internal bract whitish to red, apex acute, mostly not tinted with yellow colour or tinted with yellow, lifting one or two or more bracts at a time, revolute before falling, bract mostly very little or few wax; basal flowers carpellate, prominently ridged, yellowish green; ovules in 4 rows per locule; stigma flat, 6×6 mm, grayish; style 3 cm long, creamy; compound tepals white or cream, with 2 prominently thickened keels, light yellow, apex deep yellow and curved outward; free tepal ovate, translucent white, apex acute; little or no visible sign of development or developed; anther in the same level of the lobes on the compound tepal; staminodes 5, filaments creamy, anther lobes white, cream, rusty brown or pinkish; ovary creamy, without additional pigmentation; stigma creamy. Fruit bunch horizontal, compact, with 5-9 hands; fruits in 2 row, appearing as 2 rows when immature; individual fruit straight or straight in the distal part, up to 10 cm long (including pedicel), 1-1.5 cm in diam., prominently ridged, apex obtuse; pedicel up to 1 cm long, straight; immature fruit peel green, not waxy, fruit pulp white; seed up to per fruit, angular, wrinkled, 3-6 mm long, 3-5 mm width, 3.5 mm depth.

Type V
Plant suckering freely. Mature pseudostem 2.5-3 m tall, to up to 25 cm in diam. at base, covered with dead remains of older sheaths. Sheaths black-purple, underlying colour cream, very little to moderately waxy; sap milky. Petiole 40-50 cm long, to 2-3 cm in diam., petiole canal with erect margins, petiole margin colour purple to blue. Leaves lamina lanceolate, 150-200 × 35-60 cm, base symmetric, both sides rounded, apex truncate, abaxial light green with waxy, adaxial green not waxy, colour of abaxial midrib pink purple, colour of adaxial midrib pink purple. Inflorescence at an angle, bunch lack, peduncle 30-40 cm long, to 3-5 cm in diam., green to dark green or with brown blotches, pubescent; rachis bare. Male bud rounded, to 7-8 × 10-13 cm. Bract shape small shoulder, apex obtuse, waxy, to 7-8 x 10-13 cm, imbricate, young bract slightly overlap, external bract purple-brown, internal bract red with yellow pigmentation, not tinted with yellow colour, lifting one or two or more bracts at a time, not revolute before falling, moderate wax. Ovary cream. Ovules in 4 rows per locule. Compound tepals whitish, with 3 prominently thickened keels, light yellow; free tepal translucent white, ovate, apex acute with developed sign. Staminodes 5, filaments whitish. Fruit bunch with 8-9 hands, fruits in 2 row, 20-25 fruits. Individual fruit straight, up to 15 cm long (including pedicel), 1-1.5 cm in diam., apex blunt-tipped; pedicel ca. 2 cm long, straight or straight in the distal part. Immature fruit peel green, not waxy, fruit pulp white, seed up 200 to per fruit, flat, 6.4-7.2 mm long, 4-5.6 mm wide, 2.9-3.6 mm depth.

Chapter 3.

Male meiosis and pollen morphology in diploid Indonesian wild bananas and cultivars

Ahmad F, Poerba YS, Kema GHJ & de Jong H

Abstract

Breeding of banana is hampered by its genetic complexity, structural chromosome rearrangements and different ploidy levels. Various scientific disciplines, including cytogenetics, linkage mapping, and bioinformatics, are helpful tools in characterising cultivars and wild relatives used in crossing programs. Chromosome analysis still plays a pivotal role in studying hybrid sterility and structural and numerical variants. In this study, we describe the optimisation of the chromosome spreading protocol of pollen mother cells focusing on the effects of standard fixation methods, duration of the pectolytic enzyme treatment and advantages of fluorescence microscopy of DAPI stained cell spreads. We demonstrate the benefits of this protocol on meiotic features of five wild diploid *Musa acuminata* bananas and a diploid (AA) cultivar banana "Rejang", with particular attention on pairing configurations and chromosome transmission that may be indicative for translocations and inversions. Pollen slides demonstrate regular-shaped spores except "Rejang", which shows fertile pollen grains of different size and sterile pollen grains, suggesting partial sterility and unreduced gamete formation that likely resulted from restitutional meiotic divisions.

Keywords: Cytogenetics, translocation, *Musa acuminata* ssp. *malaccensis*, protocol, pollen mother cells, pollen

Introduction

The genus Musa L. consists of about 70 species (Häkkinen, 2013), of which *Musa acuminata* Colla (A genome) and *M. balbisiana* Colla (B genome) are the leading representatives of cultivated bananas (Simmonds & Shepherd, 1955). Most of the banana's wild relatives are diploid with 2n = 2x = 22 chromosomes (Simmonds & Shepherd, 1955), whereas cultivated forms can also be triploid or tetraploid as resulted from intra- or interspecific hybrids. Wild bananas are seeded with little pulp and so not appropriate for consumption, whereas most cultivated forms are seedless, parthenocarpic or sterile (Heslop-Harrison & Schwarzacher, 2007).

Banana is one of the essential fruits globally and is cultivated in tropical and subtropical countries. It is an important staple food with an annual production above 100 million tons and plays a vital role in its growers' social and cultural aspects (FAO, 2020). However, banana is vulnerable to bacterial, viral and fungal pathogens, urging geneticists and breeders to search for disease-resistant varieties (Ploetz, 2005; Blomme et al., 2011). Despite various breeding initiatives (Bakry et al., 2009; Khayat, 2020) and genetic modification (Dale et al., 2017), introgression of desired traits into cultivated forms is still very demanding, if not impossible (Ortiz & Vuylsteke, 1995). Factors that hamper practical banana breeding include genetic complexity, structural chromosome rearrangements and unequal ploidy levels of appropriate crossing parents (Simmonds & Shepherd, 1955; Shepherd, 1999), often leading to hybrid sterility, partial pollen fertility, linkage drag and crossing incompatibility (Fortescue & Turner, 2004; Roux et al., 2004; Morán, 2013). Various studies revealed the relevance between fertility with abnormal chromosome morphology and balanced meiotic segregation in wild *M. acuminata*, *M. balbisiana* and their hybrids (Dodds, 1943; Wilson, 1946a; Shepherd, 1999). Most cultivated triploid bananas exhibit abnormal gamete production due to unbalanced chromosome segregation at meiosis, which contrasts to most wild bananas, which are in general fertile (Dodds, 1943; Shepherd, 1999). To this end, cytogenetic analyses of chromosome pairing and transmission at male meiosis of the parental and hybrid genotypes are required to establish the course of chromosome behaviour during microsporogenesis (Dodds, 1943; Shepherd, 1999).

Despite several long-standing cytogenetics studies of banana (Dodds, 1943; Wilson, 1946a,b,c; Shepherd, 1999), the skill of producing high-quality microscopic preparations is still troublesome and time-consuming. Although various authors (Adeleke et al., 2002; De Capdeville et al., 2009; Kantama et al., 2017) claim significant improvements in making high-quality chromo-some preparation, the presence of dense cytoplasm in pollen mother cells at meiotic stages is still a major challenge in getting high-quality chromosome spreads.

In this study, we aim to further optimise chromosome preparation techniques of pollen mother cells at different meiotic stages to analyse selected Indonesian bananas. We focus on the wild *M. acuminata* ssp. *malaccensis*, a wild banana that confers resistance to one of the most devastating banana diseases, Fusarium wilt of banana (FWB). The disease is caused by a plethora of different Fusarium species (Maryani et al., 2019), of which *F. odoratissimum*, comprising the so-called Tropical Race 4 (TR4), is currently threatening "Cavendish" bananas as well as many other varieties destined for domestic markets (Ordóñez et al., 2015; Maryani et al., 2019). Hence, this species is considered a favourite for FWB resistance studies or for obtaining genetic improvement (Peraza-Echeverria et al., 2008; Kayat et al., 2009; D'Hont et al., 2012; Dale et al., 2017). Beside this accession, we also included some wild *M. acuminata* and cultivated bananas for comparison. We considered common fixations as part of slide optimisation to reduce cytoplasmic background and improved incubation time with pectolytic enzyme solution for better cell spreading. We analysed chromosome pairing abnormality, chromosome segregation, unreduced gamete and pollen viability. Eventually, we discussed the potential usage of this method for more intensive banana cytogenetics and insight into how cytogenetics can drive banana breeding programs in the future.

Materials and methods

The plant samples were obtained from the Research Center for Biology's living collection, Indonesian Institute of Sci- ences (LIPI), including *M. acuminata* ssp. *malaccensis* (LIPI-010), var. *breviformis* (LIPI-218), var. *sumatrana* (LIPI-457), var. *tomentosa* (LIPI-172), and var. *zebrina* (LIPI-043). Besides, we used the *Musa* AA "Rejang" (LIPI- 048, (Sucrier subgroup) and *Musa* ABB "Pisang Kepok" (LIPI-125, Saba subgroup) cultivated bananas, and one *M. balbisiana* (BB) accession "Pisang Klutuk" (LIPI-054). Detailed information on the material is given in Table 1.

Chromosome preparations

We collected the male buds during September 2014, and March and December 2015, and always between 10 am and 3 pm. We used anthers of 20 mm or less in the younger parts of the male bud with pollen mother cells and microsporocytes at different meiotic stages. The material was directly fixed in one of the following solutions: a) freshly prepared ethanol (96%): glacial acetic acid (3:1), b) methanol (100%): acetic acid (3:1) and c) Carnoy's solution (9 ethanol: 3 chloroform: 1 acetic acid). The volume of each fixative was about 20 × the volume of the materials.

Table 1 Details of accessions used in this study.

| Banana	Accessions	Genome	Collection site	Details
Musa acuminata ssp. *malaccensis*	LIPI-010	AA wild	Sumatra	A wild *Musa acuminata*, which is resistant to Fusarium wilt Tropical Race 4 (TR4) and Race 1 (Ahmad et al., 2020a), comparable to the fully sequenced resistant ssp. *malaccensis* DH "Pahang" (D'Hont et al., 2012; Martin et al., 2016)
M. acuminata var. *breviformis*	LIPI-218	AA wild	Java	A wild *M. acuminata* in lowland with green pseudostem and leaf
M. acuminata var. *sumatrana*	LIPI-457	AA wild	Java	A wild *M. acuminata* in mountainous rain forests with hanging verti- cal bunch
M. acuminata var. *tomentosa*	LIPI-172	AA wild	Sulawesi	A wild *M. acuminata* with light green leaves and pseudostem
M. acuminata var. *zebrina*	LIPI-043	AA wild	Java	A wild *M. acuminata* with reddish coloured pseudostem in young leaves and the lower surface of the leaves
"Rejang"	LIPI-048	AA cv	Sumatra	Fusarium TR4 resistant banana with slim fruits of about 10 cm and sweet taste
"Pisang Kepok"	LIPI-125	ABB cv	Java	A popular cooking banana in Indonesia with high starch content and resembles members of Saba sub-group
M. balbisiana "Pisang Klutuk"	LIPI-054	BB wild	Java	A wild *M. balbisiana*, commonly grown in the villages for wrapping

3

Fixation time was approximately one hour at room temperature. We replaced the fixative 1–2 times until the inflorescences and solution remain clear. Finally, we transferred the material to 70% ethanol for more extended storage at 4 °C. On the day of slide preparation, we selected the anthers and washed them three times in Milli-Q water (MQ-water) for one minute. Five to ten anthers were put in a 1.5 mL plastic tube, filled with 200 µL pectolytic enzyme solution (0.2% pectolyase Y23 (Sigma P-3026), 0.2% cellulase RS (Yakult 203,033, Yakult Pharmaceutical, Tokyo, Japan) and 0.2% cytohelicase (Bio Sepra 24,970–014)), in 10 mM sodium citrate buffer (pH 4.5), and kept for 1–5 h at 37 °C, depending on the condition of the material and the enzymatic activity. Subsequently, we washed the anthers three times with MQ-water and kept them on ice until use. For the chromosome spreading step, we transferred a 1–2 mm part of an anther onto a clean glass slide along with a tiny amount (c. 5 µL) of water. The anther was then tapped with a glass rod or needle, after which the pollen mother cells (PMCs) were squeezed out and the larger clumps of supporting tissues removed.

We then added 2–3 drops of 50% acetic acid, covered with a coverslip and incubated the cells for 5–10 min at 45 °C, while adding small amounts of acetic acid to avoid drying out of the preparation. We then squashed the material with a needle and thumb. We removed the coverslip after freezing the slide in liquid nitrogen and let it air-dry. Alternatively, we squeezed the PMCs in acetic acid 50%, then added 2–3 drops of freshly prepared ethanol—acetic acid (3:1) and flame dried the slide. Slides were stained with 100 μg/mL DAPI (4′,6-diamidino-2-phenylindole) in Vectashield® (Vector Laboratories) mounting solution (1:20) for analysis under a Zeiss fluorescence microscope equipped with appropriate filters and epifluorescence illumination. During the slide screening, we found that pollen mother cells at pachytene were most common.

In contrast, we counted 142 cells at the diakinesis/metaphase I stage, 29 cells at anaphase I and five cells at anaphase II. Digital images were eventually optimised for con- trast, brightness and sharpness in Adobe Photoshop (Kantama et al., 2017). Pollen morphology and viability were established in pollen slides stained in a drop of Lactophenol acid fuchsin (Sass, 1964). Images were captured of 100–150 pollen grains using a bright-field microscope. Pollen grains were measured with Fiji/ImageJ (https:// imagej. net/ Fiji) and data processed in Microsoft Excel.

Results

Our first goal was to optimise our protocol for minimum cytoplasm, and improved chromosome morphology in cell spread slides containing pollen mother cells (Figure 1). Most crucial is the effect of fixation and maceration in the proto- col. The best cell spreading and chromosome differentiation was obtained with the methanol—acetic acid fixative (Figs. 1H–I). The ethanol—acetic acid mixture, on the other hand, produces brighter staining of the cytoplasm and blurred chromosome boundaries (Figs. 1D–E), whereas Carnoy's (ethanol—chloroform—acetic acid) resulted in an even more opaque cytoplasm in which chromosomes are hardly discernible (Figs. 1F–G). Further optimisation was obtained with the incubation of the anthers in pectolytic enzymes. Treatment of two hours resulted in well-spread PMCs (Figure 1A). In contrast, three hours' treatments often resulted in damaged cells (Figure 1B), an artefact that becomes more severe when the enzyme treatment was extended to 4–5 h (Figure 1C).
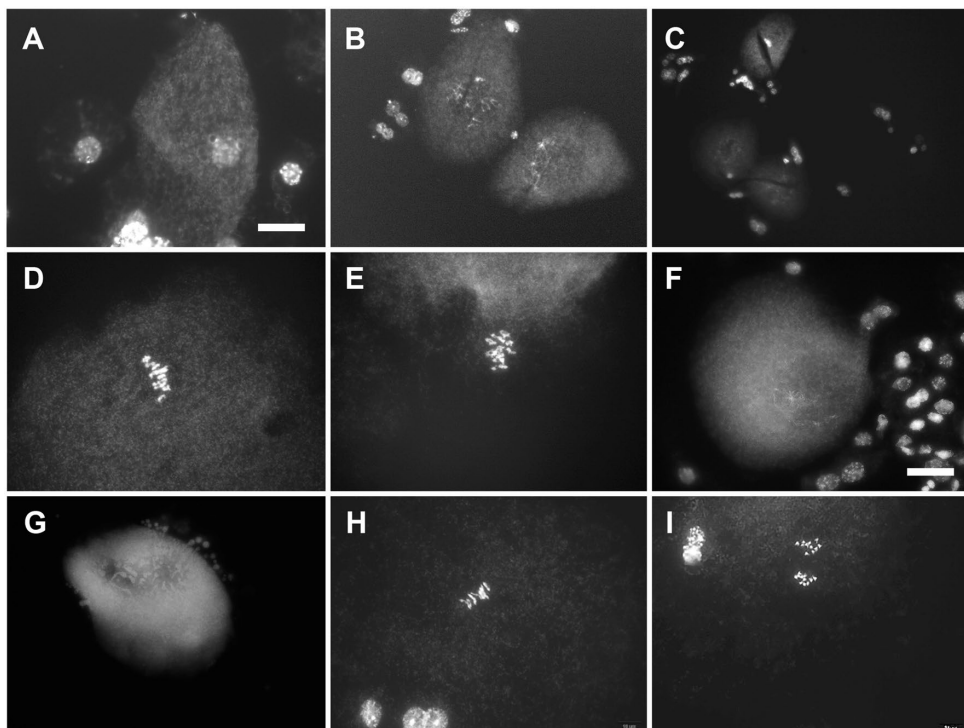
Figure 1 Optimisation of the chromosome spreading protocol. A-C: Effect of incubation time of the pectolytic enzyme solution on morphology of the pollen mother cells (PMCs) of *Musa acuminata* ssp. *malaccensis*. A Intact PMCs after 2 h of incubation; B After 3 h of enzymes treatment, cells still have relatively good and round shape, C Incubation time of 4–5 h. D-I. Effect of fixations. The cells in D and E, which were fixed in ethanol—acetic acid (3:1) solution, display relatively strong background of the cytoplasm, and hence ham per clear structures in nuclei and chromosomes. The cells in F and G came from anthers that have been fixed in Carnoy's solution (ethanol—chloroform—acetic acid (9:3:1)), which result in even denser cytoplasmic background and also less distinct chromosomes. In H and I we use a methanol—acetic acid (3:1) fixative instead, that produce a more transparent background. Scale bars in A, B, F and G equal 20 μm. Scale bar for C equals 30 μm. Scale bars in D, E, H and I equal 10 μm.

Microscopic observations of pollen mother cells and pollen

The first meiotic prophase stage with discernible chromosomes in *Musa* is pachytene, in which homologous chromosome segments are fully paired. Chromosomes display a delicate pattern of numerous minor heterochromatic knobs or chromocenters in the euchromatic part and denser, brighter fluorescing segments in most heterochromatic

pericentromeres. We measured the length of full pachytene complements in a few cells showing values of 207, 293, 329 and 339 μm, respectively. In pachytene cells of the ssp. *malaccensis* and var. *breviformis* varieties, we observed unpaired regions of the chromosomes (Figs. 2A, 2C) or longer unpaired segments (Figure 2D), such as in "Rejang". In ssp. *malaccensis*, we also detected a region with unpaired chromosomes that suggest a lack of homology or a pairing partner switch.
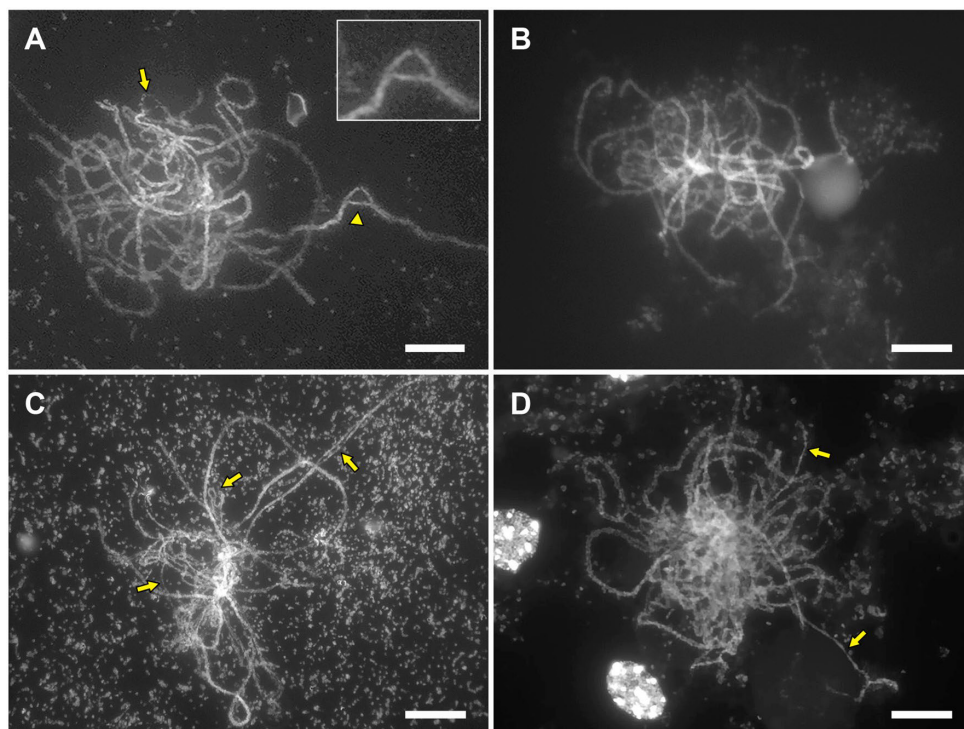
Figure 2 Pachytene complements in pollen mother cells of wild Musa acuminata and "Rejang". **A** *M. acuminata* ssp. *malaccensis*. Chromo- somes display numerous small chromomeres, in addition to brighter fluorescing segments that represents the pericentromere heterochromatin. The arrowhead denotes an open chromosome confirmation indicative of a paracentric inversion; the arrow shows a part of the chromosomes that are unpaired. **B** Var. *zebrina*. Fully paired cell complement. **C** Var. *breviformis*. This pollen mother cell is at late zygotene / early pachytene stage with several chromosome regions not fully paired (arrows). **D** "Rejang". The arrows show examples of unpaired chromosome regions. Scale bars equal 10 μm in the Figures.

Chromosomes at diplotene are largely decondensed, forming diffuse networks of threads (Figure 3A). Their axes are no longer distinguishable, and hence is this stage

not suitable for establishing chromosome configurations. After the diffuse diplotene, chromosomes recondense and display characteristic chromosome configurations, most of which with chiasmata in both chromosome arms (ring bivalent) or in one arm (rod bivalents). Incidentally, we observed univalent pairs and configurations that resemble trivalents and quadrivalents or overlapping bivalents and/or univalents (Figs. 3B−F). In "Rejang", we detected cell complements with several univalents (Figs. 3G−H). The average number of bivalents per cell in the ssp. malaccensis, var. zebrina, var. sumatrana and "Rejang" were 10.19, 10.43, 10.33 and 10.81, respectively (Table 2). Few of the cell complements of ssp. malaccensis, var. breviformis and "Rejang" clearly showed one or more univalents. In "Rejang" univalent pairs occurred at 0.44 per cell. Several cases of overlapping biva- lents were noted in the ssp. malaccensis, var. zebrina and "Rejang", structures that resemble alternate ring quadrivalents (Figure 4A), "Figure-8" like (Figure 4B) and adjacent ring quadrivalents (Figure 4C).

Table 2 Mean of chromosome configurations of univalent and bivalent at diakinesis/metaphase I in wild *Musa acuminata* and "Rejang". Some of the chromosome configurations that could not be interpreted unequivocally were not considered and so making the total number of chromosomes.

Banana	No. of cell	Diakinesis I	
		I	II
ssp. *malaccensis*	93	0.02	10.19
var. *zebrina*	7	0.00	10.43
var. *sumatrana*	15	0.00	10.33
var. breviformis	11	0.36	10.73
"Rejang"	16	0.44	10.81

Chromosome segregation at anaphase I was in most cases balanced, forming two groups of 11 chromosomes in cell spreads of ssp. *malaccensis*, var. *zebrina*, var. sumatrana. In "Rejang", we found unbalanced segregations (Table 3). Bridges and lagging chromosomes were incidentally found in ssp. *malaccensis* and "Rejang" (Figs. 5B−C). The very short metaphase II and anaphase II stages were only found in anthers of ssp. *malaccensis*. Eight pollen mother cells at metaphase II showed balanced sets of 11 + 11 chromosomes and three cells at anaphase II with four sets of four chromatids (11 + 11 + 11 + 11). Furthermore, in a sample of 16 cells at the tetrad stage in ssp. *malaccensis*, we counted one triad, six dyads and nine normal tetrads.
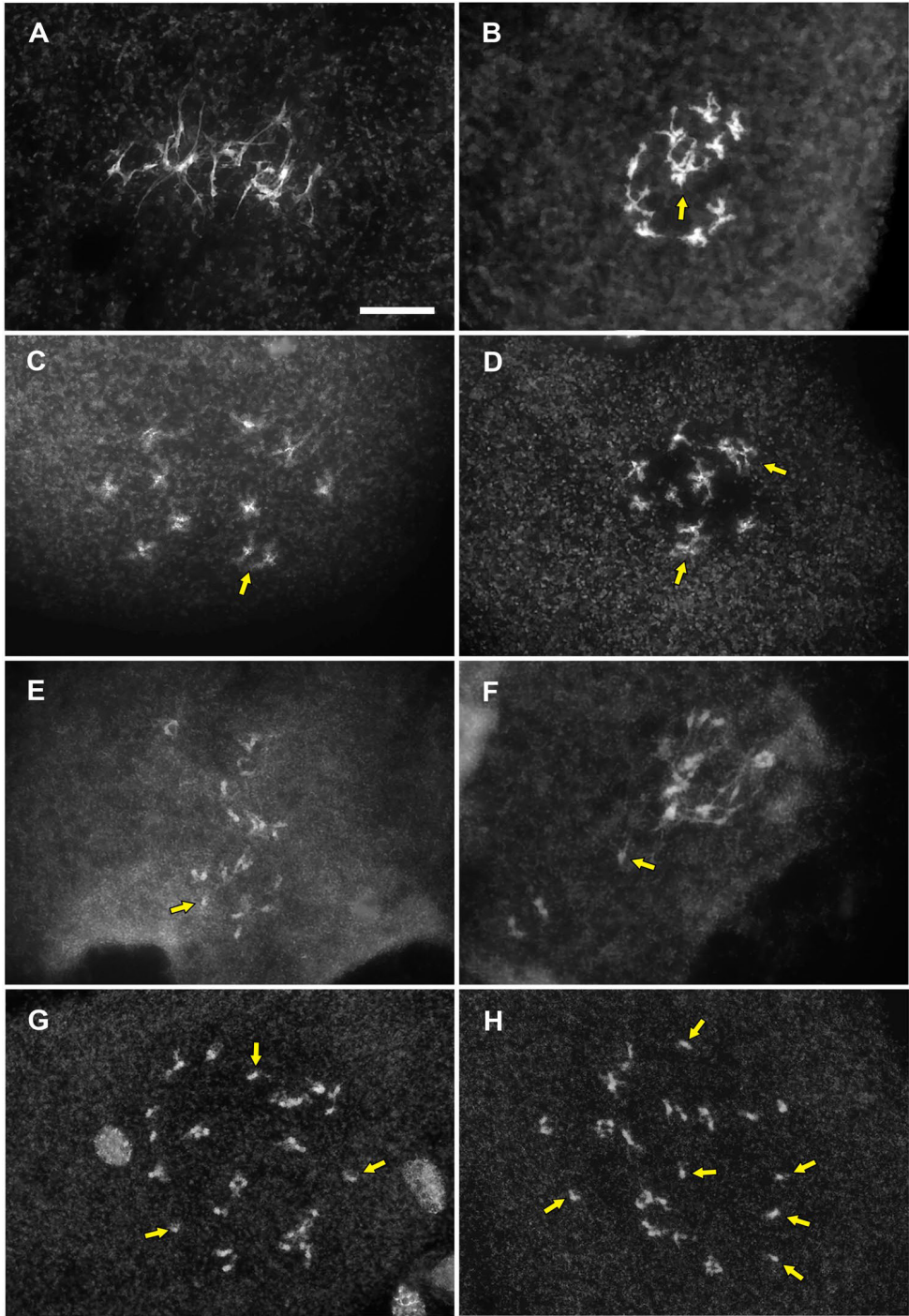
Figure 3 Pollen mother cells at late prophase I (diplotene— diakinesis) of wild *Musa accuminata*

and "Rejang". **A** Diffuse diplotene of a pollen mother cell of var. *tomentosa*. **B** Late diplotene cell of ssp. *malaccensis*. The arrow points at a quadrivalent. **C** Late diplotene of ssp. *malaccensis*. The arrow shows two univalents. **D** Late diplotene of var. *zebrina*. The two arrows show groups of 4 chromosomes, which may represent quadrivalents. **E–F** Late diplotene of ssp. *malaccensis*. The arrows show two univalents. **G–H** Late diplotene of "Rejang". The cell complement here is tetraploid (2n = 4x = 44). Arrows indicate univalents. Magnification in all Figures is the same. Scale bar equals 10 µm.

3
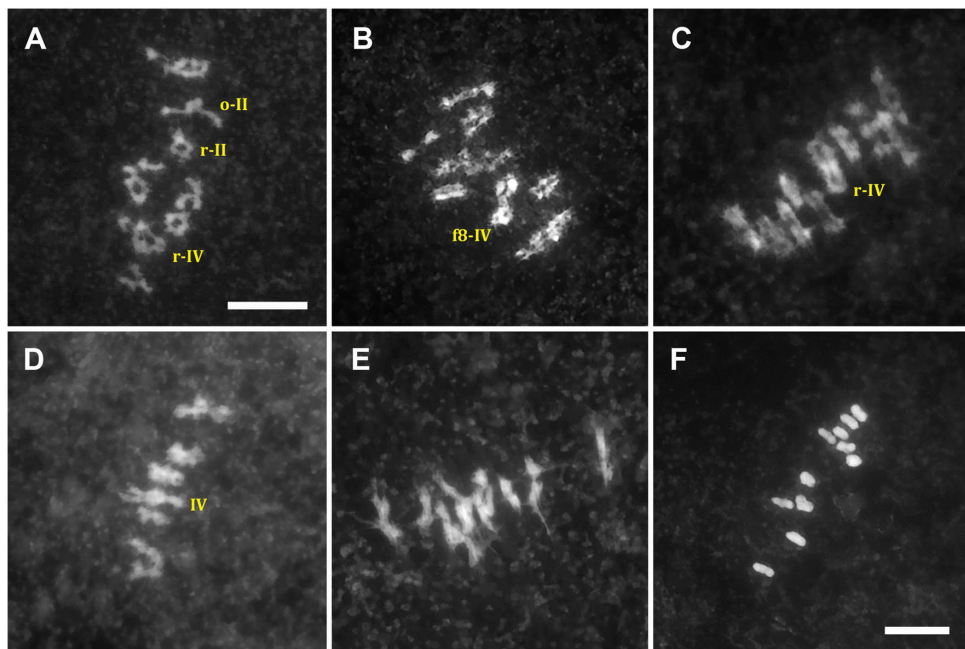
Figure 4 Pollen mother cells at (pro)metaphase I of wild *Musa accuminata*. **A–C**: Ssp malaccensis. A Prometaphase I cell that clearly shows ring bivalents (rII), with chiasmata in both chromosome arms and rod bivalents (o-II) with (mostly) one chiasma in one of the arms. The arrow shows the r-IV that represents a ring quadrivalent with associations of the four chromosomes belonging to one translocation complex. **B** In this cell a rare type of a quadrivalent is visible, the arrow shows the representing a real "Figure 8"-IV configuration, which can be formed when both interstitial translocation segments are large enough to have a chiasma simultaneously. **C** In this cell nine bivalents and one quadrivalent. The arrow shows one adjacent oriented ring-IV. **D** Var. *sumatrana*. Incomplete metaphase I cell with one quadrivalent (IV). **E** Var. *tomentosa* cell with 11 bivalents. **F** Var. *zebrina* late metaphase I cell with 11 bivalents. Scale bars equal 10 µm. Magnification in the Figures A–E is the same.

Table 3 Mean number of chromosomes with normal balanced (dip- loid like) segregation, bridges and laggards at anaphase I.

Banana	No. of cell	Anaphase I, mean of		
		Normal segregation	Bridge	Laggards
ssp. *malaccensis*	12	10	0.08	0
var. *zebrina*	5	11	0	0
var. *sumatrana*	8	11	0	0
"Rejang"	4	8.25	0.25	0.75
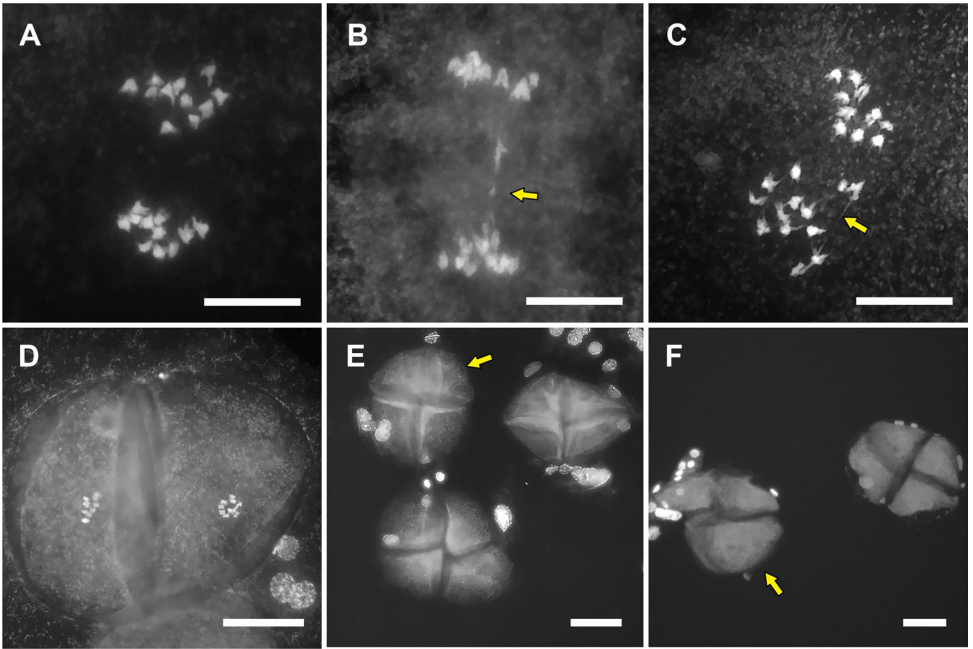
Figure 5 Pollen mother cells at anaphase I-II of wild *Musa accuminata*. A Ssp. *malaccensis*. Balanced segregation of 11 + 11 chromosomes at anaphase I. B Ssp. *malaccensis*. Anaphase I bridge (arrow), which most likely resulted from a crossover in a paracentric inversion loop. The expected acentric fragment produced during the same recombination event is not visible here. C "Rejang". Example of an unbalanced (10 + 1 + 11) chromosome segregation. The lagging chromo- some in the equatorial plane (arrow) of this anaphase I cell will most likely not be included in one of the daughter cells and so get lost as a micronucleus, or is included in the bottom group of chromosomes maken the segregation 10 + 12. D Ssp. *malaccensis*. Pollen mother dyad at interkinesis, with 11 + 11 balanced chromosome segregation (not all chromosomes are in the focal plane). E Var. *tomentosa*. Tet- rad stage. One of the pollen mother cells has no anaphase II in one part of the dyad and hence forms a larger unreduced cell complement (arrow). F Ssp. *malaccensis*. Also, in this example one half of the dyad did not form a callose wall and is supposed

to contain an unre- duced cell complement. Scale bars in A, B and C equal 10 μm. Scale bar in D equals 20 μm. Scale bars in E and F equal 30 μm.

The lactophenol acid fuchsin staining of fixed anthers was used for establishing pollen morphology and size. In most cases, pollen displayed regularly shaped grains, most likely representing viable spores, but we also observed few shrunken, irregular unstained pollen grains (Figs. 6A, B). The percentages of normal stained and regular pollen amounted to 99.5, 99.8, 99.6, 98.7 and 77.7% for the ssp. *malaccensis*, var. *breviformis*, var. *sumatrana*, and var. *zebrina* and "Rejang", respectively. Average pollen diameter varies from 85–95 μm in ssp. malaccensis to 100–115 μm in "Rejang", in "Rejang". We see a vast range of pollen size (Figure 6).
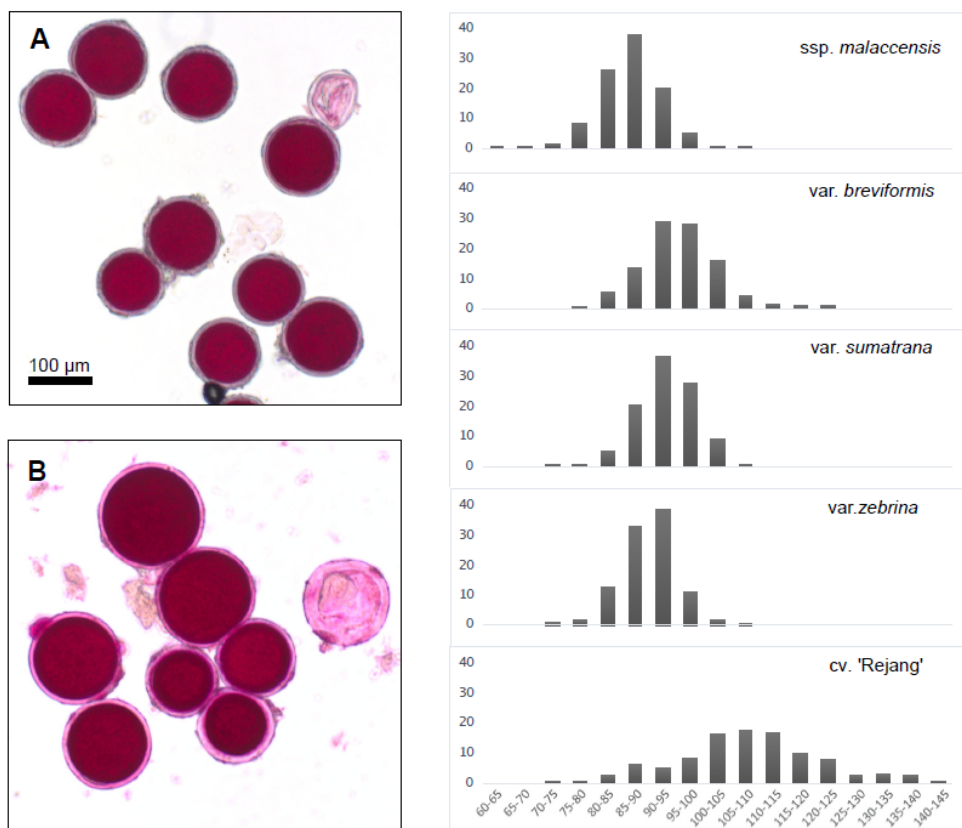
Figure 6 Pollen morphology of lactophenol fuchsin stained pollen samples of wild *Musa accuminata* and "Rejang". A Ssp. *malaccensis*, B "Rejang". Viable pollen was stained, in contrast to the dead pollen that is often shrunken, irregular shaped and unstained (arrow). Right: Graphical representation of the percentage distribution (Y-axis) of pollen grain size (X-axis) for the five bananas.

Discussion

In this study, we optimised cell spreading conditions for six Indonesian bananas to establish clear interpretations of meiotic abnormalities. The starting point was using a technique that features short fixation, enzymatic cell wall digestion and maceration of the pollen mother cells, cell spreading followed by air-drying and finally, DAPI staining for high-resolution fluorescence microscopy. The most challenging hurdle was the opaque cytoplasm of microsporocytes, in which phenolic and polysaccharide compounds mask chromosomal details from observation. We found that methanol-acetic acid (3:1) rather than the standard ethanol acetic acid (3:1) mix decreases cytoplasmic background, although some granular background always remains (De Capdeville et al., 2009). Previous studies claimed the superior property of methanol for fixation and low cytoplasmic background in mouse and human lymphocytes (Levitt & King, 1987). Besides, Iovene et al. (2008) used methanol instead of ethanol for improved cell spreading and well-differentiated morphology of tomato and potato pachytene complements. Bakry & Shepherd (2008) also used a fixation of 1 part ethanol: 4 parts glacial acetic acid: 5 parts water for banana root tip material. Still, we could not test this method in our study. The second optimised step in this study was pectolytic enzymes' time to digest anther's callose walls to obtain well-spread PMCs. The optimum incubation time of two hours was in the range of that used by Kantama et al. (2017) but differed slightly in the methods of Adeleke et al. (2002) and De Capdeville et al. (2009), most likely by the use of different genotypes and age of the fixed material that stored up to months. In our experiment, we used relatively fresh material stored in ethanol 70% for no more than two months. Despite the protocol adjustments, we still keep some cytoplasmic background and more dense in "Pisang Klutuk" (*M. balbisiana*, BB) and "Pisang Kepok" (ABB) (data unpublished), which requires further adaptation of cell spreading conditions.

The first meiotic stage that we analysed was late pachytene, in which the total chromosome complements measured 207–339 μm, about ten times longer than their mitotic counterparts (Osuji et al., 2006) and with a contraction in the range of most plant species (de Jong et al., 1999). The long pachytene chromosomes are excellent for studying the well-defined euchromatin/heterochromatin differentiation and pairing of structural chromosome rearrangements. The blue fluorescing DNA-specific DAPI is more appropriate to reveal fine details of the chromosomes (De Capdeville et al., 2009; own study) than the protein-based staining methods (Adeleke et al., 2002). *Musa* chromosomes show a characteristic heterochromatin pattern of numerous chromomeres, small heterochromatin knobs distributed in the distal and interstitial segments of the chromosomes, and larger and much brighter fluorescing segments in the pericentromere heterochromatin. Based on previous published genomic information

(D'Hont et al., 2012; Martin et al., 2016), we suggest that such chrommeres are islands of Copia retrotransposon clusters and that pericentromeres are sites enriched of Gypsy and other unclassified (satellite) repeats, regions known to be notoriously tricky in gap closure of the pseudomolecules genome assembly.

Pachytene complements are the primary sources for elucidating disturbances in pairing between homologous and homoeologous chromosomes. The late part of this stage (late pachytene) is most suitable for such analysis as chromosomes become detached from their densely clustered synizesis knot (Moens, 1964). The first class of pairing abnormalities seen in our materials are the asynaptic regions in the ssp. *malaccensis*, var. *breviformis* and "Rejang" and were also mentioned in the study of pachytene spreads in a wild *M. acuminana* "Calcutta 4" by Adeleke et al. (2002). The most obvious explanation for such structural aberrations is heterozygosity for a paracentric (or less common pericentric) inversion. Other reasons include large deletions, or insertion duplications are very unlikely, as they are (sub)lethal or infrequent (Dobzhansky & Sturtevant, 1938; Maguire & Riess, 1994). If the assumed inversion is large enough, a loop can be formed, and if a crossover in the loop takes place, a dicentric chromosome pair and acentric fragment will be created, as can be seen in metaphase I— anaphase I/II cells of ssp. malaccensis (Figure 5B), although the expected acentric fragment as previously seen in the triploid M. acuminata cv. "Hom" [20] was not detectable. Wilson (1946b, a; c) showed comparable anaphase bridges in camera lucida drawings of Musa pollen mother cells, but detailed interpretations were not given. However, Dodds (1943) gave more extensive descriptions of such bridges and fragments in a diploid *M. acuminata* clone and concluded that such meiotic disturbances resulted from crossovers in heterozygous inversion segments. Shepherd (1999) described in his monograph of *Musa* cytogenetics a large number of cases of inversion type bridges observed in many species and hybrids, making inversions common chromo-some rearrangements in this genus, in both intraspecific and interspecific hybrids. Unequivocal criteria for identifying inversions were the presence of chromatid bridges associated with a specific fragment at anaphase I, which he interpreted as the result of crossing over between relatively inverted segments. Less clear is a proper explanation of the bridges when acentric fragments were minute or even missing, assuming that inversions were very small and nearly at the chromosome's distal end. As we never observed more than a single anaphase bridge in the ssp. *malaccensis* and "Rejang" assume that only a single inversion exists in these heterozygotes and thus will not have a severe negative effect on gamete fertility.

Translocations represent the second group of major structural chromosome rearrangements. The most common type is the interchange of a reciprocal translocation

with two translocation chromosomes and their corresponding non- homologous chromosomes, forming a translocation complex at pachytene. Only in ssp. *malaccensis*, we observed an asynaptic region that may hint at a translocation breakpoint. A more accurate description of translocation complexes can be obtained by electron microscopic analysis of spread synaptonemal complex preparations [11]. Still, the low number of suitable pachytene meiocytes in an anther and the expected problems with hypotonic bursting of the dense cytoplasm made us refrain from using this technology. In the recent studies that used the sequencing database of *M. acuminata* ssp. *malaccensis* reverence chromosomes (Martin et al., 2020) and chromosome painting (Šimoníková et al., 2020), a translocation event between chromosomes 1 and 4 was present in ssp. *malaccensis*. However, convincing proof for translocation heterozygosity in our material was expected from metaphase I complements. In the ssp. *malaccensis* and var. *zebrina*, and "Rejang", we observed overlapping bivalents that may indicate alternate (zigzag oriented) quadrivalents. However, the occurrence of adjacent quadrivalents that form conspicuous configurations that lead to unbalanced spores were never found; neither did we find unbalanced chromosome numbers at anaphase I / II or many sterile pollen grains. A recent genetic study of ssp. *malaccensis* (Ahmad et al., 2020b) showed homozygosity for a $1^4/4^1$ translocation, which confirmed the lack of quadrivalents at diakinesis/metaphase I cells.

An interesting observation of "Rejang" is the occurrence of many univalents that point at partial synapsis or failure of chiasma formation. Indeed, examples of incomplete pairing were seen in the pachytene complements that at least partly could account for the observed univalents. The larger number of univalents is expected to lead to unbalanced gametes, and so to a wider variation in pollen size in the lactophenol acid fuchsin stained pollen preparations. The observed univalents suggest that this banana may be segmental polyploid. These univalents indicate a lack of homology, as observed in progenies of synthetic allodiploids of Brassica rapa and Raphanus sativus (Park et al., 2020). Further study of "Rejang" is needed to see if homoeologous differences of the parental genomes can explain the explained lack of pairing/recombination.

Our study found several examples of mitotic and meiotic restitutions leading to doubling of the chromosome numbers. Mitotic chromosome doubling was observed only once in a slide with 44 chromosomes at diakinesis in a diploid 'AA' cultivated banana, "Rejang". Such a doubling of the chromosomes likely resulted from a premeiotic restitutional mitosis as previously reported in the triploid banana "Cavendish" and "Gros Michel" (Raboin et al., 2005). A second phenomenon is meiotic restitution in banana, which give rise to monads, dyads or triads at the tetrad stage of diploid banana, caused by first divisional restitution (FDR) or second division restitution (SDR) or a combination

of both (Dodds, 1943). In the case of a triad, we assume the failure of anaphase II (SDR) in only one of the dyads. So, one daughter cell is unreduced while the other two smaller cells are haploid. In "Rejang" cells, chromosome doubling might be one of the explanations for the larger pollen size compared to other diploid bananas (Figure 6). In polyploid plants, nuclear restitution is most likely due to unbalanced segregation between a group chromosome with two complement chromosomes and a group of single complement or the complete lack of anaphase I segregation (First division segregation, FDR). Moreover, meiotic restitution in diploids and polyploids can also occur by second division restitution, SDR, mostly by deficient spindle orientation at anaphase II (Bretagnolle & Thompson, 1995; Ramanna & Jacobsen, 2003). Considering these events, it is plausible to explain how triploid bananas might have resulted from hybridisation or crosses between diploid bananas (Rekha & Hiremath, 2008).

Author's contributions

FA and HdJ designed this study and carried out the experiments. YSP helped with the selection of banana materials for this study. FA, GK and HdJ wrote the manuscript and GK and HdJ supervised the project

References

Adeleke MT V, Pillay M & Okoli BE (2002) An improved method for examining meiotic chromosomes in *Musa* L. HortScience. pp 959–961.

Ahmad F, Martawi NM, Poerba YS, de Jong H, Kema GHJ (2020a) Genetic mapping of Fusarium wilt resistance in a wild banana Musa acuminata ssp. malaccensis accession. Theor Appl Genet 133:3409–18.

Ahmad F, Bourke P, Schouten H, Vokaert HA & Kema GHJ (2020b) Cytogenetics of structural rearrangements in *Musa* hybrids and cultivars. Achieving sustainable cultivation of bananas Volume 2: Germplasm and genetic improvement. (ed by GHJ Kema & A Drenth) Burleigh Dodds Science Publishing, Cambridge, UK, pp 1–28.

Bakry F, Carreel F, Jenny C & Horry J-P (2009a) Genetic Improvement of Banana. In: Breeding Plantation Tree Crops: Tropical Species. (ed by SM Jain & PM Priyadarshan) 2nd edn. Springer, New York, NY, pp 3–50.

Bakry F & Shepherd K (2008) Chromosome count on banana root tip squashes. Fruits 63:179–181.

Blomme G, Eden-Green S, Mustaffa M, Nwauzoma B & Thangavelu R (2011) Major Diseases of Banana. In: Banana Breeding. Progress and Challenge. (ed by M Pillay & A Tenkuano) CRC Press, pp 85–120.

Bretagnolle F & Thompson JD (1995) Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants. New Phytologist 129:1–22.

De Capdeville G, Souza Júnior MT, Szinay D, Diniz LEC, Wijnker E, Swennen R, Kema GHJ & De Jong H (2009) The potential of high-resolution BAC-FISH in banana breeding. Euphytica 166:431–443.

D'Hont A, Denoeud F, Aury J, Baurens F, Carreel F, Garsmeur O, Noel B, Bocs S, Droc G, Rouard M, C DS, Jabbari K, Cardi C, Poulain J, Souquet M, Labadie K, Jourda C, Lengellé J, Rodier-Goud M, Alberti A, Bernard M, Correa M, Ayyampalayam S, Mckain M, Leebens-Mack J, Burgess D, Freeling M, Mbéguié-A-Mbéguié D, Chabannes M, Wicker T, Panaud O, Barbosa J, Hribova E, Heslop-Harrison P, Habas R, Rivallan R, Francois P, Poiron C, Kilian A, Burthia D, Jenny C, Bakry F, Brown S, Guignon V, Kema G, Dita M, Waalwijk C, Joseph S, Dievart A, Jaillon O, Leclercq J, Argout X, Lyons E, Almeida A, Jeridi M, Dolezel J, Roux N, Risterucci A, Weissenbach J, Ruiz M, Glaszmann J, Quétier F, Yahiaoui N & Wincker P (2012) The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. Nature 488:213–219.

Dale J, James A, Paul J-Y, Khanna H, Smith M, Peraza-Echeverria S, Garcia-Bastidas F, Kema G, Waterhouse

P, Mengersen K & Harding R (2017) Transgenic Cavendish bananas with resistance to Fusarium wilt tropical race 4. Nature Communications DOI: 10.10.

Dobzhansky T & Sturtevant AH (1938) Inversions i n the chromosomes of drosophila pseudoobscura. Genetics 23:28–64.

Dodds KS (1943) Geentical and cytological studies of *Musa*, V. Certain edible diploids. Genetics 45:113–139.

FAO (2020) Banana Market Review 2019. FAO, Rome.

Fortescue JA & Turner DW (2004) Pollen fertility in *Musa*: Viability in cultivars grown in Southern Australia. Australian Journal of Agricultural Research 55:1085–1091.

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

Heslop-Harrison JS & Schwarzacher T (2007) Domestication, Genomics and the Future for Banana. Annals of Botany 100:1073–1084.

Iovene M, Wielgus SM, Simon PW, Buell CR & Jiang J (2008) Chromatin Structure and Physical Mapping of Chromosome 6 of Potato and Comparative Analyses With Tomato. Genetics 180:1307–1317.

de Jong H, Fransz P & Zabel P (1999) High resolution FISH in plants - techniques and applications. Trends in Plant Science 4:258–263.

Kantama L, Wijnker E & Jong H De (2017) Optimization of Cell Spreading and Image Quality for the Study of Chromosomes in Plant Tissues. Plant Germline Development: Methods and Protocols. (ed by A Schmidt) Springer Science, pp 141–158.

Kayat F, Bonar N, Waugh R, Rajinder S, Rahimah AR, Rashid AR & Othman RY (2009) Development of a genetic linkage map for genes associated with resistance and susceptibility to *Fusarium oxysporum* f. sp. *cubense* from an F1 hybrid population of *Musa acuminata* ssp. *malaccensis*. Proceedings of the International ISHS, Acta Horticulturae 828:1–8.

Khayat E (2020) Targeted improvement of Cavendish clones. Achieving sustainable cultivation of bananas Volume 2: Germplasm and genetic improvement. (ed by G Kema & A Drenth) Burleigh Dodds Science Publishing, Cambridge, UK,.

Levitt D & King M (1987) Methanol fixation permits flow cytometric analysis of immunofluorescent stained intracellular antigens. Journal of Immunological Methods 96:233–237.

Maguire MP & Riess RW (1994) The Relationship of Homologous Synapsis and Crossing Over in a Maize Inversion. Genetics Society of America 137:281–288.

Martin G, Baurens F, Droc G, Rouard M, Cenci A, Kilian A, Hastie A, Dolezel J, Aury J, Alberti A, Careel F & D'Hont A (2016) Improvement of banana *Musa acuminata* reference sequence using NGS data and semi-automated bioinformatics methods. BMC genomics 17:243.

3

Martin G, Baurens F-C, Hervouet C, Salmon F, Delos J-M, Labadie K, Perdereau A, Mournet P, Blois L, Dupouy M, Carreel F, Ricci S, Lemainque A, Yahiaoui N & D'Hont A (2020) Chromosome reciprocal translocations have accompanied subspecies evolution in bananas. The Plant Journal 104:1698–1711.

Maryani N, Lombard L, Poerba YS, Subandiyah S, Crous PW & Kema GHJ (2019) Phylogeny and genetic diversity of the banana Fusarium wilt pathogen *Fusarium oxysporum* f. sp. *cubense* in the Indonesian centre of origin. Studies in Mycology 92:155–194.

Moens P (1964) A new interpretation of meiotic prophase in *Lycopersicon esculentum* (tomato). Chromosoma 15:231–242.

Morán AJF (2013) Improvement of Cavendish Banana Cultivars through Conventional Breeding (I Van den Bergh, Ed. by). Acta Horticulturae 986:205–208.

Ordóñez N, Seidl MF, Waalwijk C, Drenth A, Kilian A, Thomma BPHJ, Ploetz RC & Kema GHJ (2015) Worse Comes to Worst: Bananas and Panama Disease—When Plant and Pathogen Clones Meet Bananas: Their Origin and Global Rollout. PLoS Pathogens 11.

Ortiz R & Vuylsteke D (1995) Inheritance of dwarfism in plantain (*Musa* spp., AAB group). Plant Breeding 114:466–468.

Osuji J, Okoli BE & Edeoga HO (2006) Karyotypes of the A and B Genomes of *Musa* L. Cytologia 71:21–24.

Park HR, Park JE, Kim JH, Shin H, Yu SH, Son S, Yi G, Lee S, Kim HH, Huh JH & Nelson MN (2020) Meiotic Chromosome Stability and Suppression of Crossover Between Non-homologous Chromosomes in x *Brassicoraphanus* , an Intergeneric Allotetraploid Derived From a Cross Between *Brassica rapa* and *Raphanus sativus*. 11:1–12.

Peraza-Echeverria S, Dale JL, Harding RM, Smith MK & Collet C (2008) Characterization of disease resistance gene candidates of the nucleotide binding site (NBS) type from banana and correlation of a transcriptional polymorphism with resistance to *Fusarium oxysporum* f.sp. *cubense* race 4. Molecular Breeding 22:565–579.

Ploetz RC (2005) Panama disease: an old nemesis rears its ugly head, Part 1. The begining of the banana export trades. Plant Health Progress. https:// doi. org/ 10. 1094/ PHP-2005- 1221- 01- RV.

Raboin LM, Carreel F, Noyer JL, Baurens FC, Horry JP, Bakry F, Montcel HT Du, Ganry J, Lanaud C & Lagoda PJL (2005) Diploid ancestors of triploid export banana cultivars: Molecular identification of 2n restitution gamete Donors and n gamete donors. Molecular Breeding 16:333–341.

Ramanna MS & Jacobsen E (2003) Relevance of sexual polyploidization for crop improvement – A review. Eu-

phytica 133:3–18.

Rekha A & Hiremath SC (2008) Chromosome studies and karyotype analysis of some triploid banana (*Musa* species) cultivars of AAA genomic group. J. Hort. Sci 3:30–34.

Roux NS, Toloza A, Dolezel J & Panis B (2004) Usefulness of embryogenic cell suspension cultures for the induction and selection of mutants in *Musa* spp. Banana improvement: Cellular, Molecular Biology and Induced Mutations. (ed by S Jain & R Swennen) Science Publisher, New Hempshire, USA, pp 33–43.

Sass J (1964) Botanical microtechnique. Iowa State Univ. Press, Ames, Iowa.

Shepherd K (1999) Cytogenetics of the genus *Musa*. International Network for the Improvement of Banana and Plantain, Montpellier, France.

Simmonds NW & Shepherd K (1955) The taxonomy and origins of the cultivated bananas. Journal of the Linnean Society of London, Botany 55:302–312.

Šimoníková D, Nemecková A, Cížková J, Brown A, Swennen R, Doležel J & Hribová E (2020) Chromosome Painting in Cultivated Bananas and Their Wild Relatives (*Musa* spp .) Reveals Differences in Chromosome Structure. Int J Mol Sci 21:7915.

Wilson GB (1946a) Cytological studies in the Musae. I. Meiosis in some triploid clones. Genetics 31:241–258.

Wilson GB (1946b) Cytological studies in the Musae. II. Meiosis in some diploid clones. Genetics 31:475–482.

Wilson GB (1946c) Cytological studies in the Musae. III. Meiosis i n some seedling clones. Genetics 31:483–493.

3

Chapter 4.

Cytogenetics of structural rearrangements in *Musa* hybrids and cultivars

Ahmad F, Bourke P, Schouten H, Volkaert HA, Kema GHJ & de Jong H

Introduction

Edible bananas and plantains constitute a group of crops whose fruits show starchy, sweet and aromatic flesh with characteristic flavour and structures. Their seedless tasty pulp makes them popular sources of daily nutrition for people all over the world. The absence of seeds, however, comes at a price as they are sterile and hence difficult for cross-breeding. Almost all cultivated bananas have seedless fruits, either as parthenocarpic diploids or sterile triploids of *Musa acuminata*, or interspecific hybrids of *M. acuminata* (AA) and *M. balbisiana* (BB), the latter known as AA, AB, BB, AAA, AAB, ABB, AAAB and AABB genome group bananas. Gamete sterility results from the failure of balanced chromosome segregation in meiosis due to lack of pairing between the homoeologues, structural hybridity or triploidy or a combination of these (Bakry et al. 1990). Such mismatched chromosomes can lead to bridges, laggards or uneven separation of chromosomes, giving rise to infertile ovules and pollen. An additional factor that plays a role in seedless bananas is parthenocarpy, that is, the induced production of the fruit without fertilization of the ovule. This trait was likely preferred by farmers during early cultivation of the crop. The sterility problems in most of the cultivars and interspecific hybrids of banana have necessitated breeders and geneticists to analyse their material for pollen fertility, chromosome numbers and aberrant meiotic behaviour including chromosome pairing switches, lack of crossover recombination and unbalanced chromosome segregation. The classical school of cytogenetic and taxonomic analyses of *Musa* species have produced an impressive series of research articles by Dodds (1943), Dodds and Simmonds (1946), Wilson (1946), Simmonds and Shepherd (1955) and others, and the valuable monograph of Shepherd (1999). Their scrutinous work on mitotic chromosome counts and meiotic configurations revealed a wealth of information on taxonomic relationships, ploidy levels and fertility of Musa species, hybrids and cultivars. Most of these studies were based on acetocarmine or orcein chromosome squash protocols and camera lucida drawings of selected cell complements, but some of them were able to produce high-quality photomicrographs (Agarwal 1983, 1988). For increased contrast and details of the chromosomes, Pillay and Adeleke (2001) and Adeleke et al. (2002) introduced silver nitrate incubation of air-dried chromosome preparations, a method known for its specific detection of active nucleolar organizer regions (Lacadena et al. 1984) and detailed staining of synaptonemal complexes (Albini 1994). Moreover, Bakry and Shepherd (2008) published a protocol with detailed information on required laboratory materials explaining all steps in making preparation and microscopy with useful hints for troubleshooting. Osuji et al. (2006) compared the karyotypes of the AA ('Calcutta 4', *M. acuminata* Colla) and BB ('Butohan 2', *M. balbisiana* Colla) mitotic chromosome complements concluding that these 1.4–3.6 μm small chromosomes of the parental species showed great

similarity in structure and size. Microscopic interpretation of meiotic chromosome configurations is limited by the rather small size of banana chromosomes, making it difficult to interpret chiasma positions in chromosome associations visible at diakinesis – metaphase I. In addition, pollen mother cells display dense cytoplasm with granular structures that mask minute acentric fragments accompanying anaphase bridges of inversion heterozygotes. Translocation complexes are clearly recognizable in adjacent quadrivalent and trivalent + univalent configurations, but interpretation of the more compact quadrivalent configurations with alternate-oriented centromeres can be challenging. The importance of such multivalent observations is illustrated in the analyses of banana hybrids by Shepherd and others (Wilson 1946; Simmonds and Dodds 1949; Shepherd 1999). Shepherd's study on F1 hybrids obtained from crosses with diploid wild accessions of *M. acuminata* demonstrated multivalents including 4–8 chromosomes in diakinesis – metaphase I complements. In general, such hybrids display different levels of male and female fertility that are related to the number of structural change complexes, known as structural hybridity (Darlington 1929). Fertility may be from moderate to high when only one exchange is involved, depending on the ratio of adjacent (leading to unbalanced gametes) and alternate orientations (leading to balanced gametes), but fertility is progressively reduced, if the plant contains two or three translocation complexes (Shepherd 1999). Some hybrids and partial sterile diploids are known to generate small numbers of larger size pollen grains, which likely result from meiotic restitutions. These so-called 2n gametes have been described extensively (reviewed in Ortiz (1995)) and may have derived from first division restitution (FDR) or second division restitution (SDR). The sterile trisomic polyploids 'Cavendish' and 'Gros Michel' are supposed to have arisen from 2n restitution gametes and a normal haploid gamete donor (Raboin et al. 2005). Offspring from triploid × diploid crosses may give rise to predominantly tetraploid progeny produced from 2n eggs resulted from SDR in the triploid mother (Ortiz 1995). FDR 2n eggs are also reported but they are likely even rarer in Musa.

Pollen fertility

Pollen fertility tests are often used instead of or in support of meiotic studies of pollen mother cells. They are fast and straightforward and provide a direct indication about sterility and aberrations, if any, during male meiosis. Various staining protocols are used for pollen fertility studies, including Alexander staining (Alexander 1969), lactophenol acid fuchsin (Sass 1964) and 2,3,5-Triphenyltetrazolium chloride (TTC) (Damaiyani and Hapsari 2017). Unbalanced chromosome segregation, especially in few female-fertile triploid cultivars (AAA, AAB and ABB) will in general give very few stainable,

fertile pollen. Translocation heterozygotes, which occur regularly in banana, contain two translocation (e.g. 1T2 and 2T1) and two non-translocation chromosomes (1 and 2). Segregation of the chromosomes of the translocation complex will be balanced (1T2 + 2T1/1 + 2) giving rise to fertile gametes, or will be unbalanced (1T2 +1 / 2T1 + 2 or 1T2 +2 / 2T1 + 1) giving rise to sterile gametes. The second class of structural variants include heterozygoty for a paracentric inversion, which show variable levels of sterility depending on the frequency of chiasmata in the inversion loop. Unreduced and polyploid spores indicate restitutional divisions, either at anaphase I or II, or both. Damaiyani and Hapsari (2017) conducted a 1% solution of the TTC staining for testing pollen grain viability of wild *M. acuminata* ssp. *rutilifes* and *M. balbisiana* and cultivars. Their study demonstrates considerable pollen viability among the tested banana accessions making it a useful tool for banana taxonomy.

Flow-cytometry

Flow-cytometry is a high-throughput alternative in determining ploidy levels and aneuploidy in banana species and hybrids. Lysak et al. (1999) used this technology for nuclear genome size variation in *M. acuminata*, *M. balbisiana* and various triploid clones with different genomic constitutions. The *M. balbisiana* genome was estimated at 537 Mbp, with no genome size variation and about 12% smaller than the *M. acuminata* accessions that showed genome sizes ranging from 591 Mbp to 615 Mbp. The even larger variation observed in the group of triploid Musa accessions (ranging from 559 Mbp to 613 Mbp) reflects both genomic constitutions as well as differences in the size between the A and the B genomes. A further refinement of the technology was published by Roux et al. (2003), who claim a rapid and accurate flow-cytometric detection of 2n=31–32 aneuploids obtained from gamma-irradiated triploid shoot tips of the *M. acuminata* clone 'Grande Naine' (2n=3x=33). Oselebe et al. (2009) determined ploidy variation in progenies of 2x–2x, 2x–4x and 4x–2x hybrids. Ploidy was established using flow-cytometry analysis of nuclear DNA content and chloroplast characteristics, showing that progenies of 2x–2x crosses were predominantly diploid (99.7%), whereas those of 2x–4x crosses were mainly diploid (96.2%), and the 4x–2x crosses produced predominantly triploid progenies (94.1%). A comparable flow-cytometric analysis together with chromosome counts was carried out on tissue culture-induced off-type bananas by Msogoya et al. (2008) who found substantial loss of chromosomes or chromosome segments. The few examples of flow cytometry in banana show that the technology is fast and convenient, and with modern devices might be accurate enough to show aneuploidy even in triploid banana populations, segregating families or somaclonal samples, and hence will replace the much more time-consuming

chromosome counting protocols. However, for the flow-cytometric estimates of unknown *M. acuminata* accessions or a segregating population, one should keep in mind that genome size differences can be up to 12% (Lysak et al. 1999) and so may unjustly suggest the occurrence of aneuploid individuals. Thus, chromosome counting of those plants still remains indispensable.

Fluorescent in situ hybridization (FISH)

Fluorescent in situ hybridization (FISH) is a molecular cytogenetic technique using single-strand fluorescently labelled DNA probes that bind to denatured chromosomal target DNA with a high degree of sequence complementarity. It was developed in the early 1980s mostly for human cytogenetics and found its way to plant research as an outstanding and versatile tool to map repeats and single-copy sequences on chromosomes and nuclei, bridging the disciplines of genetics and genomics. Applications of FISH to banana cytogenetics have elevated the research to a higher level, although some of its potentials are yet to be exploited. The small size of the chromosomes and the dense rigid cytoplasm in spread pollen mother cells are major challenges in successful hybridizations of probe DNAs to the chromosomal targets. The composition of parental chromosomes in Musa hybrids has been disclosed by a special type of FISH technology, known as genomic in situ hybridization (GISH) or genome painting, using total genomic DNAs of the parental species as probes in a FISH on cell spread slides of interspecific or intergeneric hybrids. In a double-labelling regime, Osuji et al. (1997) identified the A genome and B genome chromosomes in several amphiploid (AAB and ABB) cultivars, with most discriminating signals coming from satellite repeats and the pericentromeric and nucleolar organizer repeats in the heterochromatin regions. In a comparable GISH study of hybrids involving three wild species of section Eumusa (*M. acuminata*, *M. balbisiana*, *M. schizocarpa*) and one wild species of section Australimusa (*M. angustigemma*), D'Hont et al. (2000) distinguished the parental chromosomes in their hybrids, although their discriminating signals were confined to the heterochromatin regions. The strength of the fluorescent hybridization signals corresponds with the assumed phylogenetic relationship between these species. In a later study, D'Hont (2005) included the genome structure of more interspecific banana cultivars revealing that the chromosome constitutions in these hybrids in most cases correspond well with their presumed genome constituents. This method has been applied to analyse chromosome pairing in metaphase from triploid interspecific cultivars, and it clearly demonstrates that interspecific recombinations between *M. acuminata* and *M. balbisiana* chromosomes do occur and may be frequent in triploid hybrids. These results shed new light on Musa cultivar evolution and provide

important implications for banana and plantain breeding. The potential of genome painting has further been explored in a study on homoeologous recombination in spread pollen mother cells at metaphase I. Bivalents and trivalents with A and B genome chromosomes could be distinguished by FISH fluorescence intensities, and so showed that A- and B-genome chromosomes display homoeologous recombination (Jeridi et al. 2011). In a follow-up analysis, Jeridi et al. (2012) studied the nature of chromosome configurations in a sterile AB Indian cultivar and in its fertile colchicine-induced allotetraploid (AABB). Genome painting demonstrated that the genomes of the AABB reveal the so-called segmental allopolyploidy (Sybenga 1975) with three chromosome sets showing a tetraploid-like pattern (i.e. showing quadrivalent pairing), whereas the remaining sets are diploid-like (bivalents only) or display undetermined patterns. Balanced and unbalanced diploid gametes were detected in progenies, with the chromosome constitution appearing to be more homogenous in pollen than in ovules. This segmental inheritance pattern exhibited by the AABB allotetraploid genotype confirms earlier findings on chromosome rearrangements between *M. acuminata* and *M. balbisiana*, which may provide the possibility of introgressing valuable alleles in banana breeding programmes. A recent genome sequence analysis of *M. balbisiana* (Wang et al. 2019) enabled the comparison with the *M. acuminata* reference genome showing details of their synteny. Both genomes display a surprising collinearity and sequence similarity. Only two translocations and two inversions between the A- and B-genomes are observed, that were also supported by the re-arrangements based on genetic mapping data (Wang et al. 2019). The large-scale collinearity between the A- and B- genomes does explain very well why the A- and B-homoeologues can undergo homologous exchange and recombination in diploid and amphiploid hybrids.

Use of repetitive sequences in banana research

The first class of repetitive sequences that were localized on chromosomes by FISH were the ribosomal DNA genes using conserved parts of their sequences as probes. Doleželová et al. (1998) performed such a hybridization using parts of the 5.8S – 18S – 25S rDNAs of *Vicia faba* and a 5S rDNA on cell spreads of two banana cultivars and six wild relatives of Musa, all of them diploid with 2n=22 chromosomes. Next to only one locus for the 18S/25S rDNA tandem array on the NOR chromosome pair, a variable number of two, three or five loci of the 5S rDNA tandem array were detected, the latter possible resulting from heterozygosity of a structural re-arrangement (Doleželová et al. 1998). Extension of these analyses on 19 diploid accessions representing the four taxonomic sections Eumusa, Rhodochlamys, Callimusa and Australimusa, and the outgroup *Ensete gilletii* was performed by Bartoš et al. (2005) with 2–8 fluorescent

spots for the 45S rDNA and 4–8 spots for the 5S rDNA. Hierarchical cluster analysis using genome size, chromosome number and 45S rDNA sites revealed a clear relationship between the accessions under study. The second group of characteristic repetitive sequences in plant genomes are the retrotransposable elements, of which the gypsy type with its species- specific long terminal repeats that occupies most of the heterochromatin of the chromosomes is the most abundant. In *M. acuminata* cv. 'Grand Nain', Balint-Kurti et al. (2000) described the gypsy-like monkey element, of which the reverse transcriptase, RNase H and integrase genes resemble retroelements from plants, fungi and yeast. Southern analysis of genomic DNAs from nine different banana cultivars, digested with *HindIII* and probed with the different fragments of the major open reading frames demonstrated hybridizations with the A genome or with both A and B genomes. FISH demonstrates that this non- autonomous element co-localizes with the 18S-5.8S- 25S rRNA genes in the nucleolar organizer region and also occurs in small regions distributed over all chromosomes. In a more comprehensive study, Valárik et al. (2002) studied the chromosomal localization of various repetitive elements on cell complements of *M. acuminata* and *M. balbisiana*, and analysed sequence similarity to known DNA sequences. The characterized clones included 26S and 5S rDNA genes, and sequences that displayed fluorescent foci in centromeric regions. All repetitive sequences were more abundant in *M. acuminata* than in *M. balbisiana*, corresponding to the 12% larger genome size of *M. acuminata* compared to *M. balbisiana* suggesting that repetitive sequences to a greater part account for the difference in genome size between both species. In a follow-up on previous repeat research, Hřibová et al. (2007) performed DNA re-association kinetics to isolate the highly repeated fraction of the banana genome (*M. acuminata* 'Calcutta 4') and constructed two Cot libraries (Peterson et al. 2002) of different repeat complexity, and a third set of DNA clones with random sequences for comparison. 'BLAST' homology searches demonstrated that the majority of the repeats represent different types of retrotransposons, of which Ty3/gypsy-type monkey retrotransposon was the most prevalent. Tandemly organized sequences are mapped to the nucleolar organizers and proximal pericentromere regions of the *M. acuminata* chromosomes. In a low-depth 454 sequencing study of the repeat part of banana (*M. acuminata* 'Calcutta 4') Hřibová et al. (2010) produce partial sequence reconstruction and characterization of repetitive DNA. Most of the repeats are identified as Ty1/copia and Ty3/gypsy retroelements, occupying 16% and 7% of the genome, respectively. In addition, some LINEs, satellite repeats including rDNAs and unclassified retroelements are observed, while DNA transposons are relatively rare. The small genomic part of satellite repeats has further been explored by Čížková et al. (2013), who compared two main DNA satellites in a set of 19 Musa accessions, including representatives of A, B and S (M. schizocarpa) genomes and their interspecific hybrids. The two DNA satellites showed a high level of sequence similarity within, and a high

homology between Musa species. Microscopic observations of these satellites in a FISH along with rRNA genes, LINE-like element and a single-copy BAC clone revealed characteristic fluorescence patterns in *M. acuminata* and *M. balbisiana*, in which up to five chromosomes could be identified. Next-generation sequencing and comparative analysis of the major repetitive sequences of representative species of the Musaceae family (*M. acuminata*, *M. ornata*, *M. textilis*, *M. beccarii*, *M. balbisiana*, and *Ensete gilletii*) demonstrated their overall intraspeific and interspecific similarities using graph-based clustering (Novák et al. 2014). The most abundant classes of highly repetitive DNA included the Maximus/SIRE and Angela lineages of Ty1/copia LTR retrotransposons and the chromovirus lineage of Ty3/gypsy elements. The differentiation of the repeats followed the taxonomic distances of the species in the Musaceae family, while the closely related species *M. acuminata* versus *M. ornata*, and *M. beccarii* versus *M. textilis* share more related populations of repetitive elements. This study did not reveal any repeat class that can help as a chromosome marker in karyotype analyses. It is still unknown if simple sequence repeats/microsatellites that are currently used in marker technology of Musa (Creste et al. 2003; Christelová et al. 2011) occur as large tandem arrays in the banana genome such that they could be used as microsatellite FISH (Bouilly et al. 2008; Cuadrado and Jouve 2010) for chromosome identification.

Chromosomal detection of single-copy sequences

In the next step of characterization of the genome and chromosomes of Musa, researchers have developed bacterial artificial chromosome (BAC) libraries. Vilarinhos et al. (2003) published the first publicly available BAC library obtained from genomic DNA of the wild diploid *M. acuminata* 'Calcutta 4' clone. The library consisted of 55 152 clones with an average insert size of 100 kb and has a coverage of 9× the haploid genome. The BAC clones were anchored on a still incomplete genetic map with eight linkage groups using a set of 13 RFLP probes. The BAC inserts were created for use as chromosome markers in FISH, but could be helpful as seed BACs during the assembly of the physical map of banana. Sequence information is now available from the BAC-end sequencing project of Cheung and Town (2007), and Arango et al. (2011), who generated 6252 and 46 080 reads, respectively. The obtained BAC-end sequences were screened against several databases and significant homology was found compared to mitochondria and chloroplasts (2.6%), transposons and repetitive sequences (36%) and proteins (11%). In addition, the sequencing revealed 352 potential simple sequence repeat (SSR) markers of which the most were AT-rich. After filtering mitochondria and chloroplast matches, thousands of BAC end-sequences had a significant BLASTN match to the *Oryza sativa* and Arabidopsis genome sequences. The syntenic relationships

between *Musa* and rice have further been elaborated by Lescot et al. (2008). Soon after Vilarinhos' study of the BAC library of *M. acuminata*, Šafář et al. (2004) released the BAC library of the second ancestral species of banana hybrid cultivars, *M. balbisiana* 'Pisang Klutuk Wulung'. Here the authors introduced an improved DNA isolation protocol and a novel flow cytometric-based nuclei isolation method as a strategy to circumvent interference of secondary metabolites (polyphenols and polysaccharides) and plastid DNA in leaf tissues for the isolation of high-molecular-weight nuclear DNA. In addition, the application of the inducible pCC1BAC vector improved the amount of BAC DNA. The two sub-libraries, for each of the two methods, were screened with twelve RFLP probes and were anchored to eight linkage groups of *M. acuminata*. Ten BAC clones were selected for FISH mapping on mitotic cell complements of *M. balbisiana*. Three BAC clones showed multiple fluorescent foci on the chromosomes, while a second group of three BAC clones, that were supposed to be rich in gypsy LTR elements, produced a uniform distribution mostly in the (peri)centromeric regions, and a third group likely rich in copia and/or gypsy elements produced almost complete labelling of the chromosome complement. The major drawback of FISH in Musa species and hybrids is the small size of the mitotic chromosomes and hence the low resolution of adjacent or partly overlapping probe signal in multiple colour FISH essays. In de Jong et al. (1999), a comparison has been made between mitotic metaphase chromosomes, pachytene bivalents, interphase nuclei and extended DNA fibres in plant species. The choice of pachytene chromosome complements for BAC FISH is undeniably one of the best choices in species with small genomes including *Musa*. In Capdeville et al. (2008) the first attempts of such a BAC clone positioning on pachytene chromosomes of 'Calcutta 4' and the wild *M. velutina* is demonstrated. The preparation of spread pollen mother cells after digestion with pectolytic enzymes and maceration with acetic acid poses considerable technical challenges for reducing the thick cytoplasm with the polyphenolics and polysaccharides that hamper probe DNA to hybridize to the chromosomal targets. In spite of these hurdles, BAC FISH is powerful enough to support meiotic studies on pairing disturbances in hybrid cultivars with chromosomal rearrangements. In a more recent approach, single copy chromosomal markers can now be obtained directly from the highly improved genome map of *M. acuminata* (Martin et al. 2016) where specific unique sequences can be selected for making synthetic probes, as used recently in developing oligo painting FISH strategies (Han et al. 2015; Li et al. 2016; Braz et al. 2018; Hou et al. 2018).

Molecular markers in linkage studies

Geneticists and breeders have employed various DNA markers for the characterization and evaluation of genetic diversity in Musa species and map-based cloning in banana breeding programmes (Fauré et al. 1993; Pillay and Tripathi 2006). For a long period, the use of SSRs has been favoured (Jarret et al. 1994; Kaemmer et al. 1997; Creste et al. 2003; Buhariwalla et al. 2005). Furthermore, sequenced tagged microsatellite site (STMS) marker technology (Kaemmer et al. 1997), inter-simple sequence repeats or ISSR (Godwin et al. 1997; Padmesh et al. 2012; Silva et al. 2017) and inter-retrotransposon amplified polymorphic (IRAP) markers (Nair et al. 2005) have also been used. Arora et al. (2018) developed BanSatDB, a whole-genome-based database of putative and experimentally validated microsatellite markers of three Musa species. Hippolyte et al. (2010) applied the diversity array technology (DArT) in combination with SSRs for a linkage map of diploid *M. acuminata*, while Sardos et al. (2016) use DArT strategy for whole genome profiling to obtain deeper understanding of the evolution, taxonomy and domestication of edible banana. The first linkage map of *M. acuminata* was obtained on the basis of a segregating F2 population from a diploid AA clone (SF265) and a diploid AA Banksii (*M. acuminata* ssp. *banksii*) line (Fauré et al. 1993). Ninety-two individuals were scored for 58 RFLPs, four isozyme and 28 random amplified polymorphic DNA (RAPD) markers. Of the 90 loci that were identified, 77 could be grouped in 15 linkage groups, while the remaining 13 segregated independently. Segregation distortion was demonstrated for 36 loci mostly favouring the male ssp. banksii parent, and is assumed the result of heterozygosity for a structural rearrangement of the parents. Lowering the LOD score between 3 and 4 allowed the merging of four out of the 15 linkage groups resulting in the expected 11 (chromosomal) linkage groups. To account for large segregation distortion, pseudo linkages and difficulties in ordering markers around putative translocation breakpoints, Hippolyte et al. (2010) designed novel genetic linkages strategies for a segregating F1 population of 180 individuals from a cross between *M. acuminata*, 'Borneo' and 'Pisang Lilin', that were known to be genetically distant accessions. Microscopic analysis of chromosome configurations in pollen mother cells at diakinesis – metaphase I showed multivalent associations in a small number of cells, including trivalent, quadrivalent, quinquevalent and a hexavalent in the 'Borneo' accessions suggesting heterozygosity for two structural re-arrangements. The 'Pisang Lilin', however, displayed only an adjacent quadrivalent assuming one structural polymorphism and an anaphase I bridge which can be explained by a crossover in the loop of a paracentric inversion heterozygote. The linkage analysis was based on 489 segregating markers (167 SSRs and 322 DArTs) for the construction of gamete recombination of each parent, leading to the two parental maps with the expected eleven linkage groups and a total genetic map length of 1197 cm. Segregation

distortion, demonstrated for 21.7% of the markers (P < 0.05) occurred in different linkage groups for each parent. The putative re-arrangements on the chromosomes 1 and 4, and on chromosome 10 were visualized by neighbour-joining trees that are designed from Kosambi's distance calculation. An extension of this genetic analysis was done for the study of the banana genome (D'Hont et al. 2012) using a set of 589 SSR markers, as much as possible equally dispersed along the genome, along with 63 DArT markers and genotyped over 180 individuals. Also, in this map the linkage groups 1 and 4 remain grouped and concentrated 72% of the distorted markers. Mbanjo et al. (2012) published a comparable linkage map of two half- sib diploid banana populations on the basis of allele-specific-polymerase chain reactions (AS-PCRs), diversity array technology (DArT) and SSR markers. The pseudo-testcross mapping strategy produced a maternal map of 15 linkage groups (LGs) covering 670 cm and a maternal map of 16 LGs with a total length of 698 cm and a combined map with 15 LGs and a total map length of 1004 cm. While sequence quality of the markers is generally good, incongruity in some cases suggests chromosomal rearrangements. Additional analysis of allelic ratios and patterns together with neighbour-joining trees (Hippolyte et al. 2010) enable the heterozygous status for the structural rearrangement to be resolved. The existence of this translocation is further substantiated in Martin et al. (2017), who used mate-pair sequencing, BAC-FISH, targeted PCR and marker (DArTseq) segregation in the progeny of a heterozygous *M. acuminata*. The heterozygous reciprocal translocation features two distal 3 Mb and 10 Mb segments from chromosomes 01 and 04 (chromosomes 1, 1T4, 4T1 and 4), respectively, and show high segregation distortion, reduced recombination and linkage between chromosomes 01 and 04 in its progeny. The rearranged chromosome structure can also be observed in triploid cultivars but is present only in wild *malaccensis* ssp. accessions, implying that this rearrangement occurred in *M. acuminata* ssp. *malaccensis*. The observed high transmission of the translocation chromosomes suggests a mechanism in which the non-translocated chromosome 1 is lost, and so favours spreading of this translocation chromosomes to other populations and accessions.

Case study

Various aspects of chromosome research and genetics are now illustrated on the basis of a study of a diploid heterozygous Fusarium wilt resistant accession of wild banana *Musa acuminata* Colla ssp. *malaccensis* (LIPI-010) from LIPI (Lembaga Ilmu Pengetahuan Indonesia, Indonesian Institute of Sciences), Bogor, Indonesia. A segregating population of 255 descendants obtained from a selfed heterozygote had been created for the mapping of QTLs (quantitative trait loci) conferring Fusarium

resistance (Ahmad et al. 2020a). As structural hybridity was presumed, a cytogenetic analysis of male meiosis was carried out. To this end, we selected flower buds of the parent plant with anthers of 20 mm or less containing pollen mother cells at all stages of meiosis or pollen grains, and fixed the material in a fresh mixture of methanol:glacial acetic acid (3:1) for about 1 h before transferring the anthers to 70% ethanol and kept at 4°C for longer storage. Cell walls of the pollen mother cells were softened by incubating anthers in a mix of pectolytic enzymes followed by maceration in 45% acetic acid following Kantama et al. (2017). Chromosome slides were prepared by a standard squash or cell spreading technique. After selecting the slides with pollen mother cells at the appropriate stage under a phase contrast microscope we left the slides to dry overnight before mounting in Vectashield® containing 300 ng/µL DAPI. Slides were then studied under a fluorescence microscope with N.A. 1.4 Plan Apochromatic objectives and epifluorescence illumination containing a narrow band filter set for DAPI. Captured images were optimized for contrast, brightness and sharpness in Adobe Photoshop (Kantama et al. 2017). Pollen morphology and viability was studied in slides of pollen samples stained in a drop of lacto-phenol acid fuchsin (Sass 1964). Images of 100−150 pollen grains were captured using a bright-field microscope. Pollen grains were measured with a home-made pollen size ImageJ plugin and the obtained data for pollen size and morphology were processed in Microsoft Excel. Further technical details of the cytogenetic procedures are in Ahmad et al. (2020b).

Banana anthers contain relatively few pollen mother cells at the same stage, while thick cloudy cytoplasm, with lots of polysaccharides and/ or polyphenolics, masks details of the chromosomes. The opacity of the cytoplasm could partly be improved by using methanol instead of ethanol in the fixative, long pectolytic digestion of the cell walls and prolonged careful maceration in 45% acetic acid (Kantama et al. 2017; Ahmad et al. 2020b). We focussed on pollen mother cells at diakinesis, metaphase I and anaphase I, but only few of them displayed sufficiently clear details. Cells that we were able to interpret show only bivalents or overlapping chromosomes that resemble alternating ring quadrivalents (Figure 1A and B). We did not detect complements with clear adjacent (open) ring quadrivalents, chain quadrivalents and trivalent + univalent combinations. Some of the configurations could be interpreted as overlapping bivalents or as compact alternate ring quadrivalents, and were further disregarded. In pollen mother cells at anaphase I, we found balanced chromosome segregations and very incidentally a dicentric chromosome bridge (lacking a clear acentric fragment) and some lagging chromosomes (Figure 1C). The anaphase bridge was explained as the result of a crossover in the loop of a chromosome pair heterozygous for a paracentric inversion. Acentric fragments, if any, remained undetectable by the thick granular cytoplasm. Pollen was in general well-stained while pollen size values displayed a relatively narrow distribution

(Figure 2). The results were interpreted in a way that the parent was homozygous for the translocation, thus explaining the high fertility of the male gametes (Sybenga 1975, 1992).
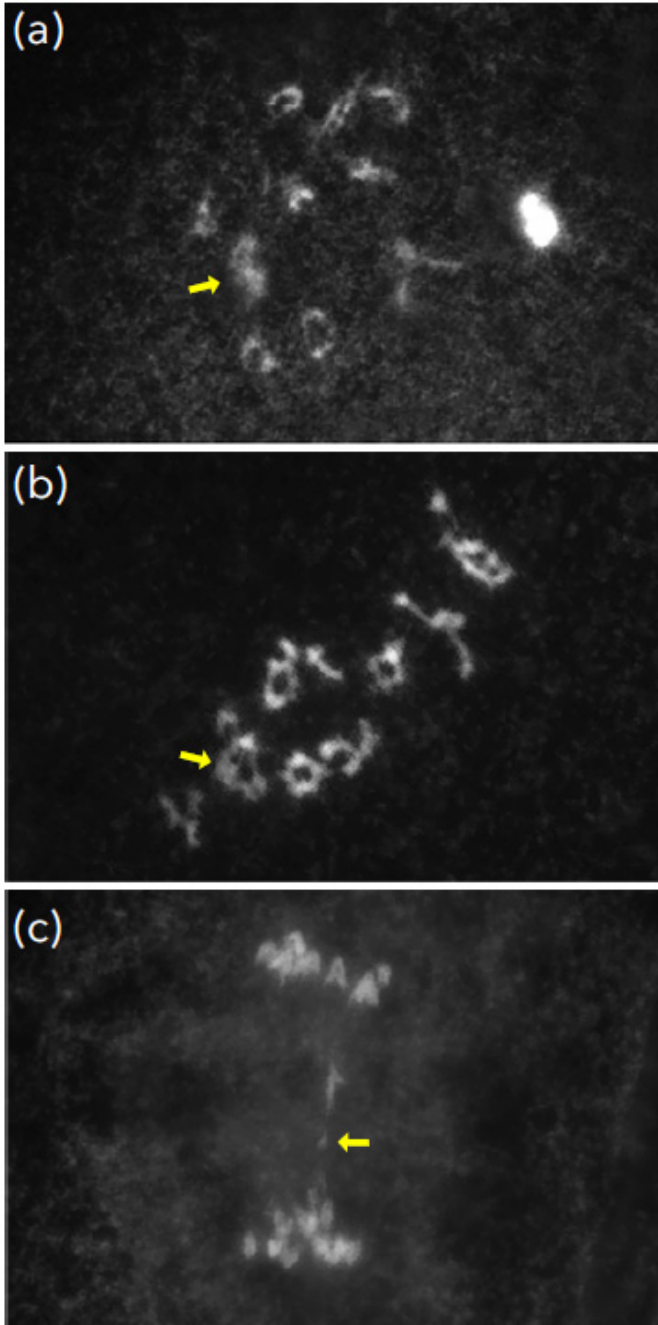
Figure 1 DAPI staining of spread pollen mother cells at meiosis of *Musa acuminata* ssp. *malaccensis*. (a) Cell complement at diakinesis. The arrow indicates two overlapping bivalents that resembles a ring quadrivalent. (b) Metaphase I complement. The arrow indicates a second example of overlapping bivalents resembling a quadrivalent. The arrowhead points at two overlapping bivalents. (c) Anaphase I cell with a dicentric chromosome bridge. The expected accompanying a-centric fragment could not be observed.
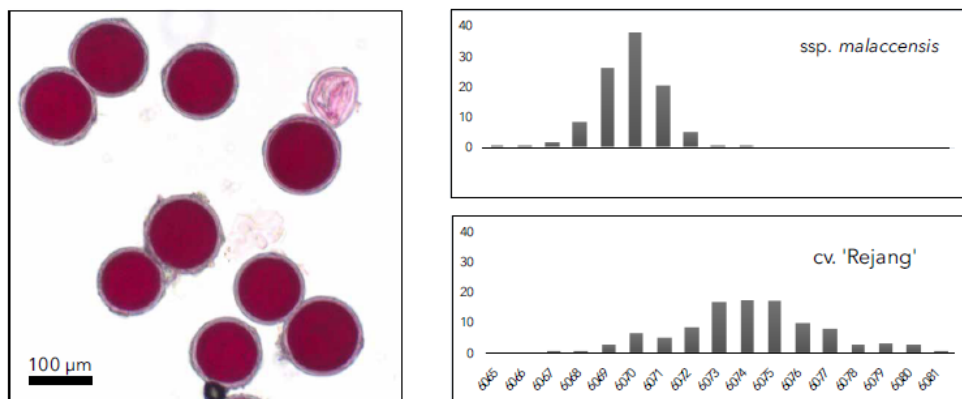
Figure 2 Pollen staining with lacto-phenol acid fuchsin (left). Most spores of the *M. acuminata* ssp. malaccensis (LIPI-010) are uniform and only very few are shrunken and colourless, and so representing dead pollen. Measurements of their size show little variation (right), in contrast to that of the M. accuminata cv. 'Rejang', which may include aneuploid and unreduced spores as well. The X-axis shows Image J pixel units of the pollen size; the Y-axis the number of measured pollen grains.

Genotyping and comparison of the genetic and physical maps

We used the DArTseq platform (https://www .diversityarrays .com) (Kilian et al. 2012) for genotyping the heterozygous parent and 217 descendants. Markers were putatively positioned on the genome assembly of the *M. acuminata* ssp. *malaccensis* DH 'Pahang', version 2 (http://banana-genome-hub.southgreen.fr/organism/Mu sa/acuminata) by aligning the DArTseq sequences to this reference genome. We collected leaf samples from the segregating population and the parent from both leaf materials from plantlets in tissue culture or from cigar leaves in the greenhouse. We freeze-dried these samples and isolated DNA using the Wizard® Magnetic DNA Purification System from Food Promega. The DNA concentration was quantified using Quant-iT PicoGreen dsDNA Assay Kit from Invitrogen and the quality was checked using 1% gel electrophoresis. The DNA was sent to Diversity Arrays Technology Pty Ltd, Australia, for scoring SNP-markers, using the DArTseq platform (http://www.diversityarrays .com/). Based on the DNA sequences flanking the SNPs, the markers were putatively positioned on the genome assembly of the Musa acuminata ssp. malaccensis DH 'Pahang', version 2 (http://banana-genome-hub.southgreen.fr/organism/Musa/acuminata). Eventually, 217 offspring and the parent were genotyped and 32 362 SNP markers were retrieved from the DArTseq (Ahmad et al. 2020a). The very high number of SNP-markers allowed us to apply a very stringent filtering, using the following criteria: (1) replicate value = 1; (2)

sequences should hit only one position on the reference genome; (3) The polymorphism information content >0.3; (4) DNA markers should be based on more than 10 calls per allele; (5) We only considered markers that were heterozygous in the parent, and therefore could segregate in the progeny. The physical positions of the filtered, segregating markers were checked on the reference genome for detection of possible homozygous regions in the self-pollinated parent. Apparently, the whole genome of the parent must have been heterozygous as there was no region devoid of markers, allowing construction of a genetic map for the whole genome. For the construction of the linkage map we initially used JoinMap® 5 software (https://www.kyazma.nl/index.php/JoinMap/) using the setting of 'F2' population type, thus regarding the heterozygous parent as an F1 from a cross between two homozygous grandparents. Duplicate markers were identified and merged where possible (conflicting marker scores made missing). As well as six marker pairs tagging the same single nucleotide polymorphism (SNP) position, 166 marker pairs with identical scores in the population were identified and merged, leaving 2630 unique markers. A single individual with a large proportion of missing values (28%) was also removed, leaving a mapping population size of 216 individuals. Marker linkage analysis was performed using the polymapR package (Bourke et al. 2018), with parental dosages specified as '1' at all markers for the selfed parent. Linkage groups were identified after clustering at a LOD threshold of 4, and marker ordering was performed using the MDSMap package with default settings (Preedy and Hackett 2016). Linkage groups were renumbered according to the *M. acuminata* ssp. malaccensis reference assembly (D'Hont et al. 2012), with the genetic positions reversed if the orientation conflicted with the order on the physical map. A comparison of the genetic and physical maps was visualized using R version 3.5.0 (R Core Team 2018). As there was evidence of map inflation (due to possible genotyping errors), the genetic map was further refined with an in-house implementation of the Smooth algorithm (Van Os et al. 2005), and marker information of 11 individuals with an exceptionally high number of re-combinations removed. Map smoothing was repeated using a decreasing delta threshold from 0.99 to 0.71 in steps of 0.02 (where delta is the threshold for declaring singletons), with marker clustering repeated only after the first round (clustering was unchanged) and marker order re-estimated at each round using the polymapR pipeline and the MDSMap ordering algorithm.

Linkage analysis

Eleven linkage groups were successfully built from the segregating population of *M. acuminata* ssp. *malaccensis* (LIPI-010) with values ranging from 103 cm to 184 cm, totalling 1,640 cm (Table 1). This result is in the range of previously published genetic map

lengths in the ssp. *malaccensis* DH 'Pahang' (D'Hont et al. 2012; Martin et al. 2017). The linkage maps compared to the physical map clearly showed that interchanges involving chromosome 1 and 4 at position 118−128 cm (between markers chr01-4,504,873 and chr04-28,778,311) on chromosome 1 and at position 174−184 cm (between markers chr04-26,534,556 and chr01-763,137) on chromosome 4. Furthermore, the comparison of the genetic maps and the physical maps of references showed markers on parts of chromosomes 1 and 4 did not map to their expected positions in linkage groups 1 and 4, also indicating the presence of a translocation between these chromosomes (Figure 3). However, the genetic map did not indicate heterozygosity for the translocation. Although microscopic observations of diakinesis complements of the parent revealed few pollen mother cells with four overlapping chromosomes that resemble an alternative ring quadrivalent, we were not sufficiently convinced about heterozygosity for the translocation of the ssp. malaccensis (LIPI-010) mother plant as adjacent quadrivalents were lacking and pollen fertility was high.

Table 1 Overview of the genetic analysis with smoothing of the diploid heterozygous *Musa acuminata* Colla ssp. *malaccensis* using the polymapR package LG1.

|	N	Original	Smoothed
LG1	274	180.16	163.47
LG2	215	113.39	103.31
LG3	295	192.66	178.76
LG4	170	168.07	138.04
LG5	245	163.71	147.08
LG6	300	202.79	184.31
LG7	212	166.89	143.43
LG8	282	182.51	162.46
LG9	247	178.17	152.05
LG10	200	162.11	142.98
LG11	176	140.66	123.67
total	2616	1851.12	1639.56

To further investigate the nature of the 1T4/4T1 translocation we performed an additional PCR with primers around the translocation breakpoints that confirmed the presence of both chromosomes 1 and 4, and the translocation chromosomes in the parent (Figure 4). A simplified genetic map with the 11 linkage groups is shown in Figure 5, along with the translocated segments and the markers flanking the translocation breakpoints. We conclude that the translocation in this study is identical to the 1T4/4T1 translocation in the PT-BA-00267 accession of *M. acuminata* ssp. *malaccensis* as described in Martin et al. (2017). Both share this translocation with two distinct genetic groups corresponding

to ssp. *malaccensis* and the Sucrier cultivar subgroup, whereas heterozygous accessions were related to both *M. acuminata* ssp. banksii, *M. acuminata* ssp. *malaccensis*, and the Sucrier cultivar subgroup. The genetical data do not support heterozygosity for the translocation; the microscopic observations do not support a translocation complex either, whereas our PCR data are in favour of heterozygosity of the translocation in the *malaccensis* parent. Strong segregating distortion for chromosome 10 in the genetic analysis might explain heterozygosity for a paracentric inversion in which crossovers in the inverted regions lead to dicentric bridges and so loss of chromosomes at anaphase I (data not shown).

4
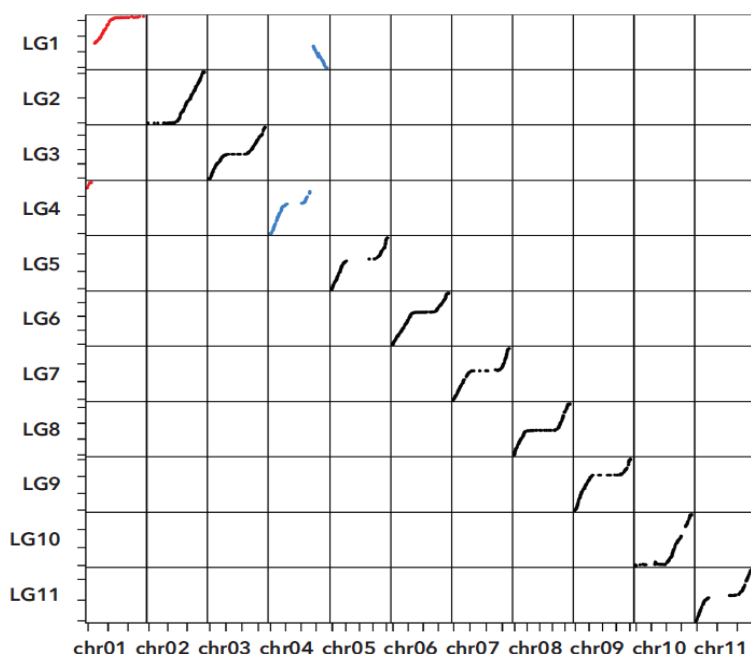
Figure 3 Comparison of the genetic maps with linkage groups LG1-11 (Y-axis) and physical (X-axis) positions according to DH 'Pahang' banana in Martin et al. (2017). In total 2,616 SNP markers are mapped in the segregating population of the heterozygous *M. acuminata* ssp. *malaccensis* (LIPI-010). The Y-axis tick marks correspond to 50 cm intervals, while X-axis tick marks correspond to 10 Mb intervals. Markers on linkage groups carrying the translocation between the chromosomes 1 and 4 have been highlighted in red and blue, respectively.
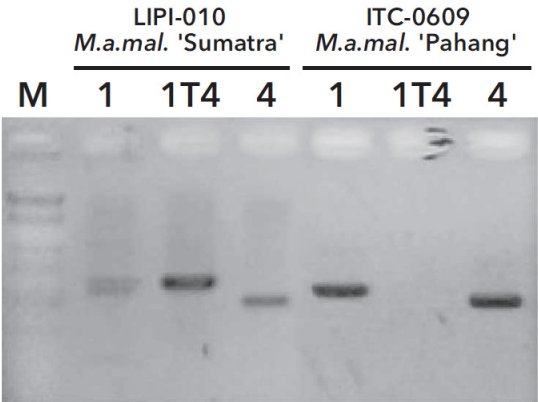
Figure 4 PCR of genomic DNA of the *M. acuminata* ssp. malaccensis (LIPI-010) and the reference genome *M. acuminata* ssp. *malaccensis* 'Pahang' (ITC-0609). The forward and reverse primers around the translocation breakpoint were described in Martin et al. (2017). The LIPI-010 plant has both chromosome 1 and 4, and the translocation chromosomes 1T4 and 4T1, and so is heterozygous for the structural rearrangement. The reference 'Pahang' does not have the translocation as previously described in D'Hont et al. (2012) and Martin et al. (2017).
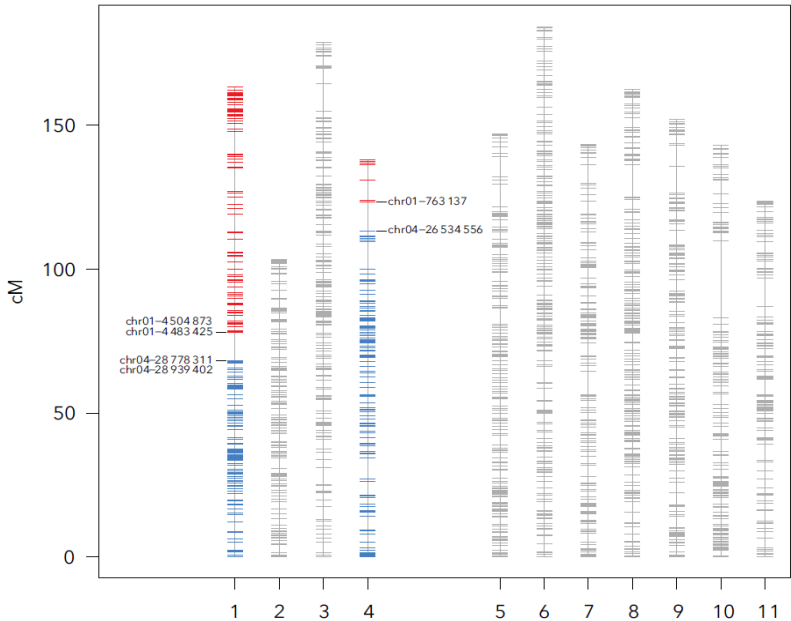
Figure 5 Genetic map of the *M. acuminata* ssp. *malaccensis* (LIPI-010) based on 2616 SNP markers mapped in the segregating population of the heterozygote. Markers proximal and distal of the translocation breakpoints are indicated and the corresponding translocation segments are displayed in red and blue.

Conclusion and future perspectives

Our study of the *M. acuminata* ssp. *malaccensis* has clearly proven the existence of the 1T4/4T1 translocation as documented in Martin et al. (2017). The observation is in line with the reports on structural hybridity as a general phenomenon in banana, which along with parthenocarpy, polyploidy and interspecific hybridizations form a serious limitation in breeding of this crop. It is currently accepted that transposons are the engines that create most structural rearrangements. Activation of these elements requires demethylation or other epigenetic changes of its DNA that enables transposition and so leads to breakage and ligation of chromosome parts (Bennetzen 2000; Lönnig and Saedler 2003; Zhang et al. 2011; Lysak and Schubert 2013). In this respect it is obvious to find the footprints of the transposon(s) that are responsible for the rearrangements in *M. acuminata* and subsequently, why these transposons are relatively active in this species. In the Columbia accession of Arabidopsis thaliana, a paracentric inversion in the short arm of chromosome 4 has breakpoints in F-box protein-coding genes of the RNI-like superfamily and Vandal5, a Mutator-like (Mule) transposon, suggesting that the inversion was created by the activity of the transposon (Fransz et al. 2016). As genomic sequences of the reference DH 'Pahang' and that of the 1–4 translocation are known, bioinformatic analyses provide excellent tools to identify such translocations at the DNA level, such as split-read mapping or paired-end-read mapping assuming that there are no long stretches of repetitive sequences around the breakpoints (Lin et al. 2015). Synteny-based comparisons of genomic sequences like the MUMmer plot tool (http://mummer .sourceforge .net/) can directly reveal the chromosomal-rearranged regions. Zooming into the breakpoints one can reveal the footprints of transposons that may have transposed in the regions of interest. If such breakpoints are caused by one and the same transposable element, it will be straightforward to identify breakpoints of yet unknown translocations and inversions. Alternatively, new, unknown translocations and inversions can be identified by FISH with chromosome-specific BACs or oligo painting pools (Han et al. 2015), microarrays (Bhat et al. 2007), whole-genome sequencing (Dong et al. 2018), long-range sequencing (Hu et al. 2018; de Coster et al. 2019). Genetic mapping is another useful tool for highlighting large structural re-arrangements. Now that most markers are sequence-based, it is possible to compare map orders and check whether there are cross-specific re-arrangements, or structural variation in comparison to the reference genome. The trend towards pangenomics would provide a convenient way to record these re-arrangements. As structural hybridity is one of the major sources of sterility in banana, it is imperative to detect translocations and inversions in common diploid subspecies and cultivars of *M. acuminata* with favourable traits like disease resistances, using a combination of the methods described above. Accessions sharing the same rearrangements can be used

4

for introgressive hybridization or traditional breeding work. In a final step, accessions differing in two or more translocations can be crossed to produce sterile diploids, or select triploid descendants from incidentally formed 2n gametes. Alternatively, recent studies have shown that CRISPR Cas9 technology is a useful alternative to correct translocations and inversions to their original wild-type arrangement (Lekomtsev et al. 2016; Jiang et al. 2016; Lynagh et al. 2018). However, the practical application of such genomic corrections still awaits acceptance of the public and policymakers for GMO plant materials for breeding purposes. We are in an exciting time where microscopy techniques are probing chromosomes in ever finer detail while at the same time molecular techniques are upscaling to long-range sequences and sub-chromosomal information. However, there are still many caveats. The current reference genomes of banana (both 'Pahang' V2 and *M. balbisiana* 'Pisang Klutuk Wulung') are still incomplete and not guaranteed reliable along all of their length. There is still too little data on the genomic sequence and structural diversity in banana cultivars and wild relatives aside from some data on one single translocation between chromosomes 1 and 4. The details of the structural variants reported by Martin et al. (2017) may not be congruent with the genetic map from the wild Indonesian *Musa acuminata* ssp. *malaccensis* probably because of these gaps in our knowledge.

Where to look for further information

Cytogenetics of Musa species, hybrids and cultivars has long been used for determining ploidy, aneuploidy and structural variations. Besides, the study of chromosomes together with genetic analyses plays a vital role in supporting banana breeding, especially in introgressive hybridization programmes, providing essential information on meiotic pairing, recombination and transmission, and formation of unreduced gametes in interspecific hybrids and their backcross derivatives. Recently advanced fluorescent in situ hybridization technologies includes molecular protocols and genomics information enabling chromosomal positioning of single-copy and repetitive sequences on mitotic and meiotic cells and nuclei. For understanding meiotic chromosome processes and aberrations in plant species and hybrids and especially in bananas we recommend:

- Shepherd, K. (1999). Cytogenetics of the Genus *Musa*. INIBAP. This monograph clearly describes the complexity of structural rearrangements between Musa (sub)species and cultivars.
- Sybenga, J. (1975). Meiotic Configurations. Springer-Verlag.
- Sybenga, J. (1992). Cytogenetics in Plant Breeding. Springer Verlag. Both monographs nicely clarify the complex meiotic behaviour of plants heterozygous for structural heterozygosity, polyploids and aneuploids.

Outstanding examples of recent advanced molecular cytogenetics of Musa are:

- Martin, G., Carreel, F., Coriton, O., Hervouet, C., Cardi, C., Derouault, P., Roques, D., et al. (2017). Evolution of the banana genome (Musa acuminata) is impacted by large chromosomal translocations. Molecular Biology and Evolution 34(9): 2140–2152. https://doi .org/10.1093/molbev/msx164. This study combines DNA sequencing, chromosome painting and genetic analysis to characterize a reciprocal translocation 1^4 /4^1 in an *M. acuminata* ssp. *malaccensis* banana accession.
- Šimoníková, D., Němečková, A., Čížková, J., Brown, A., Swennen, R., Doležel, J. and Hřibová, E. (2020). Chromosome painting in cultivated banana and their wild relatives (Musa spp.) reveals differences in chromosome structures, bioRxiv, preprint. doi: https://doi.org/10.1101/2020.08.01.232207. The authors used the recently developed powerful oligo-FISH painting technology to elucidate translocations in *M. acuminata*, *M. balbisiana* and their hybrids.

Important general websites about Musa research, breeding and agriculture:

- http://www.promusa.org/. This website is the general online platform of ProMusa a network of people promoting scientific discussions on bananas. The site also includes information on the global Musa genomics consortium.
- https://www.crop-diversity.org/mgis/organisations. The Musa Germplasm information system contains critical information on Musa germplasm diversity, including passport data, botanical classification, morpho-taxonomic descriptors, molecular studies, plant photographs and GIS information on 6548 accessions managed in 29 collections around the world.
- https://en .wikipedia .org/wiki /MusaNet. MusaNet is a global network of scientists and other stakeholders working on banana (*Musa* spp.) genetic resources.
- https://fusariumwilt.org/index.php/en/about-fusarium-wilt/. An informative website about banana and one of its most devastating threads in banana culture: fusarium wilt.

4

References

Adeleke, M. T. V., Pillay, M. and Okoli, B. E. (2002). An improved method for examining meiotic chromosomes in *Musa* L. HortScience 37(6): 959–961.

Agarwal, P. K. (1983). Cytogenetical investigations in Musaceae I. Meiotic studies in South Indian bananas. Cytologia 48(4): 847–852.

Agarwal, P. K. (1988). Cytogenetical investigations in Musaceae III. Meiotic studies in South Indian bananas. Cytologia 53(2): 359–363.

Ahmad, F., Martiwi, N. M., Poerba, Y. S., de Jong, H., Schouten, H. and Kema, G. H. J. (2020a). Genetic mapping of Fusarium Wilt resistance in a wild banana *Musa acuminata* ssp. *malaccensis* accession. In preparation.

Ahmad, F., Poerba, Y. S., Kema, G. H. J. and de Jong, H. (2020b). Meiotic studies of diploid Indonesian banana varieties and cultivars. In: F. Ahmad, PhD thesis. Wageningen University.

Albini, S. M. (1994). A karyotype of the Arabidopsis thaliana genome derived From synaptonemal complex analysis at prophase I of meiosis. The Plant Journal 5(5): 665–672.

Alexander, M. P. (1969). Differential staining of aborted and non-aborted pollen. Stain Technology 44(3): 117–122.

Arango, R. E., Togawa, R. C., Carpentier, S. C., Roux, N., Hekkert, B. L., Kema, G. H. J. and Souza, M. T. (2011). Genome-wide BAC-end sequencing of *Musa acuminata* DH Pahang reveals further insights Into the genome organization of banana. Tree Genetics and Genomes 7(5): 933–940. doi:10.1007/s11295-011-0385-3.

Arora, V., Kapoor, N., Fatma, S., Jaiswal, S., Iquebal, M. A., Rai, A. and Kumar, D. (2018). BanSatDB, a whole-genome-based database of putative and experimentally validated microsatellite markers of three *Musa* species. The Crop Journal 6(6): 642–650. doi:10.1016/j.cj.2018.01.005.

Bakry, F., Horry, J. P., Teisson, C., Tezenas Du Montcel, H. and Ganry, J. (1990). L'amélioration génétique des bananiers à l'IRFA/CIRAD, [Genetic improvement of banana plants at IRFA/CIRAD]. Fruits (numero speciale): 25–40.

Bakry, F. and Shepherd, K. (2008). Chromosome count on banana root tip squashes. Fruits 63(3): 179–181. doi:10.1051/fruits:2008008.

Balint-Kurti, P. J., Clendennen, S. K., Doleželová, M., Valárik, M., Doležel, J., Beetham, P. R. and May, G. D. (2000). Identification and chromosomal localization of the monkey retrotransposon in *Musa* sp. Molecular and General Genetics: MGG 263(6): 908–915.

Bartoš, J., Alkhimova, O., Doleželova, M., De Langhe, E. and Doležel, J. (2005). Nuclear genome size and genomic distribution of ribosomal DNA in

Musa and *Ensete* (Musaceae): taxonomic implications. Cytogenetic and Genome Research 109(1–3): 50–57. doi:10.1159/000082381.

Bennetzen, J. L. (2000). The many hues of plant heterochromatin. Genome Biology 1(1): 107.1–4. doi:10.1186/gb-2000-1-1-reviews107.

Bhat, P. R., Lukaszewski, A., Cui, X., Xu, J., Svensson, J. T., Wanamaker, S., Waines, J. G. and Close, T. J. (2007). Mapping translocation breakpoints using a wheat microarray. Nucleic Acids Research 35(9): 2936–2943. doi:10.1093/nar/gkm148.

Bouilly, K., Chaves, R., Leitão, A., Benabdelmouna, A. and Guedes-Pinto, H. (2008). Chromosomal organization of simple sequence repeats in the pacific oyster (*Crassostrea gigas*): (GGAT)(4), (GT)(7) and (TA)(10) chromosome patterns. Journal of Genetics 87(2): 119–125.

Bourke, P. M., van Geest, G., Voorrips, R. E., Jansen, J., Kranenburg, T., Shahin, A., Visser, R. G. F., Arens, P., Smulders, M. J. M. and Maliepaard, C. (2018). polymapR - linkage analysis and genetic map construction from F1 populations of outcrossing polyploids. Bioinformatics 34(20): 3496–3502.

Braz, G. T., He, L., Zhao, H., Zhang, T., Semrau, K., Rouillard, J. M., Torres, G. A. and Jiang, J. (2018). Comparative oligo-FISH mapping: an efficient and powerful methodology to reveal karyotypic and chromosomal evolution. Genetics 208(2): 513–523.

doi:10.1534/genetics.117.300344.

Buhariwalla, H. K., Jarret, R. L., Jayashree, B., Crouch, J. H. and Ortiz, R. (2005). Isolation and characterization of microsatellite markers From *Musa balbisiana*. Molecular Ecology Notes 5(2): 327–330. doi:10.1111/j.1471-8286.2005.00916.x.

Capdeville, G. de, Teixeira Souza Júnior, M., Szinay, D., Eugênio Cardamone Diniz, L., Wijnker, E., Swennen, R., Kema, G. H. J. and Jong, H. de (2008). The potential of high-resolution BAC-FISH in banana breeding. Euphytica 166(3): 431–443. doi:10.1007/s10681-008-9830-2.

Cheung, F. and Town, C. D. (2007). A BAC end view of the *Musa acuminata* Genome. BMC Plant Biology 7(1): 29. doi:10.1186/1471-2229-7-29.

Christelová, P., Valarik, M., Hřibova, E., Van den Houwe, I., Channelière, S., Roux, N. and Doležel, J. (2011). A platform for efficient genotyping in *Musa* using microsatellite markers. AoB Plants 2011: plr024. doi:10.1093/aobpla/plr024.

Čížková, J., Hřibova, E., Humplíková, L., Christelová, P., Suchánková, P. and Doležel, J. (2013). Molecular analysis and genomic organization of major DNA satellites in banana (*Musa* spp.). PLoS ONE 8(1): e54808. doi:10.1371/journal.pone.0054808.

Creste, S., Neto, A. T., Oliveira Silva, S. de and Figueira, A. (2003). Genetic characterization of banana cultivars (*Musa* spp.) From brazil using microsatellite markers. Euphytica

132(3): 259–268.

Cuadrado, Á. and Jouve, N. (2010). Chromosomal detection of simple sequence repeats (SSRs) using nondenaturing FISH (ND-FISH). Chromosoma 119(5): 495–503. doi:10.1007/s00412-010-0273-x.

Damaiyani, J. and Hapsari, L. (2017). Pollen viability of 19 Indonesian bananas (*Musa* L.) collection of Purwodadi Botanic Garden: preliminary study for breeding. Proceeding of the International Conference on Tropical Plant Conservation and Utilization, 1–10.

Darlington, C. D. (1929). Chromosome behaviour and structural hybridity in the Tradescantiae. Journal of Genetics 21(2): 207–286.

de Coster, W., De Rijk, P., De Roeck, A., De Pooter, T., D'Hert, S., Strazisar, M., Sleegers, K. and Van Broeckhoven, C. (2019). Structural variants identified by Oxford nanopore PromethION sequencing of the human genome. Genome Research 29(7): 1178– 1187. doi:10.1101/gr.244939.118.

de Jong, H., Fransz, P. and Zabel, P. (1999). High resolution FISH in plants - techniques and applications. Trends in Plant Science 4(7): 258–263.

D'Hont, A., Paget-Goy, A., Escoute, J. and Carreel, F. (2000). The interspecific genome structure of cultivated banana, *Musa* spp. Revealed by genomic DNA in situ hybridization. Theoretical and Applied Genetics

100(2): 177–183.

D'Hont, A. (2005). Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. Cytogenetic and Genome Research 109(1–3): 27–33. doi:10.1159/000082378.

D'Hont, A., Denoeud, F., Aury, J. M., Baurens, F. C., Carreel, F., Garsmeur, O., Noel, B., Bocs, S., Droc, G., Rouard, M., Da Silva, C., Jabbari, K., Cardi, C., Poulain, J., Souquet, M., Labadie, K., Jourda, C., Lengellé, J., Rodier-Goud, M., Alberti, A., Bernard, M., Correa, M., Ayyampalayam, S., Mckain, M. R., Leebens-Mack, J., Burgess, D., Freeling, M., Mbéguié-A-Mbéguié, D., Chabannes, M., Wicker, T., Panaud, O., Barbosa, J., Hribova, E., Heslop-Harrison, P., Habas, R., Rivallan, R., Francois, P., Poiron, C., Kilian, A., Burthia, D., Jenny, C., Bakry, F., Brown, S., Guignon, V., Kema, G., Dita, M., Waalwijk, C., Joseph, S., Dievart, A., Jaillon, O., Leclercq, J., Argout, X., Lyons, E., Almeida, A., Jeridi, M., Dolezel, J., Roux, N., Risterucci, A. M., Weissenbach, J., Ruiz, M., Glaszmann, J. C., Quétier, F., Yahiaoui, N. and Wincker, P. (2012). The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. Nature 488(7410): 213–217.

Dodds, K. S. (1943). Genetical and cytological studies of *Musa*, V. Certain edible diploids. Journal of Genetics 45(2): 113–138.

Dodds, K. S. and Simmonds, N. W.

(1946). Genetical and cytological studies of *Musa*, VIII. the formation of polyploid spores. Journal of Genetics 47: 223–241.

Doleželová, M., Valárik, M., Swennen, R., Horry, J. P. and Doležel, J. (1998). Physical mapping of the 18S-25S and 5S ribosomal RNA genes in diploid bananas. Biologia Plantarum 41(4): 497–505.

Dong, Z., Wang, H., Chen, H., Jiang, H., Yuan, J., Yang, Z., Wang, W. J., Xu, F., Guo, X., Cao, Y., Zhu, Z., Geng, C., Cheung, W. C., Kwok, Y. K., Yang, H., Leung, T. Y., Morton, C. C., Cheung, S. W. and Choy, K. W. (2018). Identification of balanced chromosomal rearrangements previously unknown Among participants in the 1000 genomes project: implications for interpretation of structural variation in genomes and the future of clinical cytogenetics. Genetics in Medicine: Official Journal of the American College of Medical Genetics 20(7): 697–707.

Fauré, S., Noyer, J. L., Horry, J. P., Bakry, F., Lanaud, C. and Gońzalez de León, D. (1993). A molecular marker-based linkage map of diploid bananas (*Musa acuminata*). TAG. Theoretical and Applied Genetics. Theoretische und Angewandte Genetik 87(4): 517–526. doi:10.1007/BF00215098.

Fransz, P., Linc, G., Lee, C. R., Alves Aflitos, S. A., Lasky, J. R., Toomajian, C., Ali, H., Peters, J., van Dam, P., Ji, X., Kuzak, M., Gerats, T., Schu-bert, I., Schneeberger, K., Colot, V., Martienssen, R., Koornneef, M., Nordborg, M., Juenger, T. E., de Jong, H. and Schranz, M. E. (2016). Molecular, genetic and evolutionary analysis of a paracentric inversion in *Arabidopsis thaliana*. The Plant Journal: For Cell and Molecular Biology 88(2): 159–178. doi:10.1111/tpj.13262.

Godwin, I. D., Aitken, E. A. and Smith, L. W. (1997). Application of inter simple sequence repeat (ISSR) markers to plant genetics. Electrophoresis 18(9): 1524–1528. doi:10.1002/elps.1150180906.

Han, Y., Zhang, T., Thammapichai, P., Weng, Y. and Jiang, J. (2015). Chromosome-specific painting in *Cucumis* species using bulked oligonucleotides. Genetics 200(3): 771– 779. doi:10.1534/genetics.115.177642.

Hippolyte, I., Bakry, F., Seguin, M., Gardes, L., Rivallan, R., Risterucci, A. M., Jenny, C., Perrier, X., Carreel, F., Argout, X., Piffanelli, P., Khan, I. A., Miller, R. N., Pappas, G. J., Mbéguié-A-Mbéguié, D., Matsumoto, T., De Bernardinis, V., Huttner, E., Kilian, A., Baurens, F. C., D'Hont, A., Cote, F., Courtois, B. and Glaszmann, J. C. (2010). A saturated SSR/DArT linkage map of *Musa acuminata* addressing genome rearrangements Among bananas. BMC Plant Biology 10(1): 65. doi:10.1186/1471-2229-10-65.

Hou, L., Xu, M., Zhang, T., Xu, Z., Wang, W., Zhang, J., Yu, M., Ji, W., Zhu, C.,

Gong, Z., Gu, M., Jiang, J. and Yu, H. (2018). Chromosome painting and its applications in cultivated and wild rice. BMC Plant Biology 18(1): 110. doi:10.1186/ s12870-018-1325-2.

Hřibová, E., Doleželova, M., Town, C. D., Macas, J. and Doležel, J. (2007). Isolation and characterization of the highly repeated fraction of the banana genome. Cytogenetic and Genome Research 119(3–4): 268–274. doi:10.1159/000112073.

Hřibová, E., Neumann, P., Matsumoto, T., Roux, N., Macas, J. and Doležel, J. (2010). Repetitive part of the banana (*Musa acuminata*) genome investigated by low-depth 454 sequencing. BMC Plant Biology 10: 204. doi:10.1186/1471-2229-10-204.

Hu, L., Liang, F., Cheng, D., Zhang, Z., Yu, G., Zha, J. and Wang, Y. (2018). Localization of balanced chromosome translocation breakpoints by long-read sequencing on the Oxford nanopore platform. Biorxiv .or g. doi:10.1101/419531.

Jarret, R. L., Bhat, K. V., Cregan, P., Ortiz, R. and Vuylsteke, D. (1994). Isolation of microsatellite DNA markers in *Musa*. InfoMusa 3: 3–4.

Jeridi, M., Bakry, F., Escoute, J., Fondi, E., Carreel, F., Ferchichi, A., D'Hont, A. and Rodier- Goud, M. (2011). Homoeologous chromosome pairing Between the a and B genomes of *Musa* spp. Revealed by genomic in situ hybridization. Annals of Botany 108(5): 975–981. doi:10.1093/aob/mcr207.

Jeridi, M., Perrier, X., Rodier-Goud, M., Ferchichi, A., D'Hont, A. and Bakry, F. (2012). Cytogenetic evidence of mixed disomic and polysomic inheritance in an allotetraploid (AABB) *Musa* Genotype. Annals of Botany 110(8): 1593–1606. doi:10.1093/aob/mcs220.

Jiang, J., Zhang, L., Zhou, X., Chen, X., Huang, G., Li, F., Wang, R., Wu, N., Yan, Y., Tong, C., Srivastava, S., Wang, Y., Liu, H. and Ying, Q. L. (2016). Induction of site-specific chromosomal translocations in embryonic stem cells by CRISPR/ Cas9. Scientific Reports 6: 21918. doi:10.1038/srep21918.

Kaemmer, D., Fischer, D., Jarret, R. L., Baurens, F.-C., Grapin, A., Dambier, D., Noyer, J.-L., Lanaud, C., Kahl, G. and Lagoda, P. J. L. (1997). Molecular breeding in the genus Musa: a strong case for STMS marker technology. Euphytica 96(1): 49– 63.

Kantama, L., Wijnker, E. and de Jong, H. (2017). Optimization of cell spreading and image quality for the study of chromosomes in plant tissues. In: Schmidt, A. (Ed.), Plant Germline Development: Methods and Protocols, Methods in Molecular Biology (Vol. 1669). Springer, 141–158. doi:10.1007/978-1-4939-7286-9_12.

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., Caig, V., Heller-Uszynska, K., Jaccoud, D., Hop-

per, C., Aschenbrenner-Kilian, M., Evers, M., Peng, K., Cayla, C., Hok, P. and Uszynski, G. (2012). Diversity arrays technology: a generic genome profiling technology on open platforms. Methods in Molecular Biology 888: 67–89.

Lacadena, J. R., Cermeno, M. C., Orellana, J. and Santos, J. L. (1984). Evidence for wheat-rye nucleolar competition (Amphiplasty) in Triticale by silver-staining procedure. TAG. Theoretical and Applied Genetics. Theoretische und Angewandte Genetik 67(2–3): 207–213.

Lekomtsev, S., Aligianni, S., Lapao, A. and Bürckstümmer, T. (2016). Efficient generation and reversion of chromosomal translocations using CRISPR/ Cas technology. BMC Genomics 17(1): 739. doi:10.1186/s12864-016-3084-5.

Lescot, M., Piffanelli, P., Ciampi, A. Y., Ruiz, M., Blanc, G., Leebens-Mack, J., da Silva, F. R., Santos, C. M., D'Hont, A., Garsmeur, O., Vilarinhos, A. D., Kanamori, H., Matsumoto, T., Ronning, C. M., Cheung, F., Haas, B. J., Althoff, R., Arbogast, T., Hine, E., Pappas, G. J., Sasaki, T., Souza, M. T., Miller, R. N., Glaszmann, J. C. and Town, C. D. (2008). Insights Into the *Musa* Genome: syntenic relationships to rice and Between *Musa* species. BMC Genomics 9(1): 58. doi:10.1186/1471-2164-9-58.

Li, K., Wang, H., Wang, J., Sun, J., Li, Z. and Han, Y. (2016). Divergence Between C. melo and African *Cu-cumis* species identified by chromosome painting and rDNA distribution pattern. Cytogenetic and Genome Research 150(2): 150–155. doi:10.1159/000453520.

Lin, K., Smit, S., Bonnema, G., Sanchez-Perez, G. and de Ridder, D. (2015). Making the difference: integrating structural variation detection tools. Briefings in Bioinformatics 16(5): 852–864. doi:10.1093/bib/bbu047.

Lönnig, W. E. and Saedler, H. (2003). Chromosome rearrangements and transposable elements. Annual Review of Genetics 36(4139): 389–410. doi:10.1146/annurev.genet.36.040202.092802.

Lynagh, P. G., Inagaki, S., Amundson, K. R., Marimithu, M. P. A., Pike, B. R., Henry, I. M., Tan, E. H. and Comai, L. (2018). Translocation and duplication From CRISPR-Cas9 editing in Arabidopsis thaliana. bioRxiv. doi:10.1101/400507.

Lysak, M. A., Doleželová, M., Horry, J. P., Swennen, R. and Dolezˇel, J. (1999). Flow cytometric analysis of nuclear DNA content in *Musa*. Theoretical and Applied Genetics 98(8): 1344–1350.

Lysak, M. and Schubert, I. (2013). Mechanisms of chromosome rearrangements. In: Leitch, I. J., Greilhuber, J., Doležel, J. and Wendel, J. F. (Eds), Plant Genome Diversity (Vol. 2), Physical Structure, Behaviour and Evolution of Plant Genomes. Springer-Verlag, Wien, 137–147. doi:10.1007/978-3-7091-1160-4_9.

4

Martin, G., Baurens, F. C., Droc, G., Rouard, M., Cenci, A., Kilian, A., Hastie, A., Doležel, J., Aury, J. M., Alberti, A., Carreel, F. and D'Hont, A. (2016). Improvement of the banana 'Musa acuminata' reference sequence using NGS data and semi-automated bioinformatics methods. BMC Genomics 17(243): 243. doi:10.1186/ s12864-016-2579-4.

Martin, G., Carreel, F., Coriton, O., Hervouet, C., Cardi, C., Derouault, P., Roques, D., Salmon, F., Rouard, M., Sardos, J., Labadie, K., Baurens, F. C. and D'Hont, A. (2017). Evolution of the banana genome (*Musa acuminata*) is impacted by large chromosomal translocations. Molecular Biology and Evolution 34(9): 2140–2152. doi:10.1093/molbev/ msx164.

Mbanjo, E. G. N., Tchoumbougnang, F., Mouelle, A. S., Oben, J. E., Nyine, M., Dochez, C., Ferguson, M. E. and Lorenzen, J. (2012). Molecular marker-based genetic linkage map of a diploid banana population (*Musa acuminata* Colla). Euphytica 188(3): 369– 386. doi:10.1007/s10681-012-0693-1.

Msogoya, T., Grout, B. and Roberts, R. (2008). Karyotypic and 2C nuclear DNA in vitro induced off-types of East African Highland banana (*Musa AAA East Africa*). Biotechnology 7(3): 578–581.

Nair, A. S., Teo, C. H., Schwarzacher, T. and Heslop-Harrison, P. H. (2005).

Genome classification of banana cultivars from South India using IRAP markers. Euphytica 144(3): 285–290. doi:10.1007/s10681-005-7321-2.

Novák, P., Hřibova, E., Neumann, P., Koblížková, A., Doležel, J. and Macas, J. (2014). Genome-wide analysis of repeat diversity across the family Musaceae. PLoS ONE 9(6): e98918. doi:10.1371/journal.pone.0098918.

Ortiz, R. (1995). Musa genetics. In: Gowen, S. (Ed.), Bananas and Plantains. Chapman & Hall, London, 84–109.

Oselebe, H. O., Tenkouano, A. and Pillay, M. (2009). Ploidy variation of *Musa* hybrids From crosses. African Journal of Biotechnology 5: 1048–1053.

Osuji, J. O., Harrison, G., Crouch, J. and Heslop-Harrison, J. S. (1997). Identification of the genomic constitution of Musa L. Lines (bananas, plantains and hybrids) using molecular cytogenetics. Annals of Botany 80(6): 787–793.

Osuji, J. O., Okoli, B. E. and Edeoga, H. O. (2006). Karyotypes of the a and B genomes of *Musa* L. Cytologia 71(1): 21–24. doi:10.1508/cytologia.71.21.

Padmesh, P., Mukunthakumar, S., Vineesh, P. S., Skaria, R., Kumar, K. H. and Krishnan, P. N. (2012). Exploring wild genetic resources of *Musa acuminata* Colla distributed in the humid forests of Southern Western ghats of peninsular India using ISSR markers. Plant Cell Reports 31(9): 1591–1601. doi:10.1007/s00299-012-1273-5.

Peterson, D. G., Schulze, S. R., Sciara, E. B., Lee, S. A., Bowers, J. E., Nagel, A., Jiang, N., Tibbitts, D. C., Wessler, S. R. and Paterson, A. H. (2002). Integration of cot analysis, DNA cloning, and high-throughput sequencing facilitates genome characterization and gene discovery. Genome Research 12(5): 795–807. doi:10.1101/gr.226102.

Pillay, M. and Adeleke, M. T. V. (2001). Silver staining of *Musa* L. chromosomes. Cytologia 66(1): 33–37.

Pillay, M. and Tripathi, L. (2006). Banana: an overview or breeding and genomics research in *Musa*. In: Kole, C. (Ed.), Genome Mapping and Molecular Breeding in Plants (vol. 4), Fruits and Nuts. Springer-Verlag, Heidelberg, Germany, 282–301. doi:10.1007/978-3-540-34533-6_15.

Preedy, K. F. and Hackett, C. A. (2016). A rapid marker ordering approach for high- density genetic linkage maps in experimental autotetraploid populations using multidimensional scaling. TAG. Theoretical and Applied Genetics. Theoretische und Angewandte Genetik 129(11): 2117–2132. doi:10.1007/s00122-016-2761-8.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Raboin, L., Carreel, F., Noyer, J., Baurens, F., Horry, J., Bakry, F., Montcel, H. T. D., Ganry, J., Lanaud, C. and Lago-da, P. J. L. (2005). Diploid Ancestors of Triploid Export Banana Cultivars: molecular Identification of 2n Restitution Gamete Donors and N Gamete Donors. Molecular Breeding 16(4): 333–341. doi:10.1007/s11032-005-2452-7.

Roux, N., Toloza, A., Radecki, Z., Zapata-Arias, F. J. and Doležel, J. (2003). Rapid detection of aneuploidy in *Musa* using flow cytometry. Plant Cell Reports 21(5): 483–490. doi:10.1007/s00299-002-0512-6.

Šafář, J., Noa-Carrazana, J. C., Vrána, J., Bartoš, J., Alkhimova, O., Sabau, X., Simková, H., Lheureux, F., Caruana, M. L., Dolezel, J. and Piffanelli, P. (2004). Creation of a BAC resource to study the structure and evolution of the banana (*Musa balbisiana*) genome. Genome 47(6): 1182–1191. doi:10.1139/g04-062.

Sardos, J., Perrier, X., Doležel, J., Hřibova, E., Christelová, P., Van den Houwe, I., Kilian, A. and Roux, N. (2016). DArT whole genome profiling provides insights on the evolution and taxonomy of edible banana (*Musa* spp.). Annals of Botany 118(7): 1269–1278. doi:10.1093/aob/mcw170.

Sass, J. E. (1964). Botanical Microtechnique. Iowa State University Press, Ames, IA. Shepherd, K. (1999). Cytogenetics of the Genus Musa. INIBAP, Montpellier, France. Silva, A. V. C., Nascimento, A. L. S., Vitória, M. F., Rabbani, A. R. C., Soares, A. N. R. and Lédo, A. S. (2017). Diver-

sity and genetic stability in banana genotypes in a breeding program using inter simple sequence repeats (ISSR) markers. Genetics and Molecular Research: GMR 16(1). doi:10.4238/gmr16019402.

Simmonds, N. W. and Dodds, K. S. (1949). Meiosis in seeded diploids of *Musa*. Journal of Genetics 49(3): 221–225.

Simmonds, N. W. and Shepherd, K. (1955). The taxonomy and origins of the cultivated bananas. Journal of the Linnean Society of London, Botany 55(359): 302–312.

Sybenga, J. (1975). Meiotic Configurations. Springer-Verlag. ISBN 978-3-642-80960-6. Sybenga,

J. (1992). Cytogenetics in Plant Breeding. Springer Verlag. ISBN 978-3-642-84083-8.

Valárik, M., Simkova, H., Hřibova, E., Šafář, J., Doleželová, M. and Doležel, J. (2002). Isolation, characterization and chromosome localization of repetitive DNA sequences in bananas (*Musa* spp.). Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology 10(2): 89–100.

Van Os, H., Stam, P., Visser, R. G. and van Eck, H. J. (2005). SMOOTH: a statistical method for successful removal of genotyping errors from high-density genetic linkage data. Theoretical and Applied Genetics. 112(1): 187–194.

Vilarinhos, A. D., Piffanelli, P., Lagoda, P., Thibivilliers, S., Sabau, X., Carreel, F. and D'Hont, A. (2003). Construction and characterization of a bacterial artificial chromosome Library of Banana (Musa acuminata Colla). Theoretical and Applied Genetics. 106(6): 1102–1106. doi:10.1007/s00122-002-1155-2.

Wang, Z., Miao, H., Liu, J., Xu, B., Yao, X., Xu, C., Zhao, S., Fang, X., Jia, C., Wang, J., Zhang, J., Li, J., Xu, Y., Wang, J., Ma, W., Wu, Z., Yu, L., Yang, Y., Liu, C., Guo, Y., Sun, S., Baurens, F. C., Martin, G., Salmon, F., Garsmeur, O., Yahiaoui, N., Hervouet, C., Rouard, M., Laboureau, N., Habas, R., Ricci, S., Peng, M., Guo, A., Xie, J., Li, Y., Ding, Z., Yan, Y., Tie, W., D'Hont, A., Hu, W. and Jin, Z. (2019). Musa balbisiana genome reveals subgenome evolution and functional divergence. Nature Plants 5(8): 810– 821. doi:10.1038/s41477-019-0452-6.

Wilson, G. B. (1946). Cytological studies in the Musae. II. Meiosis in some diploid clones. Genetics 31(5): 475–482.

Zhang, J., Yu, C., Krishnaswamy, L. and Peterson, T. (2011). Transposable elements as catalysts for chromosome rearrangements. In: Birchler, J. A. (Ed.), Plant Chromosome Engineering: Molecular and Protocols, Methods in Molecular Biology (Vol. 701). Humana Press, Totowa, NJ, 315–326. doi:10.1007/978-1-61737-957-4_18.

Chapter 5.

Genetic mapping of Fusarium wilt resistance in a wild banana *Musa acuminata* ssp. *malaccensis* accession

Ahmad F, Martawi NM, Poerba YS, Jong H De, Schouten H & Kema GHJ

Abstract

Banana is an important fruit and food crop, but is threatened by Fusarium wilt, one of the most devastating soil-borne fungal diseases. Only host resistance facilitates banana cultivation in infested soils around the world, but the genetic basis of Fusarium wilt of banana (FWB) is unknown. We selfed a heterozygous wild banana accession *Musa acuminata* ssp. *malaccensis* (Mam, AA, 2n = 22) to generate a mapping population and to investigate the inheritance of resistance to Race 1 and tropical race 4 (TR4) that cause FWB. Phenotyping (N = 217) revealed segregation for resistance, and genotyping by sequencing resulted in 2802 high-quality single-nucleotide polymorphic markers (SNPs) that were used for genetic mapping. Combined analyses of these data showed that a single dominant resistance locus controls resistance to Race 1 and maps near the distal part of chromosome 10. Recombinants, together with the position of the putative resistance gene, were further analysed using graphical genotyping, which retrieved markers flanking a 360 kb genetic region that associates with Race 1 resistance. The region contains 165 putative genes on the reference genome, including 19 leucine-rich repeat receptor-like kinase-like genes. At the same position and phase, we also identified a QTL for TR4 resistance, showing that the locus for resistance against Race 1 provided partial resistance to TR4. However, this effect was far less significant and hence not included in the mapping. These data support the breeding of new banana varieties with resistance to Fusarium wilt.

Introduction

Fusarium wilt is the most devastating disease in banana culture and destroyed large plantations in the tropical countries of South America since the outbreak in the early 1900s (Ploetz 2005, 2015). The causal agents of the disease are a suite of Fusarium species (Maryani et al. 2019) that previously were classified as *Fusarium oxysporum* f. sp. *cubense* (Foc). They invade banana roots and subsequently colonize and occlude the vascular system which leads to severe wilting that eventually kills the plant (De Ascensao and Dubery 2000). According to the compatibility of the fungus with groups of banana cultivars, Fusarium strains are classified into three races. Race 1 is known for its devastation of large areas of 'Gros Michel' in 1950s. Race 2 is compatible with the cooking banana subgroup Bluggoe (ABB). The third type, Race 4, is usually divided into sub-tropical race 4 (STR4) that infects banana under abiotic stress and tropical race 4 (TR4), which devastates Cavendish plantations around the world, but also affects many other banana varieties (García-Bastidas 2019). TR4 was recently identified as a new species named *F. odoratissimum* (Maryani et al. 2019), which most likely originates from the Indonesian archipelago from where it disseminated globally (Maymon et al. 2020; Özarslandan and Akgül 2020; García-Bastidas et al. 2019a; Damodaran et al. 2019; Chittarath et al. 2018; Maymon et al. 2018; Ordoñez et al. 2015, Ordoñez et al. 2016; García-Bastidas et al. 2014).

Fusarium spp. causing Fusarium wilt of banana (FWB) are soil-borne fungi with the ability to produce chlamydospores. Such inoculum can survive for more than 30 years either as spores or by hiding as endophytes in weeds (Salacinas 2019), which makes disease control extremely challenging (Blomme et al. 2011; Ploetz 2005). Planting diseasefree plantlets from in vitro culture may initially keep banana plants healthy, but after a few production cycles plants become affected and yields will dramatically drop due to massive fungal infestations (Bubici et al. 2019). Over the years, many products and conditions were trialled, including biocontrol applications with Trichoderma spp. or nonpathogenic *F. oxysporum*, in an attempt to control FWB (Dita et al. 2018; Chaves et al. 2016; Soluri 2005). However, none of these adequately managed the disease, except for the excelling resistance in Cavendish clones that were eventually embraced by the industry and quenched the Race 1 driven epidemic in 'Gros Michel' (Bubici et al. 2019). Despite the success of Cavendish, the genetic basis ofits resistance to Race 1 strains was never unveiled. Various wild and cultivated bananas were described conferring different levels of resistance to Race 1 and TR4 (García-Bastidas 2019). Interestingly, Musa acuminata ssp. *malaccensis* (Mam) 'Pahang', ssp. *burmannicoides* (Calcutta 4) and *M. itinerans* carry resistance to TR4 (D'Hont et al. 2012; Zuo et al. 2018; Zhang et al. 2018, 2019; García-Bastidas 2019), whereas the cultivar 'Rose' and 'Tuu Gia'

(ITC 0610) are examples of resistant edible bananas (Houbin et al. 2004; Zuo et al. 2018; García-Bastidas 2019). However, despite their exquisite value, such germplasm remains largely untapped in contemporary breeding programs.

In recent studies, Peraza-Echeverria et al. (2008, 2009) used PCR-based analyses of RNA data to identify resistance gene analogues (RGAs) from two resistant and two susceptible accessions of Mam. They identified five RGAs of which three were associated with TR4 resistance. Eventually, RGA2 was transformed to the susceptible Cavendish clone 'Grand Naine' which turned it resistant to TR4 (Dale et al. 2017). Actually, RGA2 is present in Cavendish, but the gene is expressed 10 × lower than in the transgenic line, indicating that the resistance of the latter is due to a dosage effect. Since resistance to Race 1 is an absolute prerequisite for any breeding program, we set out a strategy to identify the responsible genes, which enables marker-assisted breeding and avoids lengthy and expensive phenotyping assays. Genetic mapping is a powerful tool for identifying the locus of interest and the distances between genes on linkage groups. In banana, genetic maps have been constructed from segregation populations of selfed or cross-pollinated heterozygotes. Faure et al. (1993) built such a genetic map from an F2 population of two wild *M.* acuminata varieties based on RFLP, isozyme and RAPD markers, demonstrating 36% segregation distortion of the male-specific alleles. The linkage maps reported by Kayat et al. (2009) were based on two F1 populations of selfed Mam using AFLP, STMS and RAPD markers, resulting in 32 and 37 linkage groups, many more than the expected 11 groups of diploid banana (2n = 2x = 22). Kilian et al. (2012) generated a large number of markers by Diversity Array Technology (DArT) that were used by Hippolyte et al. (2010) in combination with SSR markers to build a linkage map from F1 progeny of a hybrid between *M. acuminata* 'Borneo' and 'Pisang Lilin' that resulted in the expected 11 linkage groups. Here, we developed a segregating population from selfing of a heterozygous Mam that was genotyped using the DArTseq genotyping-by-sequencing (GBS) method and was phenotyped for resistance to Race 1 or TR4. The phenotypic information was combined with SNP markers derived from the DArTseq analysis to build a linkage map. We subsequently mapped windows for Race 1 and TR4 resistance on chromosome 10. The identified flanking markers can now—for the first time—be used for marker-assisted breeding to FWB resistance.

Materials and methods
Creating a segregating population
A segregating population was made by selfing a diploid heterozygous accession of a wild *Musa acuminata* Colla var. *malaccensis* (Ridl.) Nasution (Mam) that originated from Sumatra, Indonesia. It was chosen for its resistance to TR4 in a greenhouse

bioassay and in the field (data not shown). The selfing was performed at the Research Center for Biology, Indonesian Institute of Sciences (LIPI) in the field in 2014 by hand pollination. The pollinated flowers were bagged immediately using fine insect screens to prevent cross-pollination, and fruits were harvested 12–15 weeks after pollination. Subsequently, the seeds were collected in a week when the fruits were ripe. A total of 8,077 seeds from 231 pollinated flowers was harvested from the selfed parent. The majority of the seed (90%) had black hard skins and were full of endosperm, while the remainder had brown or shrunken seed coats. The seeds were transferred to Wageningen University and Research (WUR) for embryo rescue and in vitro propagation. Prior to embryo rescue, the seeds were soaked in 96% ethanol for a minute and in 20% hypochlorite for 20 min, washed in sterile water, soaked in 10% hypochlorite for 10 min and finally washed in sterile water. Subsequently, we performed a priming by soaking the seeds in 10 ppm gibberellic acid (GA3) for three days to initiate shooting of the embryo (Arun et al., 2013). The embryos were taken out of the seeds under sterile conditions, transferred to Murashige & Skoog medium (MS) (Murashige & Skoog, 1962) with 2 ppm benzylaminopurine (BAP), 1 ppm biotin, 0.1 g/l myo-inositol and 3 g/l gelrite before autoclaving and placed in the dark to induce root elongation. After two to four weeks, the shooting embryos were exposed to light for shoot development. In total, 718 embryos were rescued, but also suffered from a bacterial contamination, whereas other embryos did not shoot (data not shown). Eventually, 255 embryos survived and developed into plants for disease assays by separating axillary shoots by subculturing on MS medium with 2 ppm BAP. Two to five times subcultures were taken to obtain 20–25 shoots per genotype. During the last subculturing, we transferred the shoots to MS medium without hormones to induce root formation. Subsequently, plants were transferred from tissue culture to the greenhouse in individual pots with soil (5% Swedish sphagnum peat, 41% grinding clay granules, 5% garden peat, 4% beam structure, 33% steamed compost and 12% PG-Mix 15-10-20) and maintained for 2 weeks under controlled conditions (100% humidity and 28 ± 2°C) to acclimatize. Subsequently, the plants were kept at 75–85% humidity and 28 ± 2°C for 2 months prior to inoculation.

Disease assays
The disease tests were performed from August 2015 until October 2016, using a Race 1 isolate of unknown vegetative compatibility (Ordóñez, 2018) originating from Brazil (coded as Foc.CNPMF.R1) that was recovered from 'Maçã' banana (ABB, Silk subgroup) and of F. odoratissimum representing tropical race 4 (TR4; isolate II-5 originating from Indonesia) (Dita et al. 2011; Maryani et al. 2019), both maintained at the Wageningen University and Research (WUR) collection. The inoculum preparation, disease assays and disease evaluation were performed according to García-Bastidas et al. (2019b). We

screened a progeny of 225 genotypes, with five replicates per genotype along with the Mam parent and 'Gros Michel' (AAA) as susceptible control for Race 1 and 'Grand Naine' as susceptible control for TR4. Seven weeks after inoculation, the disease symptoms of the leaves and the rhizomes were evaluated.

Heritability of resistance

The heritability (h^2) of resistance to Race 1 was estimated by dividing the genotypic variance (σ_g^2) by the sum of the genotypic and environmental variances ($h^2=\sigma_g^2/(\sigma_g^2+ \sigma_e^2)$) (Allard 1960). The genotypic variance (σ_g^2) and the denominator ($\sigma_g^2+ \sigma_e^2$) were estimated using an ANOVA.

Genotyping

We collected leaf samples from the segregating population and reference genotypes from either the leaves from plantlets or from the tissue culture plants or from the cigar leaves in the greenhouse. These were lyophilized and used for DNA isolation using the Wizard® Magnetic DNA Purification System for Food kit from (Promega, Madison, USA) following the manufacturer's instructions. The DNA concentration was quantified using the Quant-iT™ PicoGreen dsDNA Assay Kit (Life Technologies, USA) according to the manufacturer's instructions. Subsequently, the quantities were calculated using Tecan Infinite® M200 PRO monochromator (Tecan, Männedorf, Switzerland) using Icontrol 107 software (US, Morrisville, NC) and the quality was checked by electrophoresis on 1% agarose gels. The DNA was sent to Diversity Arrays Technology Pty Ltd, Australia, for scoring SNP markers, using the DArTseq platform (http://www. diversityarrays .com/). Based on the DNA sequences flanking the SNPs, the markers were putatively positioned on the genome assembly of Mam DH 'Pahang', version 2 (http://banana-genome-hub.southgreen.fr/organism/Musa/acuminata). Eventually, 217 progeny individuals and their parent were genotyped. The very high number of obtained SNP markers allowed us to apply a very stringent filtering, using as criteria: 1. replicate value = 1; 2. sequences should hit only one position on the reference genome; 3. the polymorphism information content > 0.3; 4. markers should be based on more than ten calls per allele; and 5. we only considered markers that were heterozygous in the parent, and therefore could segregate in the population. We checked the physical position of the filtered, segregating markers on the reference genome, for evaluation of possible homozygous regions in the self-pollinated parent. This subset of qualified markers was used for mapping the Race 1 and TR4 resistance.

Mapping
Genetic linkage maps were constructed using JoinMap® 6 software with a 'F2' population type, thus regarding the heterozygous parent as an F1 from a cross between two homozygous grandparents. Markers were assigned to homologous chromosomes, taking linkage or repulsion into consideration. Later, the refined linkage maps were the basis for the mapping analysis. We used MapQTL® 6 to find markers associated with the resistance. The LOD thresholds for significance of quantitative trait locus (QTL) were calculated with a permutation test, using the 95% confidence level and 1000 permutations. Markers above the threshold were supposed to be markers associated with the resistance. We selected genotypes that showed recombination in this region. Subsequently, these recombinants were analysed in detail, applying graphical genotyping, for genetic mapping. This yielded markers flanking the genetic region that contains the resistance gene.

Results

Disease assays
The histogram of the disease scores of leaves of the segregating population for Race 1 suggests a bimodal distribution of a susceptible group and a relatively resistant group of genotypes, segregating in a 1:1 ratio (Figure 1a). The 1:1 segregating indicates a monogenic dominant resistance gene that could be heterozygously present in the parent. However, the bimodal distribution was not present for the rhizome symptoms of the same genotypes upon inoculation with Race 1 (Figure 1b). In contrast to the leaf scores of Race 1, the histograms for disease scores upon inoculation with TR4 showed no bimodal segregation, but rather a normal distribution, both for the leaves and the rhizomes (Figure 1c, d). Moreover, regarding the leaf symptoms the parent appeared to be more susceptible to TR4 than to Race 1 (Figure 1c), similar to the response of the susceptible Cavendish 'Grand Naine' TR4 check, which was significantly dissimilar from the rhizome scores, where the parent showed a low susceptibility for TR4 (Figure 1d).

Heritability of resistance
The heritability (h^2) was calculated to estimate the genetic impact of the resistance to either Race 1 or TR4 in the segregating population. For Race 1, h^2 equalled 0.70 and 0.69 for the leaves and rhizomes, respectively, thus showing very similar levels of heritability of resistance for both plant parts. For TR4, h^2 equalled 0.43 and 0.50 for leaves and rhizomes, respectively. This analysis indicates that the heritability in the resistance of the Mam accession is higher for Race 1 than for TR4, and the genetic impact of TR4 on rhizomes was higher than on the foliage.
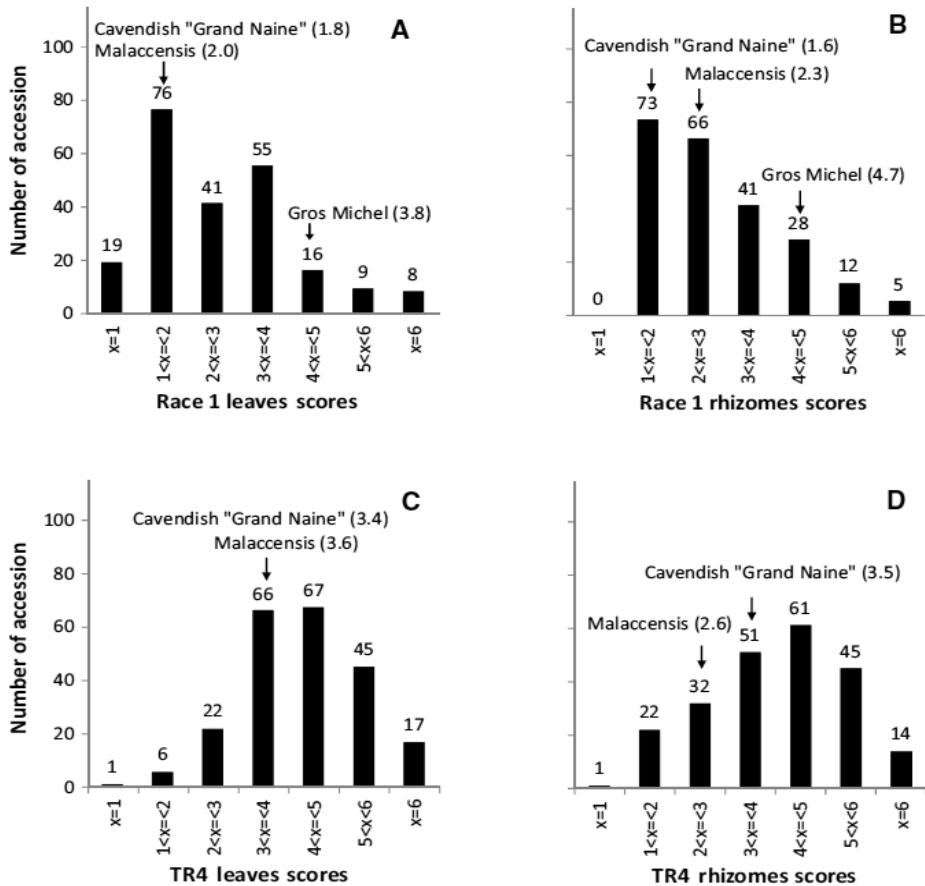
Figure 1 Frequency distribution of disease severity scores for Fusarium wilt in leaves and rhizomes upon inoculations with Race 1 and TR4. *Musa acuminata* ssp. *malaccensis* is the heterozygous parent of the segregating population. Cavendish 'Grand Naine' is the control for resistance to Race 1 and for susceptibility to TR4. 'Gros Michel' is the susceptible control for Race 1.

Genotyping

DArTseq provided 32,362 SNP markers for the segregating population. As the population consisted of 225 genotypes, with a few hundred recombinations per chromosome, we did not need that many markers. Therefore, we selected markers based on the above-mentioned very stringent quality criteria, which resulted in 2,802 SNP high-quality markers for genetic mapping. All chromosome arms harboured segregating markers, indicating that the parent has been heterozygous for the far majority of the genome

(Figure 2). As DArTseq used a methylation-sensitive restriction enzyme for making the libraries that were sequenced, the DArTseq markers were mainly located in active regions and far less abundant in the centromeric, methylated regions.
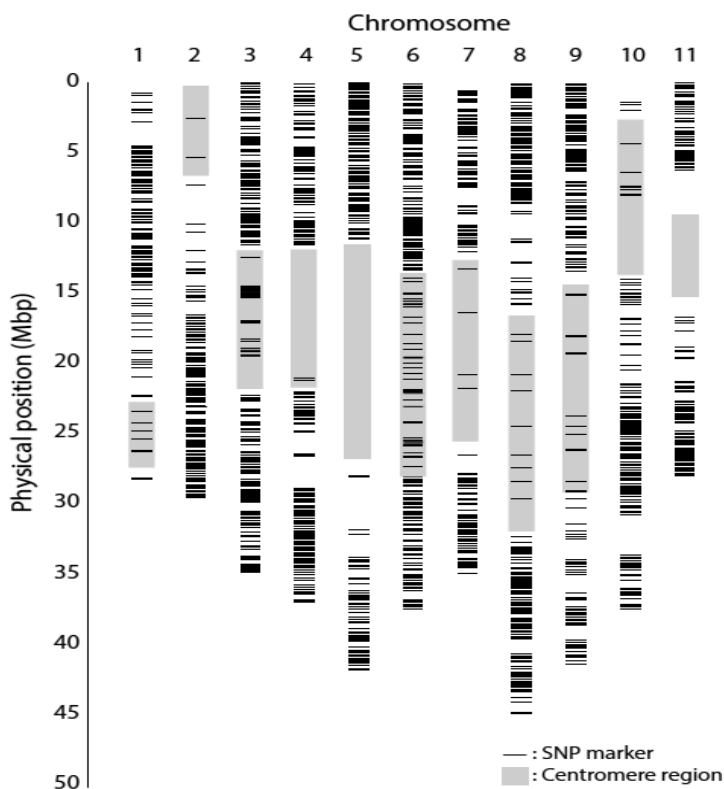
Figure 2 Distribution of SNP markers across the 11 chromosomes according to their physical positions on the reference genome of *Musa acuminata* ssp. *malaccensis* (Mam) DH 'Pahang', version 2, released January 2016 (http://banana-genome-hub.southgreen.fr/organism/Musa/acuminata). All chromosomes are covered by segregating markers, which means all chromosomes are heterozygous in the Mam parent that was self-pollinated in this study, thus allowing genetic mapping of all chromosomes in the segregating population.

We constructed linkage maps according to the selected markers, providing 11 linkage groups (Supplementary Table 1). Subsequently, we performed QTL mapping, using the linkage groups and the phenotypic disease scores for Race 1 and TR4. Only one linkage group (Chromosome 10) showed a significant QTL at the distal part (Figure 3).

Remarkably, this QTL was found for both strains (Race 1 and TR4) and both plant parts (leaves, rhizome). The highest LOD scores were obtained for Race 1 resistance, i.e. 40.9 for the rhizomes and 20.2 for the leaves. For TR4, the maximum LOD scores were far lower, i.e. 5.8 and 5.9 for leaves and rhizomes, respectively.
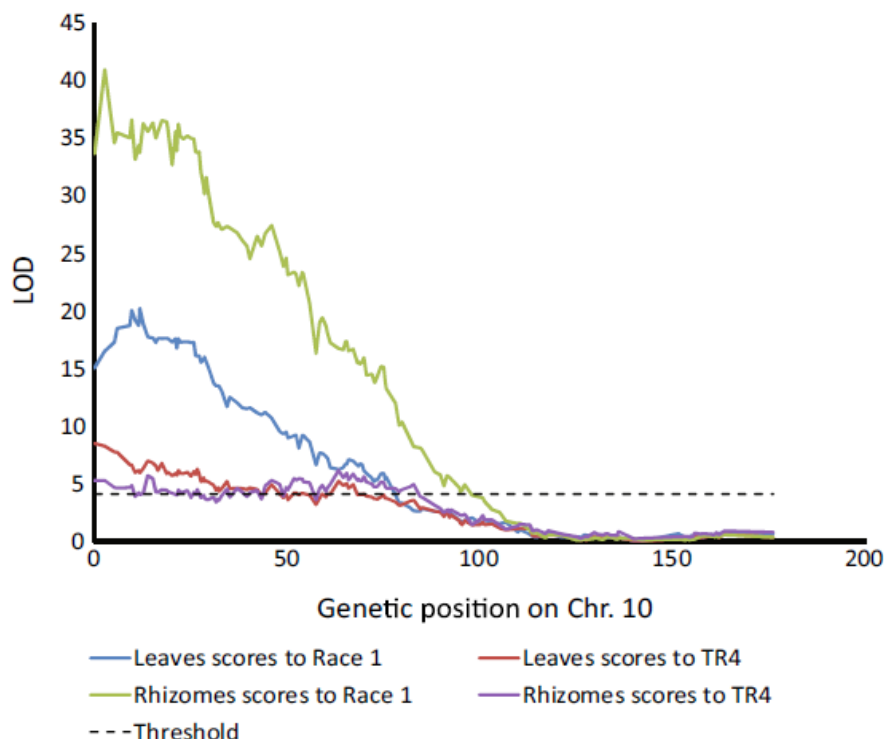
Figure 3 Interval mapping for Fusarium wilt resistance on chromosome 10 to Race 1 and TR4 on *Musa acuminata* ssp. *malaccensis* chromosomes according to leaf and rhizome severity scores. Horizontal lines represent LOD threshold (4.1) value for $\alpha = 0.05$ after 1000 permutations. Markers above the threshold line indicate a significant association with disease score.

As the combination of Race 1 inoculation and rhizome observations provided the most significant QTL, we focussed on this combination for fine mapping. For this purpose, we selected 13 genotypes that showed recombination in the QTL region and could be clearly designated as either susceptible or resistant regarding the rhizome symptoms. We used 106 SNP markers for depicting the recombination in these genotypes (Figure 4). According to the physical positions of the markers on the reference genome, the Race 1 resistance is located between 0 and 4.35 Mbp at the distal tip of chromosome 10.

This region contains 165 putative genes, and 19 of these putative genes were annotated as leucinerich repeat (LRR) receptor-like kinase-like genes (Supplementary Table 2).
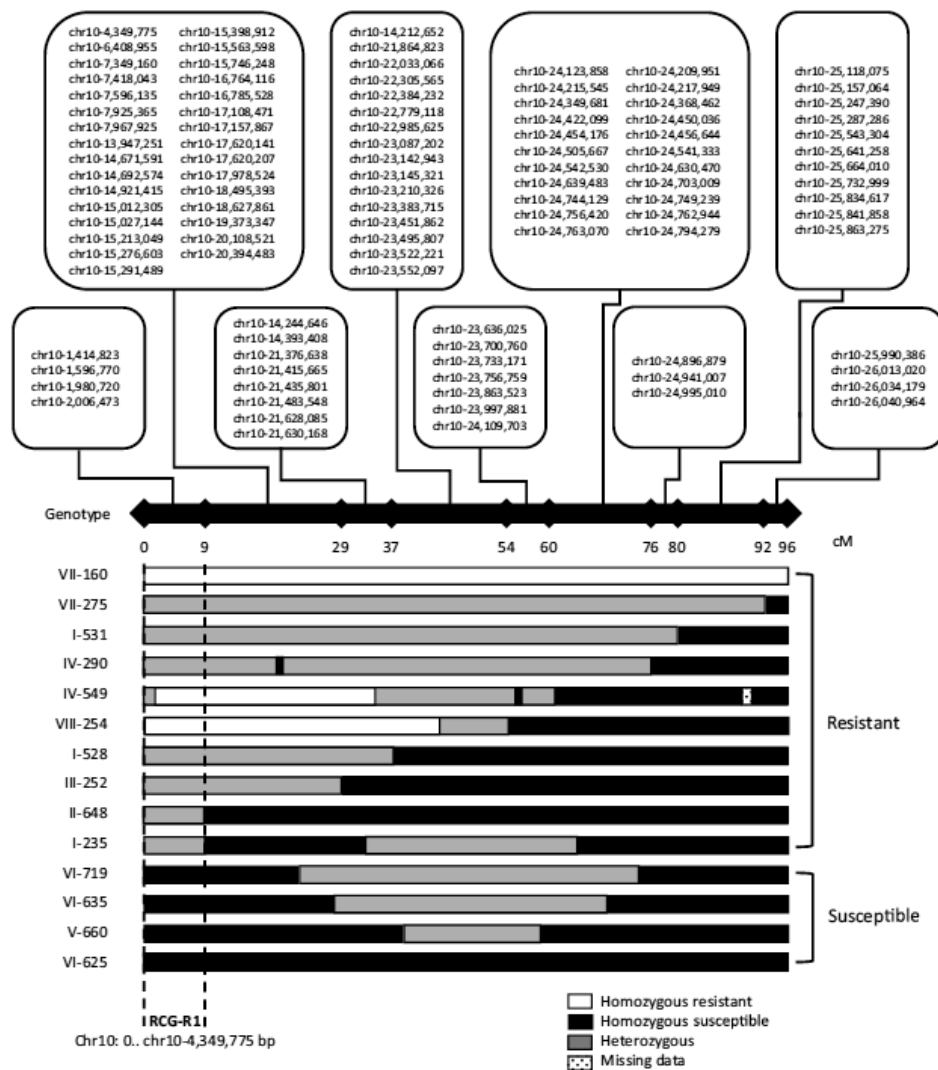
Figure 4 Graphical genotyping of Fusarium wilt resistance to Race 1 on chromosome 10, using rhizomes scores, and DArTseq SNP markers that were significantly associated with resistance. By sorting the selected progeny of selfed *Musa acuminata* ssp. *malaccensis* (Mam) according to the recombination events on chromosome 10 of Mam the Race 1 resistance could be located in the range of 0−4.3 Mbp at the distal end of the chromosome. The physical positions of the markers on the reference genome (http://banana-genome-hub.southgreen.fr/organism/Musa/acuminata) are shown in the groups of co-segregating markers at the top of the Figure.

As shown in Figure 3, in the same region as the QTL for resistance to Race 1, we found a QTL for TR4 resistance. However, that QTL was far less significant and did not allow fine mapping, due to a less clear phenotypic distinction of susceptible versus resistant genotypes. We estimated the effect of the presence of the resistance gene on the disease levels upon Race 1 inoculation (Table 1). Homozygous presence of the resistance gene reduced the rhizome symptoms from 4.6 to 2.0, so a reduction of 2.6 on a scale of six classes (Supplementary Figure 1). For leaves, the resistance QTL reduced the symptoms by 1.7. Table 1 shows that for Race 1 the resistance is dominant and also indicates the effect of the presence of the marker associated with Race 1 resistance on resistance to TR4. Although the effect on resistance to TR4 is less pronounced than for Race 1, this table suggests that the resistance for Race 1 might also slightly (reducing the leaves and rhizomes of TR4 symptoms by 1.4, when homozygously present) affect resistance to TR4.

Table 1 The mean of disease scores of progeny genotypes lacking the mapped resistance at the distal part of chromosome 10 (aa) or harbouring one copy (ab; heterozygous) or two copies (bb; homozygous) of this resistance QTL.

Genotypes	Rhizomes scores			Leaves scores		
	aa	ab	bb	aa	ab	bb
Race	2.0+0.7	2.6+1.0	4.6+1.4	2.3+1.0	2.3+1.0	4.0+1.4
TR4	3.4+1.3	4.0+1.2	4.8+1.3	3.4+1.1	4.0+0.9	4.8+0.9

aa: homozygous resistance

ab: heterozygous

bb: homozygous susceptible

Discussion

In this study, we used resistant Mam originating from Sumatra, Indonesia, which is related to the DH 'Pahang', which is resistant to TR4 and whose genome has been sequenced (D'Hont et al. 2012) and is used by many banana researchers as the reference genome for studying resistance against TR4. However, DH 'Pahang' has never been used for in mapping studies; thus, there is a possibility that this reference banana may not even have the resistance gene(s) identified in this study. Kayat et al. (2009) used progenies of Mam generated at the University of Malaya for AFLP-based mapping analysis. However, this was unsuccessful due to the limited progeny size. Later, Fraser-Smith et al. (2016) used a selfed population of TR4 resistant and susceptible Mam of unknown origin and concluded that the TR4 resistance is likely under the control of a single gene because the segregation ratio of the number of resistant (R) and susceptible

(S) genotypes was 4.67:1 in the first population and 4:1 in the second population. The Mam accession in the current study is resistant to TR4 as well as to Race 1.

In mapping or QTL analyses, molecular markers have been a standard for genetic or linkage maps analysis in many crops (Collard and Mackill 2007). Although genotypic data can utilize various molecular markers, we have chosen DArT markers as they offer a large and cost-effective set of markers from the genome (Kilian et al. 2012). In banana genetics and diversity analysis, DArT was already used and combined with SSR markers to build a linkage map from a segregating progeny of a hybrid between *M. acuminata* 'Borneo' and 'Pisang Lilin', which resulted in 11 linkage groups (Hippolyte et al. 2010). In this study, we used the SNP markers retrieved from the DArTseq analyses and revealed 11 linkage groups which were used for mapping Fusarium wilt resistance.

Our QTL mapping analysis enabled the mapping of Race 1 resistance on the distal part of chromosome 10 at 0 and 4.35 Mbp. This is the first report of the genetic basis of Race 1 resistance in banana. Previously, genetic analyses did not include mapping studies and hence conclusions were entirely based on the segregation based on phenotypic characters, which do not meet contemporary quality requirements (Vakili 1965; Ssali et al. 2013; Arinaitwe et al. 2019). Our result confirmed that the Race 1 resistance is inherited as a single gene, which accords with Vakili (1965) who used the diploid banana 'Pisang Lilin'. Our results also indicate that the resistance is controlled by a dominant gene, which contradicts with Ssali et al. (2013) who concluded that the gene was recessively inherited. They used a hybrid population of triploid dessert banana 'Sukali Ndizi' (AAB) and a resistant diploid banana 'TMB2X8075' (AA). However, this results in a mix of diploid, triploid and tetraploid progeny that complicates adequate analyses due to the complexity of the gametes and the pairing compatibility during the fertilization which may interfere with the proportion of the resistant and the susceptible genotypes (Dodds 1943; Shepherd 1999).

In contrast to Race 1 mapping, we were not able to fine map the TR4 resistance. Although some chromosome 10 markers indicated association with TR4 resistance, the LOD values are low and just significant. This indicated that the population in this study was not well suited for analysing TR4 inheritance in Mam. Nevertheless, we identified a positive interaction between the Race 1 and TR4 resistance loci in the coupling phase (Table 1). A similar observation was reported for two resistance genes (Ph-3 and Sw-5) in coupling phase in tomato resulting in progeny with resistance to tomato spotted wilt virus (TSWV) and *Phytophthora infestans* (the causal agent of late blight in potato and tomato) (Robbins et al. 2010).

As mentioned above, despite the success of Cavendish to manage Fusarium wilt in banana, the genetic basis for resistance to Race 1 remained unclear. Since Race 1 strains are globally disseminated (Ploetz 2015), resistance to Race 1 is a prerequisite for any new banana variety. Developing markers for resistance is, therefore, very valuable as it increases throughput and precision by avoiding cumbersome phenotyping assays, particularly under field conditions. The exceptional and durable nature of Race 1 resistance in Cavendish bananas still requires further study. Our subsequent analyses—19 out of 165 genes are leucine-rich repeat (LRR) receptor-like kinase-like genes—will lead to gene identification and validation, which then can be used to identify whether this gene is also present and expressed and henceforward explains the resistance to Race 1 in Cavendish. It is well established that mutations in LRR domains may predominantly be involved in regulating intramolecular interactions in defence mechanism, i.e. a mutation in the LRR coding region resulted in the loss of nematode resistance in *Nicotiana benthamiana* (Hwang and Williamson 2003). Usually, such mutants result from selection pressure on pathogen populations by deployed resistance factors (Bourguet et al. 2016; Flor 1971). Thus far, however, pathogenic strains on Cavendish bananas exclusively belong to *F. odoratissimum* (García-Bastidas et al. 2019a; Maryani 2018; Ordoñez 2018). This could indicate that the durable resistance to Race 1 in Cavendish is more complicated or that *Fusarium* spp. have alternative strategies to avoid selection pressure by, for example, growing as endophytes in weeds as was recently confirmed by Salacinas (2019). In previous studies, Ordonez et al. (2015) and Maryani et al. (2019) demonstrated that Race 1 comprises several species. Additional studies should reveal whether the identified resistance gene is also effective to the other species in the Race 1 complex.

Thus far, advanced genetic analyses have mostly focused on TR4 (Fraser-Smith et al. 2016; Kayat et al. 2004, 2009; Peraza-Echeverria et al. 2009). Kayat et al. (2009) used AFLP markers, and Fraser-Smith et al. (2016) used a selfed Mam pollination, but both groups never provided mapping data on TR4 resistance. After the preparative studies of Peraza-Echeverria et al. (2008), Dale et al. (2017) cloned RGA2 which provided resistance to TR4 that was validated by transferring it to the Cavendish cultivar 'Grand Nain'. This was successfully field-trialled for 3 years and remained free of disease, whereas all checks, including somaclonal variants, succumbed to TR4. Our study enables marker-assisted breeding for Race 1 resistance in banana, which is required in every production environment. Our study also confirms that Mam is fertile and can be a potential parent for breeding Fusarium wilt resistance. Such wild fertile bananas are necessary to create improved diploids that can be used to generate new triploids (Bakry et al. 2009). Additional genetic analyses of other banana accessions with resistance to TR4 should reveal the diversity for the sought-after TR4 resistance across a wide

panel of banana germplasm. Taken together, such data will support the breeding of new varieties to manage the threat of FWB to global banana production.

Author contribution statement

GHJK, HJ and YS designed and overviewed all experiments. FA and NMW conducted the entire experiment. FA, NMW, HJ and HS analyzed the data. FA, HJ, HS and GHJK wrote the paper.

References

Allard RW (1960) Principle of plant breeding. Wiley, New York Arinaitwe IK, Teo CH, Kayat F, Tumuhimbise R, Uwimana B, Kubiriba J, Swennen R, Harikrishna JA, Othman RY (2019) Evaluation of banana germplasm and genetic analysis of an F1 population for resistance to *Fusarium oxysporum* f. sp. *cubense* race 1. Euphytica 215:1–11

Arun K, Uma S, Saraswathi MS, Backiyarani S, Durai P, Post T, Road T (2013) Effects of whole seed priming on the in vitro germination of hybrid banana embryos (*Musa* spp.). Seed Sci Technol 41:439–451

Bakry F, Careel F, Jenny C, Horry JP (2009) Genetic improvement of banana. In: Jain S, Priyadarshan P (eds) Breeding plantation tree crops: tropical species. Springer, Berlin, pp 3–49

Blomme G, Eden-Green S, Mustaffa M, Nwauzoma B, Thangavelu R (2011) Major diseases of banana. Banana breeding. CRC Press, Boca Raton, pp 85–119

Bourguet D, Delmotte F, Franck P, Guillemaud T, Reboud X, Vacher C, Walker AS (2016) Combining selective pressures to enhance the durability of disease resistance genes. Front Plant Sci 7:1916

Bubici G, Kaushal M, Prigigallo MI, Cabanás CGL, Mercado-Blanco J (2019) Biological control agents against Fusarium wilt of banana. Front Microbiol 10:616

Chaves NP, Staver C, Dita MA (2016) Potential of *Trichoderma asperellum* for biocontrol of Fusarium wilt in banana. Acta Hort 1114:261–265

Chittarath K, Mostert D, Crew KS, Viljoen A, Kong G, Molina AB, Thomas JE (2018) First report of *Fusarium oxysporum* f. sp. *cubense* tropical race 4 (VCG 01213/16) associated with Cavendish bananas in Laos. Plant Dis 102:449

Collard BCY, Mackill DJ (2007) Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Phil Trans R Soc 363:557–572

D'Hont A, Denoeud F, Aury J, Baurens F, Carreel F, Garsmeur O, Noel B, Bocs S, Droc G, Rouard M, Da Silva C, Jabbari K, Cardi C, Poulain J, Souquet M, Labadie K, Jourda C, Lengellé J, Rodier-Goud M, Alberti A, Bernard M, Correa M, Ayyampalayam S, Mckain M, Leebens-Mack J, Burgess D, Freeling M, Mbéguié-A-Mbéguié D, Chabannes M, Wicker T, Panaud O, Barbosa J, Hribova E, Heslop-Harrison P, Habas R, Rivallan R, Francois P, Poiron C, Kilian A, Burthia D, Jenny C, Bakry F, Brown S, Guignon V, Kema G, Dita M, Waalwijk C, Joseph S, Dievart A, Jaillon O, Leclercq J, Argout X, Lyons E, Almeida A, Jeridi M, Dolezel J, Roux N, Risterucci A, Weissenbach J, Ruiz M, Glaszmann J, Quétier F, Yahiaoui N,

Wincker P (2012) The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. Nature 488:213–219

Dale J, James A, Paul J-Y, Khanna H, Smith M, Peraza-Echeverria S, Garcia-Bastidas F, Kema G, Waterhouse P, Mengersen K, Harding R (2017) Transgenic cavendish bananas with resistance to Fusarium wilt tropical race 4. Nat Commun 8:1496

Damodaran T, Mishra VK, Jha SK, Gopal R, Rajan S, Ahmed I (2019) First report of Fusarium wilt in banana caused by *Fusarium oxysporum* f. sp. *cubense* tropical race 4 in India. Plant Dis 103:1022

De Ascensao ARDCF, Dubery IA (2000) biochemistry and cell biology panama disease: cell wall reinforcement in banana roots in response to elicitors from *Fusarium oxysporum* f. sp. *cubense* Race Four. Phytopathology 90:1173–1180

Dita MA, Waalwijk C, Paiva LV, Souza MT Jr, Kema GHJ (2011) A greenhouse bioassay for the *Fusarium oxysporum* f. sp. *cubense* X 'Grand Naine' (*Musa*, AAA, Cavendish subgroup) interaction. Acta Hort 897:377–380

Dita M, Barquero M, Heck D, Mizubuti ESG, Staver CP (2018) Fusarium wilt of banana: current knowledge on epidemiology and research needs toward sustainable disease management. Front Plant Sci 9:1468. https ://doi.org/10.3389/fpls.2018.01468

Dodds KS (1943) Geentical and cytological studies of *Musa*, V. Certain edible diploids. Genetics 45:113–139

Faure S, Noyer JL, Horry JP, Bakry F, Lanaud C, De Gonzfilez, León D (1993) A molecular marker-based linkage map of diploid bananas (*Musa acuminata*). Theor Appl Genet 87:517–528

Flor HH (1971) Current status of the gene-for-gene concept. Annu Rev Phytopathol 9:275–296

Fraser-Smith S, Czislowski E, Daly A, Meldrum R, Hamill S, Smith M, Aitken EAB (2016) Single gene resistance to *Fusarium oxysporum* f. sp. cubense Race 4 in the wild banana *Musa acuminata* subsp. *malaccensis*. Acta Hort 1114:95–100

García-Bastidas FA (2019) Panama disease in banana: spread, sceen, genes. Dissertation, Wageningen University & Research

García-Bastidas F, Ordóñez N, Konkol J, Al-Qasim M, Naser Z, Abdelwali M, Salem N, Waalwijk C, Ploetz RC, Kema GHJ (2014) First report of *Fusarium oxysporum* f. sp. cubense tropical race 4 associated with Panama disease of banana outside Southeast Asia. Plant Dis 98:694

García-Bastidas FA, Quintero-Vargas JC, Ayala-Vasquez M, Schermer T, Seidl MF, Santos-Paiva M, Noguera AM, AguileraGalvez C, Wittenberg A, Hofstede R, Sørensen A, Kema GHJ (2019a) First report of Fusarium wilt tropical race 4 in Cavendish bananas caused by *Fusarium odoratissi-*

mum in Colombia. Plant Dis 107:994

García-Bastidas FA, Van der Veen AJT, Nakasato-Tagami G, Meijer HJG, Arango-Isaza RE, Kema GHJ (2019b) An improved phenotyping protocol for Panama disease in banana. Front Plant Sci 10:1–12

Hippolyte I, Bakry F, Seguin M, Gardes L, Rivallan R, Risterucci A-M, Jenny C, Perrier X, Carreel F, Argout X, Piffanelli P, Khan IA, Miller RN, Pappas GJ, Mbéguié-A-Mbéguié D, Matsumoto T, De Bernardinis V, Huttner E, Kilian A, Baurens F-C, D'Hont A, Cote F, Courtois B, Glaszmann J-C (2010) A saturated SSR/DArT linkage map of *Musa acuminata* addressing genome rearrangements among bananas. BMC Plant Biol 10:65

Houbin C, Chunxiang X, Qirui F, Guibing H, Jianguo L, Zehuai W, Molina AB (2004) Screening of banana clones for resistance to fusarium wilt in China. In: Proceeding of the 3rd BAPNET meeting: advancing banana and plantain R&D in Asia and the Pacific, pp 165–174

Hwang CF, Williamson VM (2003) Leucine-rich repeat-mediated intramolecular interactions in nematode recognition and cell death signaling by the tomato resistance protein Mi. Plant J 34:585–593

Kayat F, Javed MA, Wah HY, Othman RY, Berhad UP, Intan T, Michel G (2004) Identification of molecular markers for disease resistance genes to *Fusarium oxysporum* f. sp. *cubense*

in *Musa acuminata* ssp. *malaccensis* for marker assisted selection (MAS). In: The 4th annual seminar of national science fellowship 2004:40–44

Kayat F, Bonar N, Waugh R, Rajinder S, Rahimah AR, Rashid AR, Othman RY (2009) Development of a genetic linkage map for genes associated with resistance and susceptibility to *Fusarium oxysporum* f. sp. *cubense* from an F1 hybrid population of *Musa acuminata* ssp. *malaccensis*. Proc Int ISHS Acta Horticult 828:1–8

Kilian A, Wenzl P, Huttner E, Carling J, Xia L, Blois H, Caig V, Heller-Uszynska K, Jaccoud D, Hopper C, Aschenbrenner-Kilian M, Evers M, Peng K, Cayla C, Hok P, Uszynski G (2012) Diversity arrays technology: a generic genome profiling technology on open platforms. Humana Press, Totowa, pp 67–89

Maryani N (2018) A complex relationship banana & Fusarium wilt in Indonesia. Dissertation, Wageningen University & Research

Maryani N, Lombard L, Poerba YS, Subandiyah S, Crous PW, Kema GHJ (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

Maymon M, Shpatz U, Harel YM, Levy E, Elkind G, Teverovsky E, Gofman R, Haberman A, Zemorski R, Ezra N, Levi Y, Or G, Galpaz N, Israeli Y, Freeman S (2018) First report of

Fusarium oxysporum f. sp. *cubense* Tropical Race 4 causing Fusarium wilt of Cavendish bananas in Israel. Plant Dis 102:2655

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plant 15:473–497

Ordoñez N (2018) A global genetic diversity analysis of *Fusarium oxysporum* f. sp. *cubense* the panama disease pathogen of banana. Dissertation, Wageningen University & Research

Ordoñez N, Seidl MF, Waalwijk C, Drenth A, Kilian A, Thomma BPHJ, Ploetz RC, Kema GHJ (2015) Worse comes to worst: bananas and Panama disease—when plant and pathogen clones meet bananas: their origin and global rollout. PLoS Pathog 11:e1005197

Ordoñez N, García-Bastidas F, Laghari HB, Akkary MY, Harfouche EN, Al-Awar BN, Kema GHJ (2016) First report of *Fusarium oxysporum* f. sp. *cubense* tropical race 4 causing panama disease in cavendish bananas in Pakistan and Lebanon. Plant Dis 100:209

Özarslandan M, Akgül DS (2020) First report of *Fusarium oxysporum* f. sp. *cubense* Race 4 causing Fusarium wilt disease of banana in Turkey. Plant Dis. https://doi.org/10.1094/PDIS-09-19-1881-PDN

Peraza-Echeverria S, Dale JL, Harding RM, Smith MK, Collet C (2008) Characterization of disease resistance gene candidates of the nucleotide binding site (NBS) type from banana and correlation of a transcriptional polymorphism with resistance to *Fusarium oxysporum* f.sp. *cubense* race 4. Mol Breed 22:565–579

Peraza-Echeverria S, Dale JL, Harding RM, Collet C (2009) Molecular cloning and in silico analysis of potential Fusarium resistance genes in banana. Mol Breed 23:431–443

Ploetz RC (2005) Panama disease: an old nemesis rears its ugly head, Part 1. The begining of the banana export trades. Plant Health Prog. 6:18

Ploetz RC (2015) Management of Fusarium wilt of banana: a review with special reference to tropical race 4. Crop Prot 73:7–15

Robbins MD, Masud MAT, Gardner RG, River M, Francis DM (2010) Marker-assisted selection for coupling wilt virus and *Phytophthora infestans* (Late Blight) in tomato. HortScience 45:1424–1428

Salacinas (2019) Spot on: managing panama disease of bananas in the Philippines. Dissertation, Wageningen University & Research

Shepherd K (1999) Cytogenetics of the genus *Musa*. International network for the improvement of banana and plantain, Montpellier, France

Soluri J (2005) Banana cultures. University of Texas Press, Austin Ssali TR, Kiggundu A, Lorenzen J, Karamura E, Tushemereirwe W, Viljoen A (2013) Inheritance of resistance to *Fusarium oxysporum* f. sp. *cubense* race 1

in bananas. Euphytica 194:425–430

Vakili NG (1965) Fusarium wilt resistance in seedlings and mature plants of *Musa* species. Phytopathology 55:135–140

Zhang L, Yuan TL, Wang YZ, Zhang D, Bai TT, Xu ST, Wang YY, Tang WH, Zheng SJ (2018) Identification and evaluation of resistance to *Fusarium oxysporum* f. sp. *cubense* Tropical Race 4 in *Musa acuminata* Pahang. Euphytica 214:106

Zhang L, Cenci A, Rouard M, Zhang D, Wang YY, Tang WH, Zheng SJ (2019) Transcriptomic analysis of resistant and susceptible banana corms in response to infection by *Fusarium oxysporum* f. sp. *cubense* tropical race 4. Sci Rep 9:8199

Zuo C, Deng G, Li B, Huo H, Li C, Hu C, Kuang R, Yang Q, Dong T, Sheng O, Yi G (2018) Germplasm screening of *Musa* spp. for resistance to *Fusarium oxysporum* f. sp. *cubense* tropical race 4 (Foc TR4). Eur J Plant Pathol 4:1–12
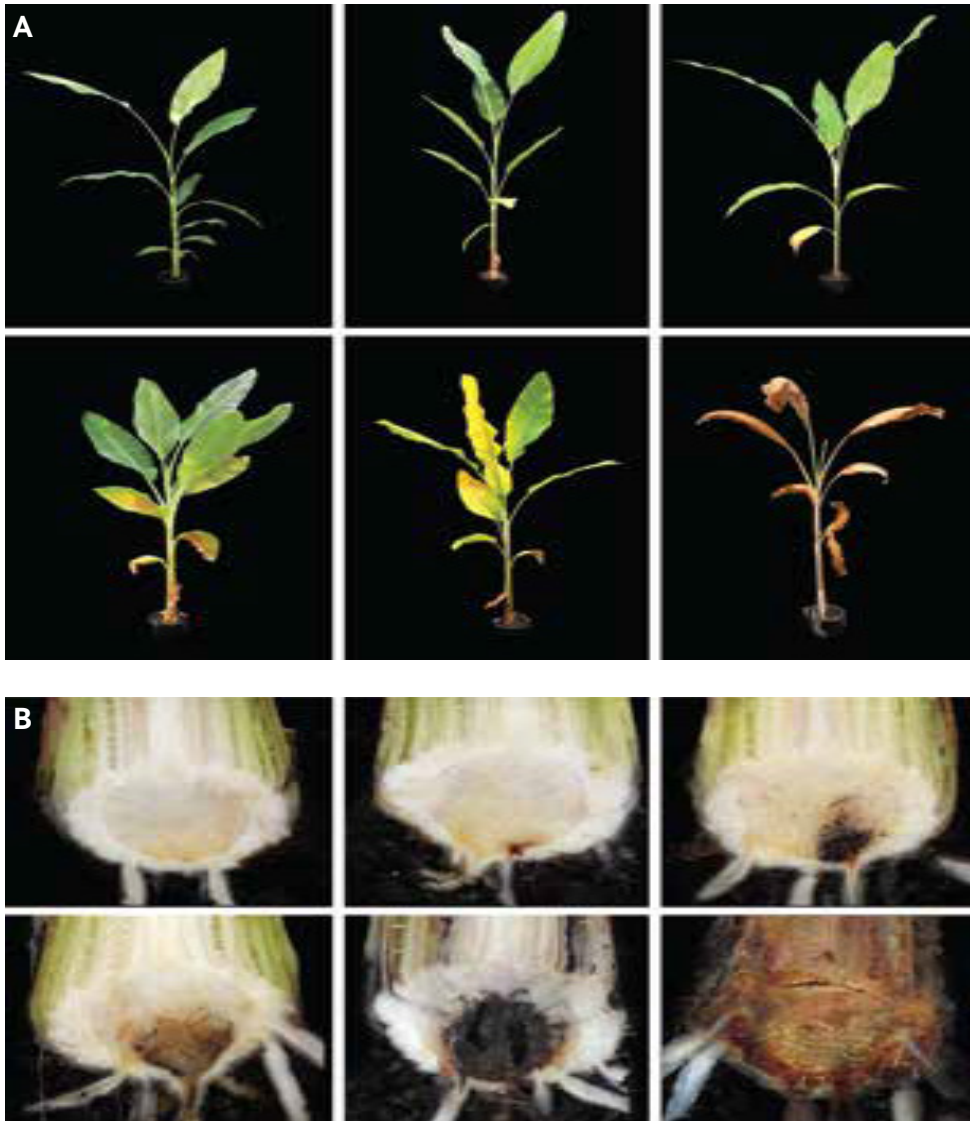
Supplementary Figure 1 Leaves and rhizomes severity scoring according to Garcia-Bastidas et al. (2019b). A. Leaves scores: Score 1 (healthy plant), Score 2 (initial yellowing/ chlorosis), Score 3 (wilting/yellowing leaves ≤10%), D. Score 4 (wilting/yellowing leaves 10%-50%), Score 5 (wilting/yellowing leaves ≤50-%-90%), Score 6 (wilting/yellowing leaves >90% or dead plant). B. Rhizomes scores: Score 1 (No discoloration), Score 2 (≤5% discoloration), Score 3 (6%-10% discoloration), Score 4 (21%-50% discoloration), Score 5 (50-%-90% discoloration), Score 6 (>90% discoloration).

Supplementary Table 1 Distribution of SNP markers, physically mapped on the chromosomes of the reference genome of *Musa acuminata* ssp. *malaccensis* DH 'Pahang', version 2 (http://banana-genome-hub.southgreen.fr/organism/Musa/acuminata).

Chromosome	Total SNP markers	Filtered SNPs markers	Lenght of genetic distance (cM)
1	1,870	162	258
2	1,973	229	165
3	2,664	328	281
4	3,088	305	207
5	2,283	254	258
6	3,021	325	302
7	2,204	224	221
8	2,732	303	256
9	2,617	277	254
10	2,441	208	440
11	1,818	187	186
Mitochondria	313		
Unanchored	5,338		
Total	32,362	2,802	2,828
Mean	2,428	255	257

Supplementary Table 2 A list of the predicted genes in the window of the resistance gene for Race 1 according to reference genome of *Musa acuminata* ssp. *malaccensis* DH 'Pahang', version 2, (http://banana-genome-hub.southgreen.fr/organism/Musa/acuminata) at the range 0-4.3 Mb..

Gene Code	Physical position in the DH "Pahang"		Putative gene
	start	stop	
Ma10_p00010.1	106217	108235	Ma10_g00010~ exocyst complex component EXO70A1-like~ unknown_gene~ missing_functional_completeness
Ma10_p00020.1	516157	519871	Ma10_g00020~ craniofacial development protein 1-like~ unknown_gene~ missing_functional_completeness
Ma10_p00030.1	597395	606009	Ma10_g00030~ endoplasmic reticulum-Golgi intermediate compartment protein 3-like~ unknown_gene~ missing_functional_completenes
Ma10_p00040.1	626059	628374	Ma10_g00040~ uncharacterized glycosyl hydrolase Rv2006/MT2062, putative, expressed~ Rv2006~ complete missing_functional_completeness
Ma10_p00050.1	752579	753205	Ma10_g00050~ mavicyanin-like~ unknown_gene~ missing_functional_completeness
Ma10_p00060.1	754512	778558	Ma10_g00060~ coiled-coil domain-containing protein 22 homolog, transcript variant X1~ unknown_gene~
Ma10_p00060.2	754512	778558	Ma10_g00060~ coiled-coil domain-containing protein 22 homolog, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00060.3	754512	778558	Ma10_g00060~ coiled-coil domain-containing protein 22 homolog, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00060.4	754512	778558	Ma10_g00060~ coiled-coil domain-containing protein 22 homolog, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00070.1	909174	911817	Ma10_g00070~ serine/threonine-protein kinase D6PKL1-like, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00070.2	909174	911817	Ma10_g00070~ serine/threonine-protein kinase D6PKL1-like, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00070.3	909174	911817	Ma10_g00070~ serine/threonine-protein kinase D6PKL1-like, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00080.1	913616	914370	Ma10_g00080~ UDP-glycosyltransferase 89A2-like~ unknown_gene~ missing_functional_completeness
Ma10_p00090.1	1076495	1079259	Ma10_g00090~ transmembrane protein 245-like~ unknown_gene~ missing_functional_completeness

Supplementary Table 2 continues

Gene Code	Physical position in the DH "Pahang"		Putative gene
	start	stop	
Ma10_p00100.3	1251994	1260969	Ma10_g00100~ formin-binding protein 4, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00100.1	1255854	1260969	Ma10_g00100~ formin-binding protein 4, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00100.2	1255854	1260969	Ma10_g00100~ formin-binding protein 4, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00110.1	1261682	1271748	Ma10_g00110~ calcineurin B-like protein 3~ unknown_gene~ missing_functional_completeness
Ma10_p00120.1	1347091	1349143	Ma10_g00120~ glucan endo-1,3-beta-glucosidase 14-like, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00120.2	1347091	1348198	Ma10_g00120~ glucan endo-1,3-beta-glucosidase 14-like, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00130.1	1352125	1355477	Ma10_g00130~ calcium-binding mitochondrial carrier protein SCaMC-1-like~ unknown_gene~ missing_functional_completeness
Ma10_p00140.1	1380269	1381801	Ma10_g00140~ 3-ketoacyl-CoA synthase 11~ KCS5~ complete
Ma10_p00150.1	1383864	1387312	Ma10_g00150~ zeaxanthin epoxidase, chloroplastic~ unknown_gene~ missing_functional_completeness
Ma10_p00160.1	1387497	1389440	Ma10_g00160~ pentatricopeptide repeat-containing protein At1g31430~ unknown_gene~ missing_functional_completeness
Ma10_p00170.1	1392114	1395221	Ma10_g00170~ root phototropism protein 3-like~ unknown_gene~ missing_functional_completeness
Ma10_p00180.1	1408918	1415692	Ma10_g00180~ putative pleiotropic drug resistance protein 7~ unknown_gene~ missing_functional_completeness
Ma10_p00190.1	1418732	1434365	Ma10_g00190~ coiled-coil domain-containing protein 93, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00190.2	1418732	1434365	Ma10_g00190~ coiled-coil domain-containing protein 93, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00200.1	1436216	1437394	Ma10_g00200~ gibberellin 3-beta-dioxygenase 1-like~ unknown_gene~ missing_functional_completeness

Gene Code	Physical position in the DH "Pahang"		Putative gene
	start	stop	
Ma10_p00210.1	1455447	1457078	Ma10_g00210~ F-box/kelch-repeat protein At1g80440-like~ unknown_gene~ missing_ functional_completeness
Ma10_p00220.1	1457117	1458548	Ma10_g00220~ uncharacterized LOC103973058, transcript variant X3~ unknown_gene~ missing_ functional_completeness
Ma10_p00220.2	1457117	1458548	Ma10_g00220~ uncharacterized LOC103973058, transcript variant X3~ unknown_gene~ missing_ functional_completeness
Ma10_p00220.3	1457117	1458548	Ma10_g00220~ uncharacterized LOC103973058, transcript variant X3~ unknown_gene~ missing_ functional_completeness
Ma10_p00230.1	1462268	1462450	Ma10_g00230~ Putative Arabinogalactan peptide 22~ AGP22~ modules
Ma10_p00240.1	1464833	1465264	Ma10_g00240~ uncharacterized LOC103973057~ unknown_gene~ missing_functional_completeness
Ma10_p00250.1	1491520	1498629	Ma10_g00250~ uncharacterized LOC103973056, transcript variant X2~ unknown_gene~ missing_ functional_completeness
Ma10_p00250.2	1492130	1498629	Ma10_g00250~ uncharacterized LOC103973056, transcript variant X2~ unknown_gene~ missing_ functional_completeness
Ma10_p00260.1	1499455	1502720	Ma10_g00260~ UDP-galactose/UDP-glucose transporter 3-like~ unknown_gene~ missing_ functional_completeness
Ma10_p00270.1	1521833	1522366	Ma10_g00270~ PRA1 family protein F3-like~ unknown_gene~ missing_functional_completeness
Ma10_p00280.1	1524018	1526418	Ma10_g00280~ putative transcription factor bHLH041~ unknown_gene~ missing_functional_ completeness
Ma10_p00290.1	1527067	1531164	Ma10_g00290~ Putative Transmembrane 9 superfamily member 4~ TM9SF4~ complete
Ma10_p00300.1	1539387	1541792	Ma10_g00300~ eukaryotic initiation factor 4A-15-like~ unknown_gene~ missing_functional_ completeness
Ma10_p00310.1	1550418	1551008	Ma10_g00310~ uncharacterized LOC103973081~ unknown_gene~ missing_functional_completeness
Ma10_p00320.1	1552114	1553667	Ma10_g00320~ zinc transporter 7-like~ unknown_ gene~ missing_functional_completeness
Ma10_p00330.1	1554631	1559684	Ma10_g00330~ reticulocalbin-2-like~ unknown_ gene~ missing_functional_completeness

Supplementary Table 2 continues

Gene Code	Physical position in the DH "Pahang"		Putative gene
	start	stop	
Ma10_p00340.1	1573871	1575950	Ma10_g00340~ zinc finger protein NUTCRACKER-like, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00340.2	1573871	1575950	Ma10_g00340~ zinc finger protein NUTCRACKER-like, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00340.3	1573871	1575950	Ma10_g00340~ zinc finger protein NUTCRACKER-like, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00350.1	1579421	1589344	Ma10_g00350~ far upstream element-binding protein 2-like~ unknown_gene~ missing_functional_completeness
Ma10_p00360.1	1589916	1590766	Ma10_g00360~ inorganic pyrophosphatase 2-like, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00360.2	1589916	1590997	Ma10_g00360~ inorganic pyrophosphatase 2-like, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00370.1	1593942	1595180	Ma10_g00370~ probable F-box protein At4g22030~ unknown_gene~ missing_functional_completeness
Ma10_p00380.1	1600629	1604047	Ma10_g00380~ Putative Whole genome shotgun sequence of line PN40024, scaffold_8.assembly12x (Fragment)~ ATH1~ modules
Ma10_p00390.1	1606305	1607438	Ma10_g00390~ zinc finger CCCH domain-containing protein 35-like~ unknown_gene~ missing_functional_completeness
Ma10_p00400.1	1614199	1615123	Ma10_g00400~ probable LRR receptor-like serine/threonine-protein kinase At1g51820~ unknown_gene~ missing_functional_completeness
Ma10_p00410.1	1618250	1621530	Ma10_g00410~ Putative Probable LRR receptor-like serine/threonine-protein kinase At1g05700~ At1g05700~ fragment
Ma10_p00420.1	1628209	1629339	Ma10_g00420~ LRR receptor-like serine/threonine-protein kinase ERECTA~ unknown_gene~ missing_functional_completeness
Ma10_p00430.1	1651242	1652705	Ma10_g00430~ Probable LRR receptor-like serine/threonine-protein kinase At1g51820~ At1g51820~ fragment
Ma10_p00440.1	1658083	1658712	Ma10_g00440~ probably inactive leucine-rich repeat receptor-like protein kinase At2g25790~ unknown_gene~ missing_functional_completeness

Gene Code	Physical position in the DH "Pahang"		Putative gene
	start	stop	
Ma10_p00450.1	1671933	1679978	Ma10_g00450~ Putative Probable LRR receptor-like serine/threonine-protein kinase At4g36180~ At4g36180~ complete
Ma10_p00460.1	1681347	1682052	Ma10_g00460~ Hypothetical protein~ At1g51820~ missing_functional_completeness; Probable LRR receptor-like serine/threonine-protein kinase At1g05700; PREDICTED: probable LRR receptor-like serine/threonine-protein kinase At4g29180 isoform X4 [Musa acuminata subsp. malaccensis]
Ma10_p00470.1	1684067	1686962	Ma10_g00470~ leucine rich repeat protein, putative~ ERL1~ complete
Ma10_p00480.1	1692198	1694093	Ma10_g00480~ probable LRR receptor-like serine/threonine-protein kinase At4g36180~ unknown_gene~ missing_functional_completeness
Ma10_p00490.1	1710726	1713763	Ma10_g00490~ Putative Probable LRR receptor-like serine/threonine-protein kinase At4g36180~ At4g36180~ complete
Ma10_p00500.1	1752896	1754926	Ma10_g00500~ leucine-rich repeat receptor-like protein kinase PEPR2~ unknown_gene~ missing_functional_completeness
Ma10_p00510.1	1754955	1761490	Ma10_g00510~ LRR receptor-like serine/threonine-protein kinase ERECTA~ unknown_gene~ missing_functional_completeness
Ma10_p00520.1	1766856	1768436	Ma10_g00520~ Putative Probable LRR receptor-like serine/threonine-protein kinase At1g05700~ At1g05700~ fragment
Ma10_p00530.1	1768529	1769229	Ma10_g00530~ Hypothetical protein~ PGIP3~ missing_functional_completeness; PREDICTED: LRR receptor-like serine/threonine-protei kinase FLS2 [Musa acuminata subsp. malaccensis] Select seq ref|XP_009385819.1|
Ma10_p00540.1	1779420	1781023	Ma10_g00540~ Hypothetical protein~ unknown_gene~ missing_functional_completeness
Ma10_p00550.1	1781038	1783746	Ma10_g00550~ probable leucine-rich repeat receptor-like protein kinase At1g35710~ unknown_gene~ missing_functional_completeness
Ma10_p00560.1	1785620	1786664	Ma10_g00560~ leucine-rich repeat receptor-like protein CLAVATA2~ unknown_gene~ missing_functional_completeness
Ma10_p00570.1	1814769	1817522	Ma10_g00570~ LRR receptor-like serine/threonine-protein kinase FLS2~ unknown_gene~ missing_functional_completeness

Supplementary Table 2 continues

Gene Code	Physical position in the DH "Pahang"		Putative gene
	start	stop	
Ma10_p00580.1	1818828	1818959	Ma10_g00580~ Hypothetical protein~ slc17a6b~ missing_functional_completeness
Ma10_p00590.1	1847632	1849029	Ma10_g00590~ protein IQ-DOMAIN 14-like, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00590.2	1847632	1849029	Ma10_g00590~ protein IQ-DOMAIN 14-like, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00600.1	1865930	1868792	Ma10_g00600~ auxin-responsive protein IAA16-like~ unknown_gene~ missing_functional_completeness
Ma10_p00610.1	1875192	1879331	Ma10_g00610~ 40S ribosomal protein S24-1-like~ unknown_gene~ missing_functional_completeness
Ma10_p00620.1	1882093	1886270	Ma10_g00620~ transcription factor MYB1R1-like~ unknown_gene~ missing_functional_completeness
Ma10_p00630.1	1887218	1921335	Ma10_g00630~ probable RNA-dependent RNA polymerase 5, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p00630.2	1887218	1918276	Ma10_g00630~ probable RNA-dependent RNA polymerase 5, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p00630.3	1887218	1921335	Ma10_g00630~ probable RNA-dependent RNA polymerase 5, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p00630.4	1887218	1921335	Ma10_g00630~ probable RNA-dependent RNA polymerase 5, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p00640.1	1926276	1928030	Ma10_g00640~ uncharacterized LOC103973075~ unknown_gene~ missing_functional_completeness
Ma10_p00650.1	1954630	1956555	Ma10_g00650~ LRR receptor-like serine/threonine-protein kinase GSO1~ unknown_gene~ missing_functional_completeness
Ma10_p00660.1	1956649	1957080	Ma10_g00660~ receptor-like protein 2~ unknown_gene~ missing_functional_completeness; PREDICTED: leucine-rich repeat receptor protein kinase MSL1-like [Musa acuminata subsp. malaccensis] Select seq ref|XP_009385818.1|
Ma10_p00670.1	1970155	1970781	Ma10_g00670~ uncharacterized LOC103973034~ unknown_gene~ missing_functional_completeness
Ma10_p00680.1	1977712	1979532	Ma10_g00680~ uncharacterized protein DDB_G0290685-like, transcript variant X2~ unknown_gene~ missing_functional_completeness

Gene Code	Physical position in the DH "Pahang"		Putative gene
	start	stop	
Ma10_p00680.2	1977712	1979532	Ma10_g00680~ uncharacterized protein DDB_G0290685-like, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p00690.1	1980557	1983954	Ma10_g00690~ monoglyceride lipase-like~ unknown_gene~ missing_functional_completeness
Ma10_p00700.1	1984946	1992137	Ma10_g00700~ phosphoribosylaminoimidazole carboxylase, chloroplastic-like, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p00700.2	1987221	1992137	Ma10_g00700~ phosphoribosylaminoimidazole carboxylase, chloroplastic-like, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p00710.1	1994710	2012932	Ma10_g00710~ vacuolar fusion protein CCZ1 homolog, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00710.2	1994710	2012932	Ma10_g00710~ vacuolar fusion protein CCZ1 homolog, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00720.1	2016911	2017750	Ma10_g00720~ zinc-finger homeodomain protein 2-like~ unknown_gene~ missing_functional_completeness
Ma10_p00730.1	2146537	2147660	Ma10_g00730~ glutaredoxin domain containing protein, putative, expressed~ hpaP~ fragment
Ma10_p00740.1	2156552	2156758	Ma10_g00740~ Protein EPIDERMAL PATTERNING FACTOR 2~ EPF2~ fragment
Ma10_p00750.1	2398024	2406142	Ma10_g00750~ nascent polypeptide-associated complex subunit alpha-like protein 1~ unknown_gene~ missing_functional_completeness
Ma10_p00760.1	2407385	2408898	Ma10_g00760~ Lichenase~ GNS1~ complete
Ma10_p00770.1	2479475	2479585	Ma10_g00770~ Acyl carrier protein, mitochondrial~ At2g44620~ fragment
Ma10_p00780.1	2653315	2654362	Ma10_g00780~ Hypothetical protein~ unknown_gene~ missing_functional_completeness
Ma10_p00790.1	2654454	2658109	Ma10_g00790~ adenosine deaminase-like protein, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p00790.2	2654454	2658173	Ma10_g00790~ adenosine deaminase-like protein, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p00800.1	2713357	2713455	Ma10_g00800~ Elongation factor G~ fusA~ fragment
Ma10_p00810.1	3063251	3065332	Ma10_g00810~ regulatory protein NPR5-like, transcript variant X2~ unknown_gene~ missing_functional_completeness

Gene Code	Physical position in the DH "Pahang"		Putative gene
	start	stop	
Ma10_p00810.2	3063251	3065332	Ma10_g00810~ regulatory protein NPR5-like, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p00820.1	3443968	3445788	Ma10_g00820~ putative pentatricopeptide repeat-containing protein At1g56570~ unknown_gene~ missing_functional_completeness
Ma10_p00830.1	3645541	3711155	Ma10_g00830~ uncharacterized LOC103999822~ unknown_gene~ missing_functional_completeness
Ma10_p00840.1	3752896	3765321	Ma10_g00840~ RRP12-like protein~ unknown_gene~ missing_functional_completeness
Ma10_p00850.1	3781738	3782094	Ma10_g00850~ Putative LTPL2 - Protease inhibitor/seed storage/LTP family protein precursor, expressed~ ZK686.2~ fragment
Ma10_p00860.1	3790199	3792038	Ma10_g00860~ neurogenic protein mastermind-like~ unknown_gene~ missing_functional_completeness
Ma10_p00870.1	3799888	3800676	Ma10_g00870~ Hypothetical protein~ uba3~ missing_functional_completeness
Ma10_p00880.1	3805682	3810351	Ma10_g00880~ DNA repair protein UVH3~ UVH3~ fragment
Ma10_p00890.1	3816940	3819301	Ma10_g00890~ ARM REPEAT PROTEIN INTERACTING WITH ABF2-like~ unknown_gene~ missing_functional_completeness
Ma10_p00900.1	3832611	3833202	Ma10_g00900~ uncharacterized LOC104000068~ unknown_gene~ missing_functional_completeness
Ma10_p00910.1	3835987	3854439	Ma10_g00910~ DNA repair protein UVH3, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00910.2	3835987	3854439	Ma10_g00910~ DNA repair protein UVH3, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00910.3	3835987	3854439	Ma10_g00910~ DNA repair protein UVH3, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00910.4	3835987	3854439	Ma10_g00910~ DNA repair protein UVH3, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p00920.1	3856484	3861715	Ma10_g00920~ Hypothetical protein~ RPS3C~ missing_functional_completeness
Ma10_p00930.1	3861722	3862640	Ma10_g00930~ Hypothetical protein~ unknown_gene~ missing_functional_completeness
Ma10_p00940.1	3873099	3873409	Ma10_g00940~ Putative Uncharacterized protein 126R~ IIV3-126R~ fragment

Gene Code	Physical position in the DH "Pahang"		Putative gene
	start	stop	
Ma10_p00950.1	3873577	3875723	Ma10_g00950~ Putative Ankyrin repeat and BTB/POZ domain-containing protein 2~ ABTB2~ fragment
Ma10_p00960.1	3875676	3878833	Ma10_g00960~ ARM REPEAT PROTEIN INTERACTING WITH ABF2-like~ unknown_gene~ missing_functional_completeness
Ma10_p00970.1	3881030	3881509	Ma10_g00970~ Hypothetical protein~ unknown_gene~ missing_functional_completeness
Ma10_p01060.1	3881062	3941845	Ma10_g01060~ ARM REPEAT PROTEIN INTERACTING WITH ABF2-like~ unknown_gene~ missing_functional_completeness
Ma10_p00980.1	3890937	3894777	Ma10_g00980~ ABTB1 - Armadillo repeats with a Bric-a-Brac, Tramtrack, Broad Complex BTB domain, expressed~ unknown_gene~ fragment
Ma10_p00990.1	3896508	3896748	Ma10_g00990~ DNA repair protein UVH3~ UVH3~ fragment
Ma10_p01000.1	3898518	3898882	Ma10_g01000~ DNA repair protein UVH3-like~ unknown_gene~ missing_functional_completeness
Ma10_p01010.1	3902908	3903290	Ma10_g01010~ Hypothetical protein~ unknown_gene~ missing_functional_completeness
Ma10_p01020.1	3904111	3909869	Ma10_g01020~ ARM REPEAT PROTEIN INTERACTING WITH ABF2-like~ unknown_gene~ missing_functional_completeness
Ma10_p01030.1	3918399	3919134	Ma10_g01030~ DNA repair protein UVH3~ UVH3~ fragment
Ma10_p01040.1	3925690	3929958	Ma10_g01040~ ABTB1 - Armadillo repeats with a Bric-a-Brac, Tramtrack, Broad Complex BTB domain, expressed~ VAC8~ fragment
Ma10_p01050.1	3934264	3935047	Ma10_g01050~ ARM REPEAT PROTEIN INTERACTING WITH ABF2-like~ unknown_gene~ missing_functional_completeness
Ma10_p01070.1	3944788	3949487	Ma10_g01070~ peroxisomal adenine nucleotide carrier 1-like~ unknown_gene~ missing_functional_completeness
Ma10_p01080.1	3952495	3988186	Ma10_g01080~ potassium transporter 7-like~ unknown_gene~ missing_functional_completeness
Ma10_p01090.1	3990712	3994653	Ma10_g01090~ serine/threonine-protein kinase PBS1-like~ unknown_gene~ missing_functional_completeness
Ma10_p01100.1	3996682	4006881	Ma10_g01100~ sodium/hydrogen exchanger 2-like~ unknown_gene~ missing_functional_completeness

Supplementary Table 2 continues

5

Gene Code	Physical position in the DH "Pahang"		Putative gene
	start	stop	
Ma10_p01110.1	4010551	4019481	Ma10_g01110~ vesicle-associated protein 1-2-like, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p01110.2	4010551	4019481	Ma10_g01110~ vesicle-associated protein 1-2-like, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p01120.1	4023728	4025648	Ma10_g01120~ secoisolariciresinol dehydrogenase-like, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p01120.2	4023728	4025648	Ma10_g01120~ secoisolariciresinol dehydrogenase-like, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p01130.1	4027514	4031525	Ma10_g01130~ 1-aminocyclopropane-1-carboxylate oxidase homolog 3-like~ unknown_gene~ missing_functional_completeness
Ma10_p01140.1	4041851	4068539	Ma10_g01140~ histone-lysine N-methyltransferase setd3~ unknown_gene~ missing_functional_completeness
Ma10_p01150.1	4074177	4110993	Ma10_g01150~ protein NRDE2 homolog~ unknown_gene~ missing_functional_completeness
Ma10_p01160.1	4114992	4116311	Ma10_g01160~ probable anion transporter 6, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p01160.2	4114992	4116311	Ma10_g01160~ probable anion transporter 6, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p01170.1	4126875	4129467	Ma10_g01170~ gamma-tubulin complex component 3-like~ unknown_gene~ missing_functional_completeness
Ma10_p01180.1	4159472	4161513	Ma10_g01180~ protein TRANSPARENT TESTA 1-like~ unknown_gene~ missing_functional_completeness
Ma10_p01190.1	4240738	4241247	Ma10_g01190~ uncharacterized LOC103999870~ unknown_gene~ missing_functional_completeness
Ma10_p01200.1	4242611	4243120	Ma10_g01200~ Hypothetical protein~ FAM129C~ missing_functional_completeness
Ma10_p01210.1	4246700	4247011	Ma10_g01210~ histone H4~ unknown_gene~ missing_functional_completeness
Ma10_p01220.1	4248690	4253026	Ma10_g01220~ sporulation-specific protein 15-like~ unknown_gene~ missing_functional_completeness

Gene Code	Physical position in the DH "Pahang"		Putative gene
	start	stop	
Ma10_p01230.1	4284105	4285001	Ma10_g01230~ protein SENSITIVITY TO RED LIGHT REDUCED 1, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p01230.2	4284105	4285001	Ma10_g01230~ protein SENSITIVITY TO RED LIGHT REDUCED 1, transcript variant X1~ unknown_gene~ missing_functional_completeness
Ma10_p01240.1	4293509	4297526	Ma10_g01240~ uncharacterized LOC103999871~ unknown_gene~ missing_functional_completeness
Ma10_p01250.1	4320226	4323566	Ma10_g01250~ protein IN2-1 homolog B-like~ unknown_gene~ missing_functional_completeness
Ma10_p01260.1	4330518	4333327	Ma10_g01260~ serine/threonine-protein kinase PBS1-like, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p01260.2	4330518	4333327	Ma10_g01260~ serine/threonine-protein kinase PBS1-like, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p01270.1	4334006	4340804	Ma10_g01270~ elongator complex protein 6, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p01270.2	4334006	4340804	Ma10_g01270~ elongator complex protein 6, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p01270.3	4334006	4340804	Ma10_g01270~ elongator complex protein 6, transcript variant X2~ unknown_gene~ missing_functional_completeness
Ma10_p01280.1	4347745	4351833	Ma10_g01280~ AP2-like ethylene-responsive transcription factor TOE3, transcript variant X3~ unknown_gene~ missing_functional_completeness
Ma10_p01280.2	4347745	4351833	Ma10_g01280~ AP2-like ethylene-responsive transcription factor TOE3, transcript variant X3~ unknown_gene~ missing_functional_completeness
Ma10_p01280.3	4347745	4351833	Ma10_g01280~ AP2-like ethylene-responsive transcription factor TOE3, transcript variant X3~ unknown_gene~ missing_functional_completeness
Ma10_p01290.1	4396454	4397485	Ma10_g01290~ uncharacterized LOC103999867~ unknown_gene~ missing_functional_completeness

Highlighted: putative resistance gene analogue

5

Chapter 6.

General Discussion

In this chapter, I review the most recent genetic knowledge of Indonesian bananas as presented in my thesis that revolves around genetic diversity, fertility and breeding and resistance to Fusarium wilt. Indonesia is well-known as the archipelago for wild bananas, making it the major source of its genetic diversity. However, research of this centre of origin has mostly disregarded the island of Sumatra, the third largest island of Indonesia. Only one species, *Musa acuminata* ssp. *sumatrana* (ITC1701), was deposited in a gene bank (Ruas et al., 2017), despite that Nasution (1991) identified at least five varieties of wild diploids on the island. In this thesis project I was encouraged to explore Sumatra to obtain more insight in the genetic diversity of wild *Musa acuminata*. Once that was accomplished, I considered the wealth of these resources for international breeding programs, but I also realized that potential bottlenecks, such as genomic instability and infertility, might hamper their deployment. I, therefore, studied meiotic behaviour of several of these accessions in more detail. Such collections are priceless, but also worthless if their potential is not unveiled. Consequently, I was eager to investigate their potential as source of resistance to Fusarium wilt of banana (FWB), one of the most devastating threats of international banana cultivation in modern times. However, to support breeding efforts and other ways of crop improvement, identification of desirable genes is required. Segregating populations were developed for genetic analyses and gene mapping, an endeavour that has been disregarded for decades despite the foundational value of resistance to FWB in Cavendish bananas that enabled its cultivation around the world. In the following paragraphs, I will discuss these results in more detail and explain how, when and why these results should be taken in a broader context to eventually support sustainable banana production.

Genetic diversity

The most positive thing that we can learn from the replacement of the susceptible "Gros Michel" by the resistant "Cavendish" is hope. Hope to save banana under the threat of FWB. However, among the thousands of banana accessions and varieties, there is no one even close to "Cavendish" in terms of quality and performance. Finding accessions amongst the natural diversity that are suitable for the industry or for breeding to improve the overall resistance to biotic threats, such as diseases, requires a thorough understanding of the genetics of bananas. Exploration of new wild banana germplasm is the start of understanding genetic diversity. In the 1950's, Simmonds reported the discovery of wild bananas during an exploration that covered India, Burma, Thailand, Malaya, Papua New Guinea, Queensland and Samoa, during what could have been the first well documented banana survey. In the period 1990-2015 the Finnish botanist, Markku Häkkinen, conducted 21 expeditions in China and

Southeast Asia and published no less than 90 publications, including the description of 15 wild species and 49 varieties (Väre, Gogoi and Arisdason, 2016). Later, Sutanto et al. (2016) explored Eastern Indonesia and added 35 accessions to the national collection at the Indonesian Tropical Fruit Research Institute (ITFRI) at Solok, and Sardos et al. (2017) reported 61 wild accessions and cultivars resulting from their exploration of Bougainville, Papua New Guinea. However, these explorations were primarily focused on collecting as many new accessions as possible but did not aim for a sampling of diversity for population analyses and possible discovery of disease resistances. Most likely, other banana explorations have been organized but were never reported. For instance, the living collections of the Research Center for Biology LIPI (RCB) contains 566 accessions of bananas including 32 accessions of wild *M. acuminata* comprising 16 varieties or subspecies, four accessions of *M. balbisiana* and some wild *Musa* spp. resulting from explorations across Indonesia (Poerba *et al.*, 2016, 2018), but details of these journeys, such as GPS locations, are not yet published. Also, the National Herbarium of the Netherlands and the Herbarium Bogoriense, 192 and 180 years old, respectively (Thiers, 2020), give us an impression of the activities of European botanists, especially Dutch explorers with regard to wild bananas as demonstrated by 108 years old herbaria in the collection (http://medialib.naturalis.nl/). Taken together, exploring genetic diversity of banana in Indonesia is still limited and hence we likely grossly underestimate its contribution to an already overwhelming genetic diversity in Southeast Asia (Perrier *et al.*, 2011; Volkaert, 2011, 2018). In this thesis, I described the first most comprehensive qualitative and quantitative banana exploration of Sumatra. Hence, I did not only focus on the total number of accessions, but primarily looked beyond the numbers into diversity and population structure as a foundation for subsequent analyses of wild *M. acuminata*.

Phenotypical or morphological characters may be intuitively expected to be associated with genetic background (Kaliontzopoulou *et al.*, 2018) and hence, should be validated by genetic analyses enabling the discovery and use of molecular markers to increase throughput and overview (Häkkinen, 2013). These markers evolved over the years (Manzo-Sánchez *et al.*, 2015) in accuracy and enabled the analysis of genetic diversity of entire populations and consequently resulted in a high-resolution revision of banana taxonomy (Häkkinen, 2013; Chapter 2). The most recent analyses of the *Musa* germplasm has been accomplished by genotyping by sequencing and whole genome analyses (Janssens *et al.*, 2016; Wu *et al.*, 2016).

Studying genetic diversity is also necessary to investigate the association between wild relatives and cultivars. For instance, we identified that the cultivar "Pisang Jari Buaya" carries alleles CAT2-A42 and IDH1-A40 that are also present in the wild accession

SSB-34. Hence, we consider SSB-34 an important *ex-situ* accession. I collected SSB-34 at a riverside cliff of the Air Besi river at Desa Curup, North Bengkulu district that is situated 40 km North of Bengkulu city (Figure 1). In Desa Curup and surroundings, there are several rivers with cliffs (Figure 1) that provide an ideal ecological niche for wild bananas, away from human activities and rubber and palm oil plantations in this district. Hence, this area (at coordinates -3.5147, 102.1878) requires further detailed exploration to map its genetic diversity. It underscores that a comprehensive genetic study of wild relatives in the centre of diversity is important to evaluate the magnitude of its diversity along with population structure. Furthermore, *ex-situ* and *in-situ* conservation warrants availability of wild relatives as well as cultivars for breeding programs.

In Indonesia, banana collections are under the responsibility of the Department of Agriculture and are mostly deposited at ITFRI. However, these bananas typically comprise cultivars and the size and number of collections do not represent the various populations across Indonesia. Wild relatives are important gene donors for plant improvement as genetic resources for alleviated biotic and abiotic stress and to enrich agronomic performance. One of the most important wild banana species is *M. acuminata* ssp. *malaccensis*, which is intensively studied due to its resistance to FWB (Dale *et al.*, 2017). Albeit that this is a top priority for breeding, its agronomic traits might be surpassed by other species such as ssp. *sumatrana* (see Chapter 2) and ssp. *banksii* (https://www.promusa.org/). These wild relatives have outstanding productivity characteristics such as the number of hands per bunch (up to 25 hands) and overall bunch form (vertical), which are important aspects for adequate yields. The plethora of other subspecies in *in-situ* or *in-vitro* collections still awaits scrutiny at the broadest level from bioassays for disease resistance to performance under abiotic stress and yield potential to deploy their richness for the benefit of growers and consumers.

Currently, the Musa Germplasm Information System (Ruas, et al., 2017) comprises 29 collections with in total 65,548 accessions that only cover a mere 249 accessions of wild *M. acuminata* and 93 accessions of *M. balbisiana,* either in *ex-situ* collections or tissue culture and/or cryo-collections (Figure 2). This is only 0.52% of the entire mapped diversity. However, the duplicates across these collections are unknown and hence, the actual representation of truly wild germplasm might be higher, but still is extraordinarily small compared to the cultivars and therefore hampers genetic studies (Volkaert, 2011). Therefore, my work is a starting point to truly explore diversity in the centre of origin of banana.

Figure 1 The collection site of accession SSB-34 at Desa Curup, North Bengkulu. **A-B** The star indicates the collection site of the SSB-34 on the map of Sumatra. **C** The accession grows among trees and bamboo. The arrow indicates the accession. **D** A spot near the site where SSB-34 was collected that illustrates the common situation around rivers in the Northern part of the Bengkulu province.
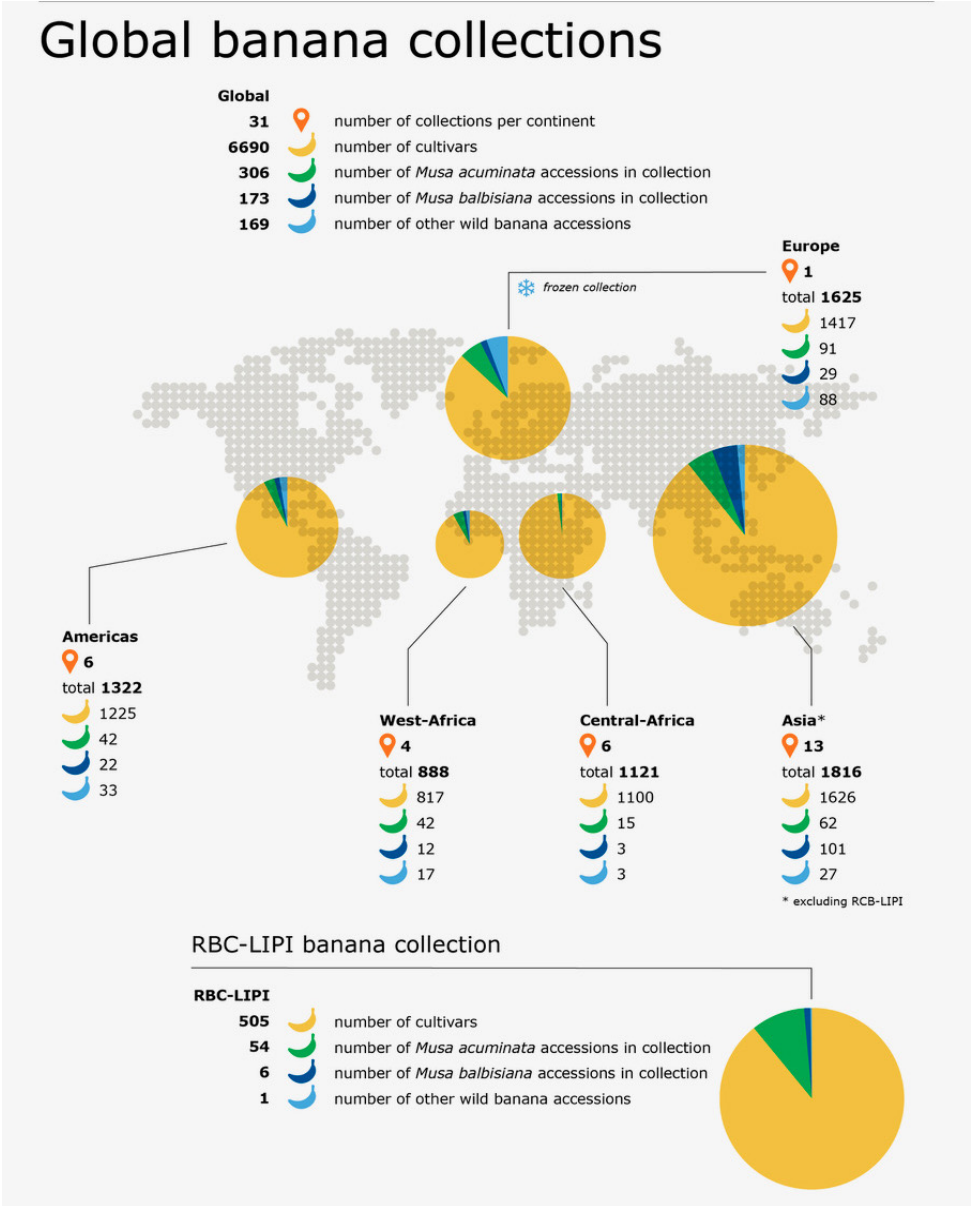
Figure 2 Global banana collections represented in the Musa Germplasm Information System, which comprise ex-situ, tissue culture and/or cryo-collections (adapted from https://www.crop-diversity.org/mgis/ and Ruas et al. 2017), excluding the RCB-LIPI collection.

Contemporary political decisions in Indonesia, however, affect the preservation and study of wild germplasm. Recently, it was decided to relocate LIPI's collection,

comprising approximately 500 accessions. This includes 32 accessions of wild relatives and 322 cultivars, collected since 2010 during various explorations across Indonesia, along with hybrids and tetraploid plants from polyploidization induction. However, the new location is still 'unknown'. Apparently, one cannot expect that the value of such collections will be readily recognized since most of these collections have never been studied for disease resistance, agronomical potential, or genetic diversity and structural chromosome variants. After investing in cross country collection missions, scientists should ensure subsequent thorough analyses of the collected materials and ensure the required budget and ascertain the necessary permits for international collaboration. Our successful identification and mapping of the first resistance gene to a Race 1 Fusarium strain – causing FWB around the world - in a wild *M. acuminata* accession from LIPI's collection illustrates the importance of germplasm collections. Hence, these collections are treasures that are already at our disposal and ready for further studies. However, maintaining living collections in the field is expensive and therefore requires effective budgeting, ideally as a joint effort and responsibility of the various institutions, such as local botanical gardens and RCB-LIPI in Indonesia. For instance, in the beginning of 2021, RCB-LIPI deposited 30 wild bananas and fifteen selected cultivars in the Cibinong Botanical Gardens to assure appropriate maintenance. Alternatively, cryo-preservation should be considered as it warrants a durable maintenance at lower costs (Escalant and Panis, 2002), not only as a national responsibility, but also as an international priority, similar to the Svalbard Global Seed Vault that contains 1,081,026 samples, from 87 gene banks in 66 countries (Westengen et al., 2013, https://www.seedvault.no/).

It is a matter of fact that Indonesia is in the core of the centre of diversity of banana (Simmonds, 1962; Perrier *et al.*, 2011; Volkaert, 2011, 2018). Hence, we should consider that these gene centres also host the richest diversity of pathogens as a result of co-evolution between host and pathogens (Takken and Rep, 2010; Möller and Stukenbrock, 2017), such as the potato pathogen *Phytopthora infestans* in the highlands of Mexico (Grünwald and Flier, 2005), the wheat pathogen *Zymoseptoria tritici* in the Fertile Crescent (Banke and McDonald, 2005; Stukenbrock *et al.*, 2007) and the cacao wilt disease caused by *Ceratocystis cacaofunesta* in the Upper Amazon (Engelbrecht, Harrington and Alfenas, 2007). Studying interactions between wild relatives and plant pathogens is therefore a corner stone to understand the evolutionary genetics of such biotic stress factors (Burdon *et al.*, 2016; Zhang *et al.*, 2017) to anticipate on their dissemination. The best example for contemporary banana production is the spread of *F. odoratissimum* Tropical Race 4 (TR4) (Ordóñez *et al.*, 2015; Ploetz, 2015) across the entire Indonesian archipelago (Maryani *et al.*, 2019).

6

In my work, we used four functional genes to study the genetic diversity and the complexity of wild *M. acuminata* on Sumatra and confirmed the presence of intra-specific hybrids among wild bananas (Volkaert, 2011). Recently, whole genome sequencing of wild banana accessions and cultivars helped to understand the genetic relationship of cultivars in more detail. Martin et al. (2020a) used single nucleotide polymorphisms across fourteen wild *M. acuminata* and ten *M. acuminata* cultivar genomes and concluded that they were derived from at least three to five ancestries. Accessions that could not be traced to these lineages either result from already extinct material or – as stipulated above – are simply absent in collections due to their limited coverage.

Soon, we should consider exploring wild banana relatives for more specific reasons, such as disease resistance or abiotic tolerance. Pathologists should join collection missions to co-collect and study the endogenous microbes found in and around the wild relatives. The rich diversity on the 4,095 metre high Mount Kinabalu in Sabah, East Malaysia (Merckx *et al.*, 2015) exemplifies the wealth of local plant and microbe communities that may be useful for disease management. In 2016, I joined two banana explorations under coordination of our project (SPIN) at Lumajang, East Java and Cianjur, West Java, in Indonesia. During these activities we identified a wealth of diverse banana cultivars as well as manifold diseases, such as Black Leaf Streak Disease (BLSD) or black Sigatoka and yellow Sigatoka disease, many cases of FWB along with blood disease and other bacterial diseases that require a better understanding of their complexity and impact for local households. Thus far, I primarily discussed *M. acuminata,* but *M. balbisiana,* which is known to be resistant to banana bunchy top disease (BBTD) and BLSD, is still poorly explored in Indonesia and elsewhere (Ahmad et al., 2014). In addition, we need to incorporate sites for abiotic stress tolerance by exploring arid areas, for instance the islands in Nusa Tenggara, which have a relatively low precipitation of approximately 1,000 mm/year compared to provinces in western Java or Sumatra with high annual precipitation reaching above 7,000 mm/year. In such regions, we should be able to find drought tolerant accessions. Temperate regions may harbour germplasm with cold tolerance at different altitudes. Clearly, the nations in the centre of diversity of banana have an important responsibility and role in banana research, which traditionally has been spearheaded and dominated by countries that once colonized Africa and Southeast Asia such as Belgium, France, and The Netherlands. This responsibility requires establishing collaborative programs, such as the KNAW-SPIN program, and effective respectful communication between national and international researchers to unveil, preserve and utilize genetic potentials of the crop for the benefit of national and global consumers and producers thereby embracing diversity and sustainable cultivation methods.

The cytogenetics of banana

One of the main challenges in banana breeding is the sterility of cultivars (Bakry *et al.*, 2009). Later, breeders realized that the role of sterility of banana is due to heterozygosity for structural chromosome abnormality (Dodds, 1943; Shepherd, 1999; Bakry, Horry and Jenny, 2020). Cytogenetic analyses are required to establish potential causal relationships between meiotic aberrations, chromosome transmission problems and sterility in hybrid genotypes. Ploidy levels are routinely determined by chromosome counting in mitotic cell complements and flow cytometry, but heterozygosity for structural chromosome variants should be addressed by chromosome analyses of spread pollen mother cells at various stages of mitosis and meiosis (Dodds, 1943; Shepherd, 1999). The occurrence of univalents and meiotic configurations involving three and more chromosomes are indicative of numerical and structural chromosome variants.

However, chromosomes of banana are relatively small, and the dense cytoplasm frequently complicates optical microscopic details. Protocol optimization is necessary for adequate observations of structural rearrangements. Even the best protocols require modifications depending on the germplasm of the studied material. I optimized the fixation method and subsequently the cell spreading and staining techniques for five wild *M. acuminata* i.e., ssp. *malaccensis* (LIPI-010), var. *breviformis* (LIPI-218), var. *sumatrana* (LIPI-457), var. *tomentosa* (LIPI-172), and var. *zebrina* (LIPI-043) and a cultivar *Musa* AA 'Rejang' (LIPI-048, (Sucrier subgroup) (Chapter 3). With the advent of genome data of *Musa* (D'Hont *et al.*, 2012; Martin *et al.*, 2016; Wu *et al.*, 2016; Belser *et al.*, 2018; Rouard *et al.*, 2018), we can now combine molecular information with microscopy techniques to develop chromosome painting to visualize their structure, including crossing-overs and synapsis (Simonikova *et al.*, 2019). Such techniques also confirmed translocations between chromosome 9 and 11 in Indica rice (Hou *et al.*, 2018) and multiple translocations between chromosomes 1, 3, 5, 7, 9 and 10 in maize (Albert *et al.*, 2019). Of course such information is important for genome assemblies where chromosome painting can identify unanchored scaffolds (Simonikova *et al.*, 2019).

Recently, comparative genome analyses showed at least six structural rearrangements (Martin et al., 2020b), while classical cytology revealed nine translocations (Shepherd, 1999). Chromosome painting techniques by fluorescent in-situ hybridization (FISH) using chromosome specific probes established the identification of all chromosomes in bananas and the anchoring with the sequence information of available reference genomes. Simonikova et al. (2019) applied this advanced chromosome painting to reciprocal translocations between wild banana accessions and cultivars. Such

6

advanced molecular cytogenetic techniques are powerful tools to evaluate the sequence orientation by next generation sequencing (NGS) and cast light on the evolution of *Musa* species at the chromosome level (Baurens et al., 2019; Martin et al., 2020a, 2020b). This cytogenetic knowledge is also directly applicable in breeding programs, as it enables efficient selection and evaluation of parents and progenies, for such chromosomal aberrations may impair chromosomal involving genes of interest. Although classical breeding supported by modern techniques is a guarantee for diversifying banana crops, we presently can consider genetic modification and genome editing as accessible and proven techniques for repairing the most disadvantageous characteristics of contemporary banana cultivars. These include susceptibility to FWB (Dale *et al.*, 2017), and BLSD (Vishnevetsky *et al.*, 2011) in Cavendish bananas and *Xanthomonas* wilt in "Sukala Ndiizi" and "Mpologoma" (Tripathi *et al.*, 2010) and banana streak virus in "Gonja Manjaya" (Tripathi *et al.*, 2019), which are important for domestic markets.

Genome diversity, gene identification and mapping

So far, four *M. acuminata* accessions and one *M. balbisiana* accession, as well as three banana relatives *i.e.*, *M. schizocarpa*, *M. itinerans* and *Ensete glaucum*, have been sequenced (Table 1). Although this contributes to the overall knowledge of the genome structure and core statistics of the number of genes and how they are categorized, including probable resistance genes, it still requires validation and marker development to effectively deploy such genes in breeding (Peraza-Echeverria *et al.*, 2008; Sutanto *et al.*, 2014). Such a strategy is common sense in many other crops including wheat, which has resulted in the iconic Catalogue of Gene Symbols for Wheat, https://wheat.pw.usda.gov/GG3/wgc that lists 8,578 mapped genes with appropriate gene nomenclature and symbolization that is frequently supplemented by a group of wheat experts, and so contributes to overview and use of these genes in breeding. Moreover, such repositories drive progress, summarize all required genetic information, and stimulate scientific exchange and discussion across communities. However, in banana this procedure still has not been initiated. Now, https://banana-genome-hub.southgreen.fr/ (Droc *et al.*, 2013) is the primary reference for banana genomes, but in-depth information about wild relatives, cultivars and breeding results, let alone mapping data and gene validations are not available.

Presently, my work (Chapter 5) is the only report of a mapped Fusarium wilt Race 1 and TR4 resistance gene in banana. Previously, five resistance gene analogues , *RGA1 through RGA5*, have been identified using degenerate primers to amplify NBS-type

sequences from another accession of *M. acuminata* ssp. *malaccensis* (Peraza-Echeverria *et al.*, 2008). One of these, *RGA2*, was cloned and validated and provides resistance to TR4 under field conditions (Dale *et al.*, 2017). Sardos et al. (2016) used a *M. acuminata* panel that comprised 80 cultivars and 25 wild bananas accessions from Papua and by GWAS identified genomic regions harbouring genes for seedlessness on chromosomes 1, 3, 4, 7, 9 and 11. However, neither of these genes were yet validated. Once genes have been fine-mapped, marker-assisted breeding can finally start (Jiang, 2013; Bahadur *et al.*, 2015). A similar trend developed in wheat breeding to *Z. tritici*, where targeted breeding only kicked-in after resistance genes were mapped (Brown *et al.*, 2015) and breeders gradually abandoned the dependency on infections by highly diverse natural populations for selection. We expect a similar trend for banana breeding to BLSD since *Pseudocercospora fijiensis*, as a Dothideomycete, has a very similar lifestyle as *Z. tritici* (Arango *et al.*, 2016). Hence, our work on resistance to Race 1 strains that cause FWB in banana is a good and necessary start to professionalize banana breeding (Chapter 5).

6

Table 1 The number of genes or predicted genes of five bananas species based on whole genome sequencing as described in https://banana-genome-hub.southgreen.fr/ (accessed 19th March 2021).

Species	ssp/accession	Gene (predicted)	Gene (predicted) associated to resistance	Gene (predicted) associated to disease resistance	Reference
Musa acuminata	ssp. *banksii*	32,692	177	130	(Rouard *et al.*, 2018)
	ssp. burmannicoides "Calcutta 4"	45,069	308	239	(Rouard *et al.*, 2018)
	ssp. *malaccensis* "DH Pahang"	45,856	280	170	(D 'hont *et al.*, 2012; Martin *et al.*, 2016)
	ssp. *zebrina*	44,702	255	180	(Rouard *et al.*, 2018)
Musa balbisiana	"DH-PKW"	33,021	146	88	(Wang *et al.*, 2019)
Musa schizocarpa *Musa itinerans*		32,809 32,456	172	101	(Belser *et al.*, 2018) (Wu *et al.*, 2016)
Ensete glaucum		36,836	121	87	https://banana-genome-hub.southgreen.fr/

Fortunately, *Fusarium* collections from a comprehensive sampling in Indonesia are now available (Maryani et al., 2018). Bioassays exposing banana diversity to a representative panel of *Fusarium* strains are necessary to estimate the efficacy of resistance. Recent data have shown that only 36% of a diploid panel of wild banana germplasm was resistant to TR4 (García, 2019). This prioritizes the exploration of germplasm for genetic analyses and underscores the necessity to explore the pathogenic capacity of endemic *Fusarium* spp. affecting banana. Our work showed that mapping resistance in populations resulting from self-pollination of heterozygous parents is possible. Alternatively, segregation populations can be generated by crossing two heterozygous fertile parents or by creating double haploids from diploid resistant accessions and cross these with susceptible accessions (Prins *et al.*, 2005). Finally, genome wide association studies (GWAS) are possible once large genomic data sets become available. As mentioned above, Sardos et al. (2016) used a *M. acuminata* panel to identify loci associated with seedlessness. Similar analyses should be conducted by generating databases with the response of such panels to (a)biotic stress to build a database with candidate genes, which will require validation before they can be used in marker-assisted breeding programs or genetic modification strategies. Such techniques and strategies underscore the above plea for collecting, preserving and analysing germplasm collections. Clearly, marker-assisted breeding brings a huge economic advantage as it circumvents cumbersome and time-consuming phenotyping analyses for FWB (Dita *et al.*, 2011; García *et al.*, 2019). Similarly, this will eventually also be applicable for other disease such as BLSD, BBTV, Moko and agronomic traits and increases the throughput of breeding programs. It will revolutionize banana breeding and shorten the horizon of the release of new, diverse germplasm, surely when combined with genomic prediction that can reduce the regular breeding trajectories of 11 - 14 years with seven years (Rateng, 2018).

In conclusion, the genetic diversity study and phylogenetic analysis in this thesis and those of my predecessors underscore the continued need for more extensive explorations to capture the overall genetic variation that consequently will increase our understanding of the evolutionary track towards domestication. Such an analysis is very much in line with the foundational work of Wallace, described in The Malay Archipelago, who travelled South East Asia and observed - earlier than Darwin – the remarkable features of speciation and isolation (Wallace, 1869). As mentioned above, indeed the existence of collections prevents extinction, but their exploration leads to exultation: the discovery and deployment of genes. This is particularly true for banana. A multidisciplinary stakeholder-wide approach is required for the transition towards a sustainable banana production.

References

Ahmad F, Martawi NM, Poerba YS, Jong H De & Kema GHJ (2020) Genetic mapping of Fusarium wilt resistance in a wild banana *Musa acuminata* ssp. *malaccensis* accession. Theoretical and Applied Genetics 133:3409–3418.

Ahmad, F., Megia, R. and Poerba, Y. S. (2014) 'Genetic Diversity of *Musa balbisiana* Colla in Indonesia Based on AFLP Marker', *HAYATI Journal of Biosciences*, 21(1), pp. 39–47. doi: 10.4308/hjb.21.1.39.

Albert, P. S. *et al.* (2019) 'Whole-chromosome paints in maize reveal rearrangements, nuclear domains, and chromosomal relationships', *Proceedings of the National Academy of Sciences of the United States of America*, 116(5), pp. 1679–1685. doi: 10.1073/pnas.1813957116.

Arango, R. E. I. *et al.* (2016) 'Combating a Global Threat to a Clonal Crop : Banana Black Sigatoka Pathogen *Pseudocercospora fijiensis* (Synonym *Mycosphaerella fijiensis*) Genomes Reveal Clues for Disease Control', *PLoS ONE Genet*, 12(8), pp. 1–36. doi: 10.1371/journal.pgen.1005876.

Bahadur, Bir *et al.* (2015) 'Genetic markers, trait mapping and marker-assisted selection in plant breeding', in Bahadur, B et al. (eds) *Plant Biology and Biotechnology: Volume II: Plant Genomics and Biotechnology*. New Delhi: Springer, pp. 65–88. doi:

10.1007/978-81-322-2283-5.

Bakry, F. *et al.* (2009) 'Genetic Improvement of Banana', in Jain, S. M. and Priyadarshan, P. M. (eds) *Breeding Plantation Tree Crops: Tropical Species*. 2nd edn. New York, NY: Springer, pp. 3–50. doi: doi:10.1007/978-0-387-71201-7_1.

Bakry, F., Horry, J. P. and Jenny, C. (2020) 'Making banana breeding more effective', in Kema, G. and Drenth, A. (eds) *Achieving sustainable cultivation of bananas Volume 2: Germplasm and genetic improvement*. Cambridge: Burleigh Dodds Science Publishing, pp. 217–256.

Banke, S. and McDonald, B. A. (2005) 'Migration patterns among global populations of the pathogenic fungus Mycosphaerella graminicola', *Molecular Ecology*, 14(7), pp. 1881–1896. doi: 10.1111/j.1365-294X.2005.02536.x.

Baurens, F. C. *et al.* (2019) 'Recombination and Large Structural Variations Shape Interspecific Edible Bananas Genomes', *Molecular Biology and Evolution*, 36(1), pp. 97–111. doi: 10.1093/molbev/msy199.

Belser, C. *et al.* (2018) 'Chromosome-scale assemblies of plant genomes using nanopore long reads and optical maps', *Nature Plants*. Springer US, 4, pp. 879–887. doi: 10.1038/s41477-018-0289-4.

Brown, J. K. M. *et al.* (2015) 'Genetics of resistance to *Zymoseptoria tritici*

and applications to wheat breeding', *Fungal Genetics and Biology*. Elsevier Inc., 79, pp. 33–41. doi: 10.1016/j.fgb.2015.04.017.

Burdon, J. J. *et al.* (2016) 'Addressing the Challenges of Pathogen Evolution on the World ' s Arable Crops', *Phytopathology*, 106(10), pp. 1117–1127.

D'Hont, A. *et al.* (2012) 'The banana (Musa acuminata) genome and the evolution of monocotyledonous plants', *Nature*, 488(7410), pp. 213-+. doi: 10.1038/nature11241.

D 'hont, A. *et al.* (2012) 'The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants'. doi: 10.1038/nature11241.

Dale, J. *et al.* (2017) 'Transgenic Cavendish bananas with resistance to Fusarium wilt tropical race 4', *Nature Communications*. Nature Publishing Group, 8(1), p. 1496. doi: 10.1038/s41467-017-01670-6.

Dita, M. A. *et al.* (2011) 'A Greenhouse Bioassay for the *Fusarium oxysporum* f. sp. cubense X "Grand Naine" (Musa, AAA, Cavendish Subgroup) Interaction', *Acta Horticulturae*, (897), pp. 377–380. doi: 10.17660/ActaHortic.2011.897.51.

Dodds, K. S. (1943) 'Geentical and cytological studies of *Musa*, V. Certain edible diploids', *Genetics*, 45, pp. 113–139.

Droc, G. *et al.* (2013) 'The banana genome hub', *Database*, 2013. doi: 10.1093/database/bat035.

Engelbrecht, C. J., Harrington, T. C. and Alfenas, A. (2007) 'Cacao Diseas-es : Important Threats to Chocolate Production Worldwide Ceratocystis Wilt of Cacao — A Disease of Increasing Importance', *Phytopathology*, 97(1), pp. 1648–1649.

Escalant, J. and Panis, B. (2002) 'Biotechnologies toward the genetic improvement in Musa', in *Proceedings of XV ACORBAT meeting, Cartagena de Indias, Colombia, 2002/10/27-11/02. AUGURA.*, pp. 68–85.

García-Bastidas, F. (2019) *Panama Disease in Banana: Spread, Screen and Genes*. Wagenuingen Univeristy & Research.

García, F. A. *et al.* (2019) 'An Improved Phenotyping Protocol for Panama Disease in Banana', *Frontiers in Plant Science*, 10(August), pp. 1–12. doi: 10.3389/fpls.2019.01006.

Grünwald, N. J. and Flier, W. G. (2005) 'The biology of Phytophthora infestans at its center of origin', *Annual Review of Phytopathology*, 43, pp. 171–190. doi: 10.1146/annurev.phyto.43.040204.135906.

Häkkinen, M. (2013) 'Reappraisal of sectional taxonomy in *Musa* (Musaceae)', *Taxon*, 62(4), pp. 809–813. doi: 10.12705/624.3.

Hou, L. *et al.* (2018) 'Chromosome painting and its applications in cultivated and wild rice'. BMC Plant Biology, pp. 1–10.

Janssens, S. B. *et al.* (2016) 'Evolutionary dynamics and biogeography of Musaceae reveal a correlation between the diversification of the banana family and the geological and

climatic history of Southeast Asia', *New Phytologist*, 210, pp. 1453–1465. doi: 10.1111/nph.13856.

Jiang, G.-L. (2013) 'Molecular Markers and Marker-Assisted Breeding in Plants', in Andersen, S. B. (ed.) *Plant Breeding from Laboratories to Fields*. London, UK: InThechOpen, pp. 45–83. doi: DOI: 10.5772/52583.

Kaliontzopoulou, A., Pinho, C. and Martínez-freiría, F. (2018) 'Where does diversity come from? Linking geographical patterns of morphological , genetic , and environmental variation in wall lizards', *BMC Evolutionary Biology*. BMC Evolutionary Biology, 18, p. 124.

Manzo-Sánchez, G. *et al.* (2015) 'Genetic Diversity in Bananas and Plantains (*Musa* spp.)', in Caliskan, M. et al. (eds) *Molecular Approaches to Genetic Diversity*, pp. 93–121. doi: DOI: 10.5772/59421.

Martin, G. *et al.* (2016) 'Martin 2016 Improvement of banana musa acuminata reference sequence using ngs data and semi-automated bioinformatics methods', *BMC genomics*, 17(243). doi: DOI 10.1186/s12864-016-2579-4.

Martin, G. *et al.* (2020a) 'Genome ancestry mosaics reveal multiple and cryptic contributors to cultivated banana', *The Plant Journal*, 102(5), pp. 1008–1025. doi: 10.1111/tpj.14683.

Martin, G. *et al.* (2020b) 'Chromosome reciprocal translocations have accompanied subspecies evolution in bananas', *The Plant Journal*, 104, pp. 1698–1711. doi: 10.1111/tpj.15031.

Maryani, N. *et al.* (2019) 'Phylogeny and genetic diversity of the banana Fusarium wilt pathogen *Fusarium oxysporum* f. sp. *cubense* in the Indonesian centre of origin', *Studies in Mycology*, 92, pp. 155–194. doi: 10.1016/j.simyco.2018.06.003.

Merckx, V. *et al.* (2015) 'Evolution of endemism on a young tropical mountain', *Nature*, 524, pp. 347–350. doi: 10.1038/nature14949.

Möller, M. and Stukenbrock, E. H. (2017) 'Evolution and genome architecture in fungal plant pathogens', *Nature Reviews Microbiology*. Nature Publishing Group, 15(12), pp. 756–771. doi: 10.1038/nrmicro.2017.76.

Nasution, R. E. (1991) *A taxonomy study of the species* Musa acuminata *Colla with its intraspecific taxa in Indonesia*. Tokyo University of Agriculture

Ordóñez, N. *et al.* (2015) 'Worse Comes to Worst: Bananas and Panama Disease—When Plant and Pathogen Clones Meet', *PLOS Pathogens*. Public Library of Science, 11(11), p. e1005197. Available at: https://doi.org/10.1371/journal.ppat.1005197.

Peraza-Echeverria, S. *et al.* (2008) 'Characterization of disease resistance gene candidates of the nucleotide binding site (NBS) type from banana and correlation of a transcriptional polymorphism with resistance to Fusarium oxysporum f.sp. cubense race 4', *Molecular Breeding*, 22, pp. 565–579. doi: 10.1007/s11032-008-9199-x.

6

Perrier, X. *et al.* (2011) 'Multidisciplinary perspectives on banana (*Musa* spp.) domestication', *Proceedings of the National Academy of Sciences*, 108(28), pp. 11311–11318. doi: 10.1073/pnas.1102001108.

Ploetz, R. C. (2015) 'Fusarium Wilt of Banana.', *Phytopathology*, 105(12), pp. 1512–21. doi: 10.1094/PHYTO-04-15-0101-RVW.

Poerba, Y. S. *et al.* (2016) *Katalog Pisang: Koleksi Kebun Plasma Nutfah Pisang Pusat Penelitian Biologi-Lembaga Ilmu Pengetahuan Indonesia*. Jakarta: LIPI Press.

Poerba, Y. S. *et al.* (2018) *Deskripsi Pisang Koleksi Pusat Penelitian Biologi LIPI*. Jakarta: LIPI Press.

Prins, R. *et al.* (2005) 'Development of a doubled haploid mapping population and linkage map for the bread wheat cross Kariega × Avocet S Development of a doubled haploid mapping population and linkage map for the bread', *South African Journal of Plant and Soil*, 22(1), pp. 1–8. doi: 10.1080/02571862.2005.10634672.

Rateng', B. (2018) *Genomic models predict shorter time for banana breeding*. Available at: https://www.scidev.net/global/news/shorter-time-for-banana-breeding-1/. (Accessed: 8 April 2021).

Rouard, M. *et al.* (2018) 'Three New Genome Assemblies Support a Rapid Radiation in Musa acuminata (Wild Banana)', *Genome Biol Evol*, 10(12), pp. 3129–3140. doi: 10.1093/database/bat035.

Ruas, M. *et al.* (2017) 'MGIS: managing banana (*Musa* spp.) genetic resources information and high-throughput genotyping data', *Database : the journal of biological databases and curation*, 2017, pp. 1–12. doi: 10.1093/database/bax046.

Sardos, J. *et al.* (2016) '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*, 11(5), p. e0154448. doi: 10.1371/journal.pone.0154448.

Sardos, J. *et al.* (2017) *Report from the exploration of wild banana populations in Papua New Guinea*. Available at: https://hdl.handle.net/10568/89857.

Shepherd, K. (1999) *Cytogenetics of the genus Musa*. Montpellier, France: International Network for the Improvement of Banana and Plantain.

Simmonds, N. W. (1962) *The evolution of the bananas*. London, UK: Longman.

Simonikova, D. *et al.* (2019) 'Chromosome Painting Facilitates Anchoring Reference Genome Sequence to Chromosomes In Situ and Integrated Karyotyping in Banana (Musa Spp .)', *Frontiers in Plant Science*, 10, pp. 1503. doi: 10.3389/fpls.2019.01503.

Stukenbrock, E. H. *et al.* (2007) 'Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation', *Molecular Biology and Evolution*,

24(2), pp. 398–411. doi: 10.1093/molbev/msl169.

Sutanto, A. *et al.* (2014) 'Isolation and characterization of Resistance Gene Analogue (RGA) from Fusarium resistant banana cultivars', *Emirates Journal of Food and Agriculture*, 26(6), pp. 508–518. doi: 10.9755/ejfa.v26i6.17219.

Sutanto, A. *et al.* (2016) 'Collecting banana diversity in eastern Indonesia', *Acta Horticulturae*, 1114, pp. 19–25. doi: 10.17660/ActaHortic.2016.1114.3.

Takken, F. and Rep, M. (2010) 'The arms race between tomato and *Fusarium oxysporum*', *Molecular Plant Pathology*, 11(2), pp. 309–314. doi: 10.1111/j.1364-3703.2009.00605.x.

Thiers, B. M. (2020) *The World's Herbaria 2019: A Summary Report Based on Data from Index Herbariorum*, *The third Index Herbariorum annual report 2016-2018*. Available at: http://sweetgum.nybg.org/science/ih/ (Accessed: 14 April 2021).

Tripathi, J. N. *et al.* (2019) 'CRISPR/Cas9 editing of endogenous banana streak virus in the B genome of *Musa* spp. overcomes a major challenge in banana breeding', *Communications Biology*. Springer US, 2, p. 46. doi: 10.1038/s42003-019-0288-7.

Tripathi, L. *et al.* (2010) 'Expression of sweet pepper Hrap gene in banana enhances resistance to *Xanthomonas campestris* pv . *musacearum*', *Molecular Plant Pathology*, 11(6), pp. 721–731. doi: 10.1111/J.1364-3703.2010.00639.X.

Väre, H., Gogoi, R. and Arisdason, W. (2016) 'Markku Häkkinen, 1946–2015', *Taxon*, 65(4), pp. 938–939. doi: 10.12705/654.52.

Vishnevetsky, J. *et al.* (2011) 'Improved tolerance toward fungal diseases in transgenic Cavendish banana (*Musa* spp. AAA group) cv. Grand Nain', *Transgenic Research*, 20(1), pp. 61–72. doi: 10.1007/s11248-010-9392-7.

Volkaert, H. (2011) 'Molecular analysis reveals multiple domestications of edible bananas', *Acta Horticulturae*, 897, pp. 143–152. doi: 10.17660/ActaHortic.2011.897.15.

Volkaert, H. (2018) 'The origin, domestication and dispersal of bananas', in Kema, G. H. J. and Drenth, A. (eds) *Achieving Sustainable Cultivation of Bananas*. Canbridge, UK: Burleigh Dodds Science Publisihing Limited, pp. 1–18. doi: http://dx.doi.org/10.19103/AS.2018.0020.01.

Wallace, A. R. (1869) *Malay Archipelago: The Land of the Orang Utan and the Bird of Paradise*, *Oxford University*. New York: Harper & Brothers.

Wang, Z. *et al.* (2019) '*Musa balbisiana* genome reveals subgenome evolution and functional divergence', *Nature Plants*, 5(8), pp. 810–821. doi: 10.1038/s41477-019-0452-6.

Westengen, O. T., Jeppson, S. and Guarino, L. (2013) 'Global Ex-Situ Crop Diversity Conservation and the Svalbard Global Seed Vault: Assessing the Current Status', *PLoS ONE*,

6

8(5), p. e64146. doi: 10.1371/journal.pone.0064146.

Wu, W. *et al.* (2016) 'Whole genome sequencing of a banana wild relative *Musa itinerans* provides insights into lineage- specific diversification of the Musa genus', *Scientific Reports*. Nature Publishing Group, 6, p. 31586. doi: 10.1038/srep31586.

Zhang, H. *et al.* (2017) 'Back into the wild — Apply untapped genetic diversity of wild relatives for crop improvement', *Evolutionary Applications*, 10, pp. 5–24. doi: 10.1111/eva.12434.

Summaries

English summary
Nederlandse samenvatting
Ringkasan

English summary

Chapter 1 is an introductory section and describes the aim of this thesis with a focus on providing insight into the rich genetic diversity of wild bananas in the center of diversity to eventually contribute to overall sustainability of global banana production.

Chapter 2 describes an exploration of wild *Musa acuminata* on Sumatra, one of the largest islands of the Indonesian archipelago and the world. In total 164 accessions from 164 locations in eight provinces of Sumatra and of 20 wild banana accessions in the collection of the Research Center for Biology (LIPI), originating from other regions of Indonesia were studied using morphological characteristics and multilocus genotyping. Analysis of the data revealed five types of wild *M. acuminata* on Sumatra. A principle component analysis showed that these were distributed over three major clusters. Based on these data it seems appropriate to merge the previously described var. *halabanesis* (Meijer) Nasution and var. *alasensis* Nasution into one subspecies *halabanesis* (Meijer) Hotta and to consider var. *sumatrana* (Becc.) Nasution as a subspecies. The multilocus genotyping showed a high genetic diversity of *M. acuminata* across Sumatra, often in isolated subpopulations of each subspecies.

Chapters 3 and 4 detail the cytogenetic complexity of bananas by describing the optimization of a chromosome spreading protocol for pollen mother cells and demonstrates its benefits by studying the meiotic features of five wild diploid *M. acuminata* bananas and the diploid (AA) cultivar banana "Rejang", with particular attention on pairing configurations and chromosome transmission. Pollen analyses suggested partial sterility and unreduced gamete formation that likely resulted from restitutional meiotic divisions. How this affects breeding efforts is further described in a literature study in Chapter 4.

In chapter 5, the mapping of a resistance gene to the so-called Race 1 Fusarium strains is described, based on the genotyping and phenotyping of a selfed *M. acuminata* ssp. *malaccensis* population. The combined analyses of these data indicate a single dominant resistance locus near the distal part of chromosome 10. Graphical genotyping revealed molecular markers flanking a 360 kb genetic region associated with the resistance locus. The mapped region contains 165 putative genes on the reference genome, including 19 leucine-rich repeat receptor-like kinase-like genes. These data are the basis for a fine-mapping strategy towards cloning the resistance gene. Chapter 6 is a general discussion, which puts the data of the entire thesis in perspective of the urgently required diversification of the banana crop and how this can be achieved to eventually contribute to a sustainable banana production. This

study supports millions of small holders in their livelihoods and consumers that need banana as a staple crop or fruit snack.

Nederlandse samenvatting

Hoofdstuk 1 is een introducerende sectie en beschrijft het doel van het onderzoek dat in dit proefschrift wordt beschreven met een focus op het verbeteren van het inzicht in de rijke genetische diversiteit van wilde bananen in het centrum van diversiteit, om uiteindelijk bij te dragen aan de verduurzaming van de mondiale bananenproductie.

Hoofdstuk 2 beschrijft een verkenning van wilde Musa acuminata op Sumatra, één van de grootste eilanden in de Indonesische archipel en van de wereld. In totaal, werden 164 accessies afkomstig van 164 locaties in acht provincies van Sumatra bestudeerd, tezamen met 20 accessies uit de collectie van het onderzoekscentrum van LIPI, die afkomstig zijn uit andere Indonesische gebieden. Analyse van de morfologische en multilocus genotypering data onthulde vijf wilde *M. acuminata* soorten op Sumatra, die op basis van een principal component analyse verdeeld zijn over drie grote clusters. Deze data geven aan dat het voor de hand ligt om de voorheen beschreven soorten var. *halabanesis* (Meijer) Nasution en var. *alasensis* Nasution samen te voegen in één ondersoort *halabanesis* (Meijer) Hotta en om te overwegen om var. *sumatrana* (Becc.) Nasution als een ondersoort te beschouwen. De multilocus genotypering laten zien dat M. acuminata op Sumatra heel divers is en vaak voorkomt in geïsoleerde clusters van elke ondersoort.

Hoofdstukken 3 en 4 geven details van de cytogenetische complexiteit van banaan en beschrijven het optimaliseren van een chromosoomspreidingsprotocol voor pollen moedercellen en laten de voordelen ervan zien in het bestuderen van meiotische kenmerken van vijf wilde diploïde *M. acuminata* bananen en de diploïde (AA) cultivar "Rejang", met bijzondere aandacht voor paringsconfiguraties en chromosoom transmissie. Pollen analysen suggereren gedeeltelijke steriliteit en ongereduceerde gameten formatie die waarschijnlijk voortkomen uit meiotische restitutie. Hoe dit de veredeling beïnvloed wordt verder beschreven in hoofdstuk 4 die een literatuurstudie omvat.

In hoofdstuk 5 wordt de kartering van een resistentiegen tegen Race 1 Fusarium stammen beschreven, gebaseerd op de genotypering en fenotypering van een uit zelfbestuiving voortgekomen populatie van *M. acuminata* spp. *malaccensis*. De analyse van deze data geven aan dat er een enkel dominant resistentielocus ligt op het distale gedeelte van chromosoom 10. Grafische genotypering resulteerde in moleculaire merkers in een flankerend gebied van 360 kb dat geassocieerd is met dit locus. De gekarakteriseerde regio bevat 165 veronderstelde genen in het referentiegenoom, inclusief 19 leucine-rich repeat receptor-like kinase-like genen. Deze data zijn de basis voor een fijn-

karteringstrategie om het resistentiegen uiteindelijk te kloneren. Hoofdstuk 6 is de algemene discussie die de resultaten het gehele onderzoek in het perspectief plaatst van de dringende noodzaak tot diversificering van het gewas banaan en hoe dit kan worden bereikt om uiteindelijk bij te dragen aan een duurzame bananenproductie. Dit onderzoek ondersteunt miljoenen kleine boeren in hun bestaan en consumenten voor wie banaan een basis voedsel òf fruit snack is.

Ringkasan

Bab 1 adalah bagian pendahuluan yang menjelaskan tujuan dari disertasi ini, bab ini memberikan gambaran tentang keragaman genetik pisang liar di pusat keragaman dan kontribusinya dalam budidaya dan produksi pisang yang berkelanjutan.

Bab 2 menjelaskan kegiatan eksplorasi pisang liar *Musa acuminata* di Sumatera, salah satu pulau terbesar di Indonesia dan juga di dunia. Secara keseluruhan 164 aksesi dari 164 lokasi pengambilan sampel di delapan provinsi di Sumatera ditambah dengan 20 pisang liar milik koleksi Pusat Penelitian Biologi-LIPI yang berasal dari berbagai tempat di Indonesia dipelajari keragaman genetikanya berdasarkan karakter morfologi dan molekuler menggunakan metode *multilocus genotyping*. Berdasarkan analisa karakter morfologi, di Sumatera terdapat lima tipe *M. acuminata*, sedangkan hasil analisa *principle component analysis* menunjukkan bahwa keragaman pisang di Sumatera terbagi dalam tiga kelompok utama. Berdasarkan data-data ini, dua varietas yaitu var. *halabanensis* dan var. *alasensis* selayaknya digabungkan menjadi satu subspesies *halabanensis* (Meijer) Hotta dan mempertimbangkan var. *sumatrana* (Becc,) Nasution sebagai subspesies. Analisa *multilocus genotyping* menunjukkan tingginya keragaman genetika *M. acuminata* di Sumatera dan menunjukkan adanya subpopulasi dari tiap subspesies yang ada.

Bab 3 dan 4 menjabarkan detail dari kerumitan sitogenetika pada pisang, dalam bab ini dijelaskan optimalisasi protokol dalam pembuatan preparat kromosom dari sel induk serbuk sari dan menunjukkan manfaatnya untuk studi meiosis, terutama pada saat terjadinya pasangan kromosom dan hasil pembelahannya, lima pisang liar *M. acuminata* dan satu pisang diploid (AA) "Rejang". Analisa polen menunjukkan adanya sterilitas sebagian dan formasi gamet yang tidak tereduksi (*unreduced gamete*) hasil dari pembelahan meiosis tanpa pengurangan tingkat ploidi (*restitutional meiotik division*). Selanjutnya, bagaimana hal-hal ini mempengaruhi usaha pemuliaan pisang pada masa yang akan datang dijabarkan dalam studi literatur di bab 4.

Di dalam bab 5, dijelaskan pemetaan genetik gen ketahanan terhadap Fusarium strain Race-1 berdasarkan karakterisasi genotipe (*genotyping*) dan fenotipe (*phenotyping*) dari populasi hasil penyerbukan sendiri *M. acuminata* ssp. *malaccensis*. Analisa dari dari dua data ini mengindikasikan adanya satu lokus ketahanan yang bersifat dominan yang terdapat pada ujung kromosom 10. *Graphical genotyping* menunjukkan bahwa jarak dua marka molekuler yang mengapit area yang terkait dengan sifat ketahanan tersebut adalah 360 kb. Dalam area ini terdapat 165 *putative genes* pada *reference genome* yang 19 diantaranya merupakan *leucine-rich repeat receptor-like*

kinase-like genes. Data ini merupakan dasar untuk pemetaan lanjut menuju kloning gen ketahanan tersebut.

Bab 6 adalah pembahasan, dalam bab ini data-data dalam disertasi digunakan untuk membuka sudut pandang pentingnya penganekaragaman pisang dan usaha-usaha untuk mencapainya yang pada akhirnya menunjang produksi pisang yang berkelanjutan. Produksi pisang yang berkelanjutan akan menopang jutaan petani kecil di tempat tinggalnya dan konsumen yang membutuhkan buah pisang sebagai makanan.

About the author

About the author

Fajarudin Ahmad was born on December 25, 1979, in Yogyakarta, Indonesia where he completed his primary and secondary education. In 2003, he finished his BSc degree in Botany with a major in Plant Genetics at the University of Gadjah Mada, Yogyakarta. As a fresh graduate, he got a job in a shrimp company, PT Central Pertiwi Bahari, Lampung, with the duty of virus diseases detection from 2003 to 2006. He was then recruited by the Research Center for Biology and since then works in the Plant Genetics and Breeding Laboratory. To continue his training, he obtained a scholarship of the Ministry of Research and Technology and finished his MSc degree in Plant Genetics with a major in Plant Biology at the IPB University in 2013. His MSc thesis focused on the genetic diversity of the Indonesian wild banana *Musa balbisiana* Colla. To further his studies, he joined the group of Prof. Dr G.H.J. Kema as a PhD student in 2014 under the Scientific Program Indonesia Netherlands of the Royal Netherlands Academy of Arts and Sciences (KNAW). His research project focused on the genetic diversity of Indonesian bananas with a focus on cytogenetics and resistance to Fusarium wilt.

Acknowledgments

Acknowledgments

Doing a PhD was just like an exhausting journey, sometimes I took a breath that made me stop for quite some time, but all stage of the work is full of excitement. It is obvious that I would not have survived without all people who contribute to my PhD during the past seven years.

First of all, I would like to offer my sincere gratitude to my both promotors and supervisors **Prof. Dr GHJ (Gerrit) Kema** and **Prof. Dr JHGSM (Hans) de Jong**. Your both contribution are equal to my PhD. Dear **Gert**, thank you for entrusting me to be part of KNAW-SPIN project to pursue my PhD project. You brought me to the international academic community that change my mind perspective of research. I debt of your times in supervising me from the scratch of my thesis until the complete shape of story. Dear **Hans**, thank you for your guidance that full of wisdom and always be my side in all situation to identify the problem and get the silver lining. You though me the techniques of slide preparation to scientific artwork pieces by pieces that make me can learn more about genetics.

To my co-promotor **Dr Ir. HJ (Henk) Schouten**, thank you for introduce me the philosophy of genetic and mapping. You make me realize the challenge and hope in saving the plant against diseases.

I would also express my gratefulness to my external supervisor **Dr Hugo A. Volkaert**. Dear **Hugo**, thank you for your kindness and warm welcome to your laboratory. Thank you for helping me to get a better understanding of banana diversity and much more appreciate our talent as human being.

A special thankful for **Dr Yuyu S Poerba**, you just like my parent in Laboratorium Plant Genetic at Cibinong, you encouraged me since the beginning of my PhD project until the end of the program. You are the best friend in my all banana explorations in Indonesia.

I would like to express my appreciation to **Dr Witjaksono**, you have opened the door for an international experience and prepared all fundamental knowledge on banana that essential for me to begin this project.

I would like to thanks **Prof. Dr Siti Subandiyah**, the Indonesian project leader of the KNAW-SPIN project. I am pleased to all your trust to me to be part of this project, your supports made lab works and exploration in Indonesia smoothly done.

Diatas semua rasa terima kasih, saya haturkan untuk keluargaku. Istriku, Nurul, terima kasih atas kesabaran dan cinta yang selalu kamu sediakan untukku dan anak-anak kita, terutama kesabaran di tiga tahun pertama program PhD-ku ini. Untuk Ramiy, anakku paling ganteng dan peka, dan Zada, anakku yang paling cantik dan kritis, kalian adalah bagian motivasi hidupku untuk selalu bersemangat dan berpikir positif. Untuk Ibu, terima kasih doanya yang engkau panjatkan setiap hari untuk anak-anakmu, aku yakin tidak akan bisa seperti saat ini tanpa restumu. Buat Bapak, semoga Bapak mendapat tempat yang tebaik di sisi Allah SWT, terima kasih atas ajaran dan nasehatmu sejak aku masih kecil, aku yakin ini adalah modal dasar untuk menjalani hidupku sekarang.

Education Statement of the Graduate School

Experimental Plant Sciences

Issued to: **Fajarudin Ahmad**
Date: **23 August 2021**
Group: **Laboratory of Genetics & Laboratory of Phytopathology**
University: **Wageningen University & Research**

1) Start-Up Phase	*date*
► **First presentation of your project**	
Cytogenetics of banana and its implications for breeding for resistance to Fusarium wilt disease	3 Feb 2015
► **Writing or rewriting a project proposal**	
Cytogenetics of banana and its implications for breeding for resistance to Fusarium wilt disease	20 Oct 2015
► **Writing a review or book chapter**	
► **MSc courses**	
Population and Quantitative Genetics (GEN-30806)	May- Jun 2015

Subtotal Start-Up Phase

2) Scientific Exposure	*date*
► **EPS PhD student days**	
EPS PhD student days (GET2GETHER)	29-30 Jan 2015
EPS PhD student days (GET2GETHER)	09-10 Feb 2017
► **EPS theme symposia**	
EPS Theme 4 Symposium 'Genome Biology', Wageningen, the Netherlands	03 Dec 2014
EPS Theme 2 Symposium 'Interactions between Plants and Biotic Agents' and Willie Commelin Scholten Day, Amsterdam, the Netherlands	24 Jan 2018
EPS Theme 1 Symposium 'Developmental Biology of Plants', Wageningen, the Netherlands	30 Jan 2018
EPS Theme 3 Symposium 'Metabolism and Adaptation', Wageningen, the Netherlands	13 Mar 2018
► **Lunteren Days and other national platforms**	
Annual Meeting 'Experimental Plant Sciences', Lunteren, the Netherlands	11-12 Apr 2016
Annual meeting 'Experimental Plant Sciences', Lunteren, the Netherlands	09-10 Apr 2018
► **Seminars (series), workshops and symposia**	
Symposium: All-inclusive Breeding: Integrating high-troughput science, Wageningen, the Netherlands	16 Oct 2014
Symposium: 29th Meeting of the Fusarium working group of the KNPV, Utrecht, the Netherlands	29 Oct 2014
Symposium: 30th Meeting of the Fusarium working group of the KNPV, Utrecht, the Netherlands	26 Oct 2015
Symposium: 31st Meeting of the Fusarium working group of the KNPV, Utrecht, the Netherlands	26 Oct 2016
Symposium: Wageningen Indonesia Scientific Expose 2016, Wageningen, the Netherlands	28 Oct 2016
Symposium: Wageningen Indonesia Scientific Expose 2017, Wageningen, the Netherlands	08 Mar 2017
Symposium: 32nd Meeting of the Fusarium working group of the KNPV, Utrecht, the Netherlands	25 Oct 2017
Symposium: Host-Microbes Genetics Meeting, Wageningen, the Netherlands	27 Oct 2017
► **Seminar plus**	
► **International symposia and congresses**	
1st International Conference on Biodiversity, food Security and Health, Yogyakarta, Indonesia	22-23 Nov 2016
Plant and Animal Genomic XXVI Conference, San Diego, USA	13-17 Jan 2018
6th Plant Genomics & Gene Editing Congress, Rotterdam, the Netherlands	14-15 May 2018
International Congress of Plant Pathology (ICPP) 2018, Boston, USA	29 Jul - 03 Aug 2018
► **Presentations**	
Talk: 1st International Conference on Biodiversity, Food Security and Health, Yogyakarta, Indonesia	22-23 Nov 2016
Talk: Wageningen Indonesia Scientific Expose 2017, Wageningen, the Netherlands	08 Mar 2017
Talk: Host-Microbes Genetics Meeting, Wageningen, the Netherlands	27 Oct 2017
Talk: Plant and Animal Genomic XXVI Conference, San Diego, USA	16 Jan 2018
Talk: Annual meeting 'Experimental Plant Sciences', Lunteren, the Netherlands	09-10 Apr 2018
Talk: International Congress of Plant Pathology (ICPP) 2018, Boston, USA	28 Jul 2018
Poster: Wageningen Banana Day, Wageningen, the Netherlands	18 Nov 2014
Poster: Annual meeting 'Experimental Plant Sciences', Lunteren, the Netherlands	11-12 Apr 2016
Poster: Wageningen Indonesian Scientific Expose (WISE) 2016, Wageningen, the Netherlands	28 Oct 2016
Poster: Wageningen Indonesia Scientific Expose 2017, Wageningen, the Netherlands	8 Mar 2017
Poster: Plant and Animal Genomic XXVI Conference, San Diego, USA	13-17 Jan 2018
Poster: 6th Plant Genomics & Gene Editing Congress, Rotterdam, the Netherlands	14-15 May 2018
► **IAB interview**	
► **Excursions**	
EPS PhD council company visit to breeding company 'Enza Zaden', Enkhuizen, the Netherlands	13 Jun 2015
EPS PhD council company visit to tomato company 'Tomato World', Honselersdijk, the Netherlands	14 Oct 2016

Subtotal Scientific Exposure

3) In-Depth Studies	*date*
► **Advanced scientific courses & workshops**	
COMREC Bioinformatics Workshop, Wageningen, the Netherlands	04-06 Feb 2015
International workshop on Musa Cytogenetics and Molecular Taxonomy, Kamphaeng Saen, Thailand	14-16 Aug 2019
► **Journal club**	
A weekly journal club/meeting group, every Wednesday 12pm-1pm	2015-2018
► **Individual research training**	
Cytogenetic and molecular training, Kasetsart University, Kamphaeng Saen, Thailand	03-31 Aug 2014

Subtotal In-Depth Studies

4) Personal Development	*date*
► **General skill training courses**	
EPS Introduction course, Wageningen, the Netherlands	20 Jan 2015
Advanced Course Guide to Scientific Artwork, Wagenigen, the Netherlands	23-24 Mar 2015
Scientific Writing, Wageningen, the Netherlands	Apr - Jun 2015
Project and Time management, Wageningen, the Netherlands	Sep - Oct 2015
Brain Training, Wageningen, the Netherlands	20 Sep 2016
► **Organisation of meetings, PhD courses or outreach activities**	
Organizing Committee, Wageningen Indonesian Scientific Expose (WISE) 2016, Wageningen, the Netherlands	28 Oct 2016
Organizing Committee, Wageningen Indonesian Scientific Expose (WISE) 2017, Wageningen, the Netherlands	08 Mar 2017
► **Membership of EPS PhD Council**	

Subtotal Personal Development

TOTAL NUMBER OF CREDIT POINTS*
Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS minimum total of 30 ECTS credits.
* A credit represents a normative study load of 28 hours of study.

Propositions belonging to the thesis, entitled

Genetics and diversity of Indonesian bananas

1. The future of banana improvement lies in exploring and exploiting wild germplasm from the centre of diversity, and not in mutation breeding or genome editing.
   (this thesis)

2. Efficient breeding of banana requires a thorough understanding of numerical and structural chromosome variants between the crossing parents.
   (this thesis)

3. Artificial intelligence can handle large genome data sets to predict genes and proteins, but always needs human intelligence.

4. Implementing any regulation for sustainable food production should consider the sustainability of the natural ecosystem.

5. Food not only connects people, it also has the ability to destroy societies.

6. Life science scientists could learn more about human life via TikTok.

Fajarudin Ahmad
Wageningen, 23 August 2021