Diversity and origins of an underutilized tropical fruit tree, cempedak (*Artocarpus integer*, Moraceae), and its wild relative bangkong (*Artocarpus integer* var. *silvestris*)

A THESIS SUBMITTED TO THE FACULTY OF THE PROGRAM IN PLANT BIOLOGY AND CONSERVATION

BY MARIA WANG MEI HUA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN PLANT BIOLOGY AND CONSERVATION FROM NORTHWESTERN UNIVERSITY AND THE CHICAGO BOTANIC GARDEN

November 20, 2014

COMMENTS ON THESIS DEFENSE FOR MASTER OF SCIENCE IN PLANT BIOLOGY AND CONSERVATION NORTHWESTERN UNIVERSITY AND THE CHICAGO BOTANIC GARDEN

The student Hana Wang Mei Hua defended his/her thesis, titled Diversity & origins of an underutilized tropical fruit tree complex (Artocarpus integer) & its wild relatives in Hong Kong on the date of 20th Nov 2014. The committee members have rendered the following decision:

☒ Pass unconditionally

☐ Pass conditionally and requires no re-examination (conditions are detailed on the next page)

☐ Requires re-examination (conditions are detailed on the next page)

☐ Fail to pass (comments are attached on the next page)

Committee members:

Printed name	Signature	Department/Affiliation
Nyree Zerega Major advisor		PBC
JEREMIE FANT		CBG/PBC
Andrea Kramer		CBG/PBC

In the case of revisions or re-examination, specify details on the next page. Once all requirements have been met, the major advisor and program director should sign below.

Signature of major advisor confirming final approval — Date 11/20/14

This form, along with a pdf file of the final approved thesis, must be submitted to the PBC program assistant and the program director, at which point the latter will sign below and approval will be communicated to TGS.

Plant Biology and Conservation Program Director Signature — Date 11/20/14

Abstract

The underutilized tropical fruit tree, "cempedak" (*Artocarpus integer* (Thunb.) Merr. , Moraceae), is a major crop in Malaysia but the distribution of genetic diversity in the crop and its putative wild progenitor, "bangkong" (*Artocarpus integer* var. *silvestris* Corner), have not been studied. To investigate the diversity and origins of cempedak, we assessed the genetic diversity of cempedak and bangkong, explored geographic patterns in genetic structure and differentiation, and conducted phylogenetic analyses. We also tested whether cempedak cultivars were true-to-type. We used 12 microsatellite loci to genotype 269 cempedak and 187 bangkong individuals from Peninsular Malaysia, 64 cempedak individuals from Borneo, and 12 cempedak samples from other regions. We found that the number of effective alleles (Ne), allelic richness (Ar), Shannon's Information Index (I), and expected heterozygosity (He) were higher in bangkong populations than cempedak populations, but cempedak populations harbored alleles not found in bangkong populations. In contrast, Borneo cempedak showed overall lower genetic diversity compared to Peninsular cempedak. Both bangkong and cempedak exhibited weak geographic structure within Peninsular Malaysia; but we found evidence for isolation-by-distance among bangkong populations, as well as increased genetic structure in cempedak at a broader geographic scale. Borneo cempedak represented a gene pool distinct from Peninsular cempedak. Results of phylogenetic analyses varied based on the methods used, but indicate that bangkong or Borneo cempedak may be possible progenitors of Peninsular cempedak. Lastly, clonal analysis showed that cempedak cultivar names were not always indicative of genotype, which may be due to multiple reasons such as mislabeling, hybridization, and scoring errors. Although we did not find conclusive evidence regarding the origins of cempedak, this study confirmed that bangkong and Borneo cempedak populations harbor unique genetic variation that should be conserved. We hope that the results of this study will inform further research and conservation strategies for genetic resources of cempedak and other fruit crops in Malaysia and beyond.

Acknowledgements

This research project was funded by the Plant Biology and Conservation Award from Northwestern University, the Botanical Society of America (BSA) Graduate Student Research Award, and a National Science Foundation grant awarded to my advisor, Nyree Zerega (NSF DEB REVSYS 0919119**)**. My Masters education was funded in part by the Shaw Fellowship in Plant Biology and Conservation.

I would like to thank my advisor Nyree Zerega, for being an all-round supportive, fantastic, and super-woman advisor and my thesis committee, Jeremie Fant (for his patience with my endless questions in the lab and during data analysis), and Andrea Kramer (for her kindness and apt advice) for all their valuable advice and help in this project. I would like to sincerely thank my Malaysian collaborators: the Economic Planning Unit (EPU) of Malaysia, Forestry Department of Peninsular Malaysia, Department of Orang Asli Development (JAKOA) for enabling this project; personal thanks to Dr. Richard Chung, Miss Chew Ming Yee, Dr. Lee Soon Leong, Angan, Apok, and others at FRIM; Dr. Abdul Rahman Milan, Dr. Salma Idris, Pn. Noor Baiti, En. Khadzir Darus, En. Mat Nazim, En. Mohd. Noor, En. Ai, En. Mohd. Azram, and others at MARDI; Pn. Zaliza and others at UPM; Dr. Ong Hean Chooi from USM for advice early on in the project; the Orange Asli communities, especially Norli (Sungai Buloh), Insan (Ulu Groh), whose knowledge of the forests is indispensable to forest research and conservation; Orang Asli guides Norli (Sungai Buloh), Insan (Ulu Groh); officers at JAKOA Jeli and Ayer Hitam; all of the cempedak tree/orchard owners and guides (Uncle Lim Boon Huat, Mr. Gan, Mr. Cheng, Aunty Michelle, Mr. James ‘Ah Fook’, Liew brothers, Mr. Lim from Balik Pulau, and others I have not mentioned), friends and family who have provided information, contacts, and assistance with travel/accommodation, with special thanks to my parents and their whole-

hearted support that was crucial to the success of this project. With all my heart, I would like to thank all the wonderful members of the Zerega lab who have been awesome in big and little ways, especially Elliot Gardner who deserves a gold star award for his help in various aspects of the project including sequencing, providing R code, and helping obtain additional leaf material from herbarium specimens; members of the CBG genetics lab; members of the Eco-Climate lab at SCBI; CBG folks and the trusty PACE 213 bus for ferrying me to and from CBG; and everyone who have helped in any way to make this project possible – again, *Ribuan Terima Kasih*!!!

Table of Contents

List of Figures

List of Tables

1. Introduction

Plant biodiversity provides food, raw materials, medicines, and many other ecosystem services that are crucial to life on earth. Between 300,000 to 400,000 plant species are predicted to be found on earth (Chapman, 2009; Mora et al., 2011). Out of an estimated 30,000 edible plants, only about 7,000 species have been cultivated for food at some point in history (Wilson, 1999). Only about 103 species provide 90% of the global plant food supply (Prescott-Allen and Prescott-Allen, 1990), and only 3 species (rice, corn and wheat) provide more than 50% of the plant-sourced calories and proteins in the human diet (FAO, 1991). A recent study by Khoury et al. (2014) demonstrated an increase in global homogeneity composition of food supplies and global dependence on 50 crop commodities or an estimated 94 species. This represents a large untapped potential of plant genetic resources, including genetic diversity, which must be preserved as it provides the tools for species to adapt to future challenges, such as disease and climate change (Harlan, 1976; Altieri and Merrick, 1987). Commercialization of crops, dependence on a few select uniform, high-yielding varieties, and subsequent loss of landraces have led to genetic erosion in cultivated crops (Gepts, 2006; FAO, 2010). For example, in India, many local varieties of rice in Orissa and Chattisgarh, medicinal rice in Kerala, and various millet species in Tamil Nadu are no longer cultivated in their native habitats or have been replaced by modern varieties (Chaudhuri, 2005). In Italy, almost 400 botanical varieties of wheat were lost in 45 years from 1945 to 1990 (Hammer and Laghetti, 2005). More recently, Dyer et al. (2014) found that varietal richness of corn in its center of origin, Mexico, had decreased nationwide from 1.43 to 1.22 varieties per farm between 2002 and 2007.

Genetic erosion may lead to crop failure, which has severe consequences for the economy, people's livelihoods, and human health (Heal et al., 2004). This is most evident when a pathogen attacks a genetically homogenous commercial variety of a food crop that is grown on

a large scale. For example, the 1970 southern corn leaf blight epidemic in the United States caused serious losses in the South and Mid-West, due to the widespread genetic homogeneity of the corn crop (Ullstrup, 1972). About 85% of hybrid seed corn grown in the United Sates had been bred to contain Texas cytoplasm male sterility (Tcms), a trait which caused the corn plants to be hyper-susceptible to a new race of southern corn leaf blight, resulting in an overall loss of one billion USD in the United States (Ullstrup, 1972). The consequences would be even graver if such a massive crop failure occurred in nations whose agriculture depends on only a few crops. For example, when Ireland depended solely on the potato as a staple crop, the potato blight epidemic in the 1840s resulted in great famine and the suffering and loss of many lives (Moore and Tymowski, 2005). Since then, blight-resistance genes have been found in wild relatives of potato from South America, the center of diversity of potato; and more recently, transgenically introduced into a potato cultivar (Jones et al., 2014).

This example demonstrates the importance of identifying and conserving wild relatives of crops for food security. Wild relatives, including progenitors of domesticated crops and other more or less closely related species, provide valuable genetic resources for farmers and plant breeders to improve cultivars for higher yield, greater tolerance to drought, flooding, and climate change, or to maintain pest and disease resistance (Harlan, 1976; Lane and Jarvis, 2007). Wild relatives have long been used to improve crops, and they make significant contributions to agro-economies by providing resistance to biotic and abiotic stresses; for example, wild relatives contributed more than $340 million/year in yield and quality improvements to US-grown or imported crops (Prescott-Allen and Prescott-Allen, 1986). However, the genetic resources of wild relatives are also under threat of extinction and climate change. Expecting this, scientists have long called for targeted efforts to identify and locate wild relatives around the world, as

well as *ex situ* and *in situ* conservation of known populations (Hoyt, 1988; Meilleur and Hodgkin, 2004). In addition to conserving wild relatives, crop diversification is equally important to food security. For example, growing underutilized, native crops which are adapted to local conditions, can decrease the carbon footprint of agriculture and contribute to food security in developing nations. Many conventional crops like wheat, corn and rice are annuals, which require heavy inputs of water and inorganic fertilizers (FAO, 2012). In contrast, perennial crops like trees and locally adapted plants require less input and can be grown more sustainably. In short, increased efforts to identify, study, and promote wild relatives and underutilized crops are critical to preserve genetic resources and maintain crop diversity.

1.1 Malaysia

Malaysia is a developing country in Southeast Asia, consisting of Peninsular Malaysia and the states of Sabah and Sarawak on the island of Borneo ('East Malaysia'), separated by the South China Sea (Figure 1). Peninsular Malaysia borders Thailand to the north and Singapore to the south, while Sabah and Sarawak border Brunei and Indonesia. Malaysia has a total land area of 32.86 million ha. Situated between 2º and 7º north of the equator, Malaysia experiences an equatorial climate, with southwest (April - October) and northeast (October to February) monsoons. Humidity is high all year round and annual rainfall is 2000 - 2500 mm.

Malaysia has a multi-racial, multi-cultural population of 30.07 million, including Malay, Chinese, Indian, Dayak, Kadazan (Dusun), Bajau, Melanau, Murut and various other ethnic groups and indigenous people (Central Intelligence Agency, n.d.)Agriculture plays a major role in the national economy, with rice, oil palm, and rubber making up the largest percentage of cultivated land (Nordin et al., 2007). Crops are produced by smallholders, government-aided land development schemes such as the Federal Land Development Authority (FELDA), and

large commercial estates or plantations. Research and development on industrial and food crops are actively conducted by government institutions such as the Malaysian Agricultural Research and Development Institute (MARDI), universities, and private companies. Fruit trees and cash crops are also commonly planted in home-gardens and villages for personal consumption or for sale at the local market.

Malaysia is party to the Rio Convention on Biological Diversity (CBD), and is among the mega-diverse countries in the world with high levels of endemism (CBD, 2014). The Malaysian rainforest holds diverse species of wild trees with edible fruits and wild relatives of cultivated species (Saw et al., 1991). For example, the 50-ha primary rainforest of Pasoh contains 12 wild species of mango (*Mangifera*, Anacardiaceae), 13 wild species of mangosteen (*Garcinia*, Clusiaceae), and five wild species of rambutan (*Nephelium*, Sapindaceae) (Saw et al., 1991). Malaysia has pledged to maintain at least 50% of its land under forest cover at the Rio Earth Summit in 1992, and has established a number of policies and acts to manage natural resources (Ministry of Natural Resources and Environment of Malaysia, 2014). Approximately 44% of the land area on the Peninsula is still forested (Forestry Department of Peninsular Malaysia, 2013). In Sabah and Sarawak, about 53% and 38% of the land area are respectively designated as Permanent Forest Estate (Sabah Forestry Department, 2012; Sarawak Forestry Department, 2012). At the same time, satellite-based remote sensing studies showed that "hotspots" of forest clearing were found in every state of Malaysia from 2000 to 2005, and that the country has experienced the highest percent net loss of forest cover in the world from 2000 to 2012, and (Hansen et al., 2008, 2013). Hence, serious efforts must be taken to sustain and effectively conserve forest areas and the biodiversity within these areas.

In summary, the coexistence of homegardens and commercial agriculture, active research institutions, and large areas of forests harboring many edible wild species, makes Malaysia a valuable study system of the genetic diversity of cultivated crops and their wild relatives.

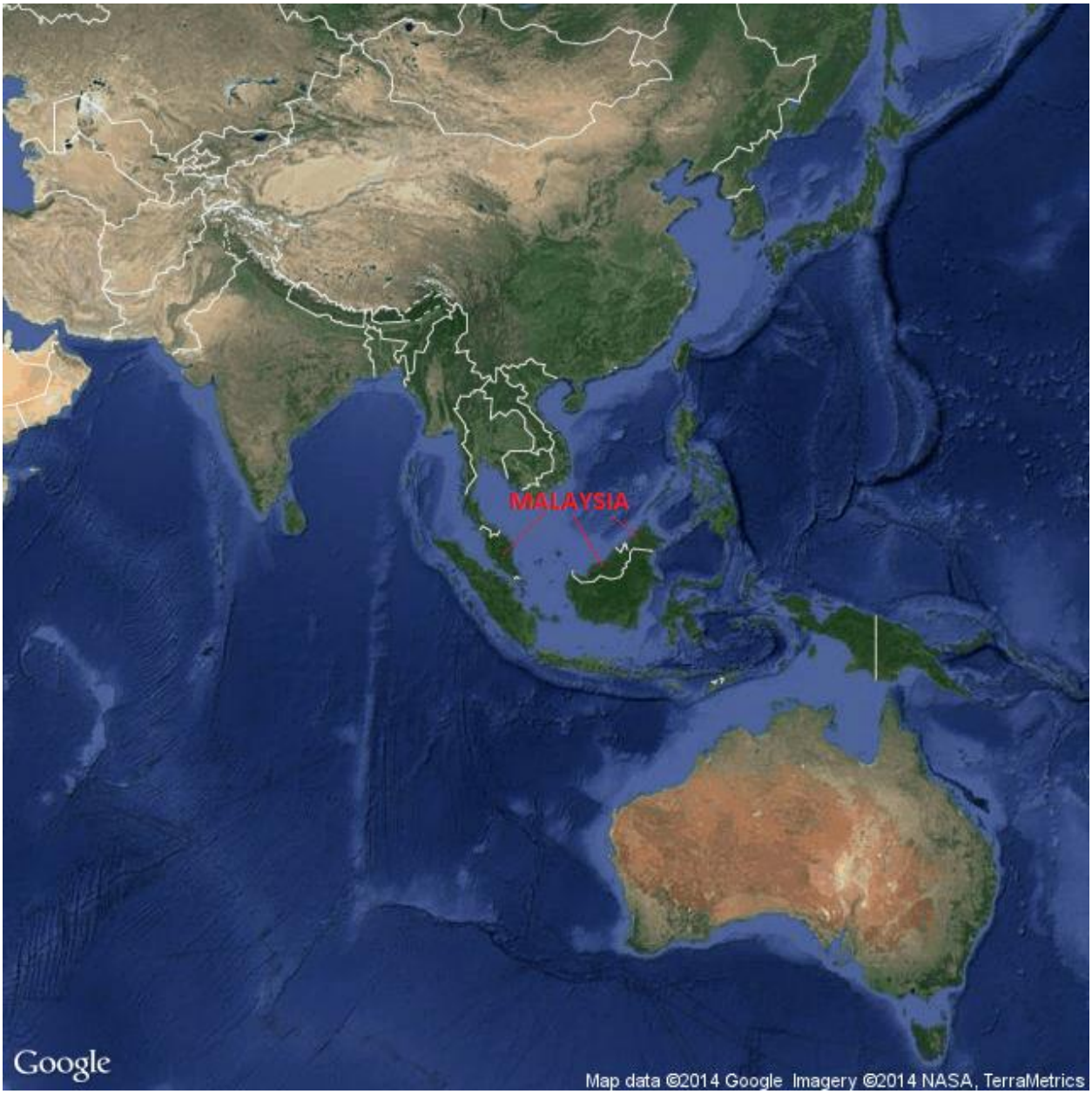


Figure 1. Location of Peninsular Malaysia and the states of Sabah and Sarawak on the island of Borneo ('East Malaysia'), separated by the South China Sea.

1.2 Study Species

Cempedak (*Artocarpus integer* (Thunb). Merr.) is a member of the Moraceae family, which contains many ecologically and economically important species, including figs, mulberries, and breadfruit (Zerega et al., 2010). The genus *Artocarpus* contains nearly 70 species, which are native to South and Southeast Asia (Zerega et al., 2010). More than a dozen *Artocarpus* species are cultivated throughout this range, and many of them also occur naturally in Malaysian rainforests, some being endemic to the region (Jarrett, 1959, 1960; Saw et al., 1991; Zerega et al., 2010). Cempedak is also known as chempedak, campedak, sempedak, temedak (Malaysia); Bukah (Iban); champedak, chepedak, chubadak, kakan, buda (Indonesia); campedak, comedak (Javanese); baroh (Lingga); champada (Thai); sonekadat (Burma); mit to nu (Vietnamese); kathal, kathar (Hindi); and chakka, pilual (Tamil) (Jansen, 1991; Janick and Paull, 2008). It is distributed across Southeast Asia and New Guinea (Janick and Paull, 2008). It is cultivated in those areas, southern India, and in Australia, but is most popular in Malaysia, southern Thailand, and some regions of Indonesia (Janick and Paull, 2008).

Cempedak is an evergreen, monoecious, medium-sized tree, reaching 10-20 m in height (Jansen, 1991; Yaacob and Subhadrabandhu, 1995). The bark is smooth and grey-brown when young, becoming thick and rough with age (Yaacob and Subhadrabandhu, 1995). The trunk and main branches are often bumpy, indicating where fruit-bearing twigs emerge (Jansen, 1991). The tree exudes sticky white latex when cut, including the trunk, twigs, and fruits (Yaacob and Subhadrabandhu, 1995). Leaves are simple, spirally alternate, stipulate, and entire; obovate to elliptic; dull to medium green in color; with characteristic brown wiry hairs on leaf surfaces and twigs (Jansen, 1991; Yaacob and Subhadrabandhu, 1995). Inflorescences are solitary in leaf axils, cauliflorous or ramiflorous (Jansen, 1991). Fruits are very large multiple aggregate syncarps, varying in shape from cylindrical to clavate or ellipsoid, usually 20-40 cm long and 10-

15 cm in diameter (Jarrett, 1959; Jansen, 1991). The outer rind consists of fleshy spines, and is green when unripe, gradually turning dark yellow-brownish green when ripe. The fleshy perianth tissue and the achenes, commonly referred to as "seeds", are both edible. Cempedak is thought to be pollinated by gall midges that feed and breed on a fungus that grows on the male inflorescences (Sakai et al., 2000). Interestingly, figs are also used as a brood site by their pollinating fig-wasps. Cempedak saplings typically start bearing fruit after 3-6 years, while clonal trees that have been grafted onto mature rootstock can bear fruit at about 2-4 years old. The seasonality of cempedak varies, but the main harvest is between June and September in Peninsular Malaysia, and towards the end of the year in Borneo (Jansen, 1991; pers. obs.). Most cempedak pests and diseases also affect jackfruit (Janick and Paull, 2008). Cempedak fruits are mostly attacked by fruit flies, and sometimes by the larvae of a fruit-boring moth (*Margaronia* caesalis); but these pests can be controlled by bagging fruits (Jansen, 1991; Department of Agriculture of Peninsular Malaysia, 2001). Immature fruit rot caused by *Rhizopus artocarpi* can be controlled with fungicide (Department of Agriculture of Peninsular Malaysia, 2001). Boring beetles (*Apriona flavescens)* may also attack the bark and stem of the trees (Jansen, 1991; Department of Agriculture of Peninsular Malaysia, 2001). Bacterial dieback caused by (*Erwinia carotovora*) is the most serious disease affecting cempedak, which first infects the growing shoots but rapidly spreads to the main trunk and kills the tree (Jansen, 1991; Department of Agriculture of Peninsular Malaysia, 2001).

1.3 Uses of cempedak

Cempedak is a major crop in Malaysia, with a total planted area of 7,297 ha and production value of MYR 82 million (USD 25 million) in 2013 (Ministry of Agriculture and Agro-Based Industry of Malaysia, 2014). However, it is underutilized in most other places. This may be because cempedak is tricky to harvest at the right time, difficult to assess the quality of unopened whole fruits, and poses challenges for shipping and export. Cempedak is less known than its sister species, jackfruit, as one can infer from the lack of an English name for cempedak. Cempedak has been reported to hybridize with jackfruit, and cempedak-jackfruit hybrid cultivars are commercially available (Jansen, 1991; Yaacob and Subhadrabandhu, 1995). Additionally, cempedak is commonly used as a rootstock for jackfruit grafts, and apparently performs better than jackfruit rootstock (Yaacob and Subhadrabandhu, 1995).

Cempedak has a strong characteristic smell when ripe, described as harsh and penetrating like durian, or musky and sweet like overripe jackfruit (Corner, 1938). Cempedak is harvested primarily for its fleshy perianth, which is typically yellow or orange in color and rich in vitamin A, but some varieties range from cream to pinkish to orange-red (Primack, 1985; Jansen, 1991; also see Appendix 7). The flesh ranges from fibrous to soft and custardy in texture; while the taste is sweet and reminiscent of durian and mango (Jansen, 1991). Because of the strong aroma and flavor, cempedak is sometimes described as a cross between durian and jackfruit. The pulpy fleshy perianth is usually eaten raw or dipped in batter and fried, but can also be creamed to make jams, cakes, and ice cream. The hard achenes or "seeds" (hereafter, simply referred to as "seeds") can be boiled in salty water or roasted, or ground into flour (Janick and Paull, 2008). Chempedak achene flour is a good source of dietary fiber and may be used as a fiber-rich partial substitute for wheat flour in baking, and has successfully been incorporated in bread (Aziz and Zabidi, 2011). Young fruits and leaves can be cooked as vegetables, similar to jackfruit (Jansen,

1991). Cempedak, which is rich in carbohydrates and low in fat, has potential to be processed into various food products like flour, desserts, snacks, and candy, which can attract a wider audience/population to increase its utilization and marketability (Tee et al., 1988; Department of Agriculture of Peninsular Malaysia, 2001). Potential biomedically useful compounds have been isolated from cempedak plant parts and seeds; for example, an antimalarial stilbene (Boonlaksiri et al., 2000) and a human immunoglobulin-reactive lectin (Lim et al., 1997; Abdul Rahman et al., 2002). The timber is dark yellow to brown and strong and durable. It can be used for construction, furniture, and boats (Jansen, 1991). Ropes can be made from the bark and the latex can be used to make birdlime for catching birds (Jansen, 1991).

1.4 Wild relative of cempedak: Bangkong

"Bangkong" is a wild relative of cempedak, which was first described by Corner (1938) as *Artocarpus integer* var. *silvestris*. Bangkong is most commonly found in primary and secondary forests of Peninsular Malaysia, although it has been recorded in Sumatra, Lingga Archipelago, and South/Southeast Borneo (Jarrett, 1959). Bangkong is very difficult to differentiate from cempedak when not in fruit. Morphological differences between cempedak and bangkong include variation in hairiness, color of withering leaves, size of male inflorescences, size of fruits, but most notably the lack of fragrance and insipid taste of the bangkong fruit (Table 1; Appendix 7). The common name, "bangkong", is sometimes also used to refer to other *Artocarpus* species that grow in the forest, or used in combination with other names, e.g. "cempedak-bangkong", or "keledang-bangkong". Other common names for bangkong include "cempedak hutan" and "cempedak ayer", meaning "forest cempedak" and "water cempedak". The indigenous tribes of Peninsular Malaysia, collectively known as the Orang Asli, also have their own names for bangkong, including temenguk (Temuan), dekoh (Semai), and tekik (Jahai) (pers. obs.). In contrast with the cultivated cempedak, the bangkong fruit is harvested only for its seeds while the fleshy perianth is not eaten. The seeds are boiled or roasted, and taste like potato or chestnuts. It is an important part of the Orang Asli diet, especially when they go hunting for days in the forest. It also provides emergency food rations for rural villages (pers. obs.). Bangkong timber can also be used for construction and furniture; previous generations of the Temuan in Selangor used bangkong wood to make a mortar for pounding rice (pers. comm.). Bangkong is considered rare and underutilized in Malaysia. It is not known to most of the public, and only the Orang Asli utilize and know where to locate the wild fruit trees, which may be deep in the forest. Only three bangkong trees are kept as living

germplasm in the Rare Fruits Field Genebank at MARDI, Serdang, despite its importance as a genetic resource for cempedak breeding and as a rootstock (pers. obs.).

The absence of odor and taste in the bangkong fruit led Corner to suggest that bangkong was the wild progenitor of cempedak. However, this hypothesis was disputed by Primack (1985) based on his observations of cempedak in Sarawak, a Malaysian state on Borneo Island. Primack argued that the abundance of cempedak trees growing in the wild, including in undisturbed primary forests (e.g. Semengoh Forest Reserve Arboretum, and Mount Mulu National Park where the only people around were the Penan nomad tribe), indicate that cempedak is a native wild forest species in Borneo. He compared characteristics of leaf, twig, and male inflorescences between two populations of wild trees from secondary forest and two village populations, and found no consistent differences between wild and cultivated populations (Table 2 and Table 3). He also compared fruit characteristics of wild fruits from secondary forests, cultivated fruits from a village, a market in Sarawak, and a market in Peninsular Malaysia. Again, he did not observe obvious differences between wild and cultivated fruits, because all of the samples were highly variable in fruit quality. He concluded that the cultivated cempedak is not a highly derived variety, but virtually identical to the wild cempedak in the forests of Borneo. Based on his observations, he concluded that it was unlikely that cempedak was domesticated from bangkong in Peninsular Malaysia; rather, he noted that it was more likely that cempedak was brought to Peninsular Malaysia from Sumatra or Borneo via trade or migration of peoples, and bangkong was subsequently derived from cempedak in Peninsular Malaysia. However, it should be noted that Primack did not consider other morphological and fruit features such as hairiness, color, and fruit smell and taste that were used by Corner to distinguish cempedak and bangkong (Table 2 and Table 3). Comparison of the data from Primack and Corner suggests that wild and cultivated

cempedak from Sarawak may be distinct from Peninsular cempedak and bangkong measured by Corner; however, the data is limited and any interpretation is merely preliminary. The relationship between wild and cultivated forms of cempedak, in Peninsular Malaysia as well as Borneo, is unclear and warrants further investigation (Jarrett, 1959).

Table 1. Morphological characters distinguishing cempedak and bangkong, as described by Corner (1938).

Plant part	**Cempedak (*A. integer*)**	**Bangkong (*A. integer* var. *silvestris*)**
1. Hairiness of leaf, twigs, petioles, stipules, leaf veins, stalks of flower and fruit-heads	Long, wiry, spreading, pale to dark brown hairs, 1-7mm long	Varying from wholly glabrous to hairy
2. Color of leaf	**Underside:** dull middle-green, slightly yellowish **Midrib:** pale greenish yellow to greenish ochre **Withering**: orange-ochre to brown-ochre	**Underside:** usually ochre-brown **Midrib**: distinct ochre on both sides **Withering**: green, slightly yellowish green, dingy yellowish, then turning dry and brown, never rich ochre or orange
3. Male inflorescence	Peduncles 2-3mm wide Cylindrical head with minute furrow at base Length 3-5.3cm Width 0.9-1.2cm Stamen 1-1.3mm long	Peduncles 4-6mm wide Conical head with broad furrow at base Length 3-4.5cm Width 2-2.5cm Stamen 1.7-2.4mm long
4. Fruit (Ripe Syncarp)	Length 20-35cm Width 10-15cm Oblong cylindric Cream yellow to ochre or brownish ochre Strong harsh penetrating stench	Length 15-30cm Width 10-15.5cm Oblong elliptic Often uneven, golden then light orange-ochre to orange brown, or golden brown No smell
5. Rind	Spines slightly prominent, 2-4mm long	Spines distinctly prominent (3-5mm long) or not (2-3mm long)
6. Edible Pulp (Perianth)	Strong sweet taste of durian and mango	Slight sour-sweet taste and no smell
7. Seed (Matured Ovule)	Plump, slightly flattened	Like cempedak, but seed sometimes subcylindric
8. Embryo (Kernel)	Cotyledons less unequal, 2/3-3/4 length ratio	Cotyledons often nearly equal

Table 2. Morphological characteristics measured by Corner and Primack to compare bangkong and cempedak in peninsular Malaysia (Corner) or wild and cultivated cempedak in Sarawak (Primack). Only five measurements, highlighted in bold, were shared between the two studies.

Plant part	**Corner (1938)**	**Primack (1985)**
Leaf	Hairiness of leaf, petioles, leaf veins Color of leaf	Leaf length and width Number of lateral veins Petiole length and width
Twig	Hairiness of twigs, stipules, stalks of flower and fruit-heads	Bud length Twig thickness Node length
Male inflorescence	**Length and width** **Peduncle width** Stamen length Shape of head and size of furrow at base*	**Length and width** **Peduncle** length and **width**
Fruit	**Length** and width **Shape** Strong harsh penetrating stench Rind color	**Length** **Shape** Weight Core thickness
Rind	Length of spines	Rind thickness
Perianth	Presence or absence of strong sweet taste of durian and mango	Number of flowers Fruit ball length Thickness of fleshy perianth Weight of fleshy perianth % water content of fleshy perianth
Seed	Plump, slightly flattened	Seed length Weight of seed Number of seeds per fruit % water content of seeds % seed set

* In addition, Primack (1985) also examined herbarium material of bangkong and observed that all male inflorescences were cylindrical rather than conical as described by Corner, but it was unclear if he examined the same collections Corner used.

Table 3. Comparisons between measurements by Corner (1938) for cempedak and bangkong in Peninsular Malaysia, and by Primack (1985) for cultivated and wild cempedak in Sarawak.

	Corner (1938)		Primack (1985)				
	Cempedak	Bangkong	Cempedak wild (Semengoh)	Cempedak wild (Kampong Penang)	Cempedak cultivated (Kampong Segah)	Cempedak cultivated (Penrissen Road)	Cempedak Cultivated (market in Penin. M.)
Characteristics of Male Inflorescences							
Length (mm)	30 – 53	30 – 45	72.3 ± 17.8	56.6 ± 9.7	63.3 ± 14.4	55.3 ± 9.4	n/a
Width (mm)	9 – 12	20 – 25	13.7 ± 1.9	13.9 ± 2.6	14.5 ± 2.0	11.1 ± 1.8	n/a
Peduncle width (mm)	2 – 3	4 – 6	2.8 ± 0.5	2.9 ± 0.4	3.5 ± 0.7	2.7 ± 0.7	n/a
Sample size	Not specified	Not specified	15 (5 trees, 3 samples each)	15 (5 trees, 3 samples each)	15 (5 trees, 3 samples each)	15 (5 trees, 3 samples each)	n/a
Fruit Characteristics							
Fruit length (cm)	20 – 35	15 – 30	31.5 ± 5.76	26.8 ± 4.75	25.9 ± 4.73	22.3 ± 4.56	26.1 ± 2.63
Fruit shape	Oblong cylindric	Oblong elliptic	Varying from almost round to ovoid, to cylindrical to almost long cylindrical				
Sample size	Not specified	Not specified	16	11	14	9	11

1.5 Cempedak Research in Malaysia

In Malaysia, cempedak is most commonly grown in mixed-fruit orchards of varying sizes, sometimes in home gardens and villages, and less commonly in larger fruit plantations. Cempedak is usually propagated by bud-grafting to maintain clonal plants that are true-to-type and for faster fruiting (Dr. Abdul Rahman, MARDI, pers. comm.). Cempedak scions are normally grafted on cempedak rootstock, preferably from the same variety. The Department of Agriculture has registered at least 37 clonal varieties since 1951 and is growing these varieties in agricultural field stations around the country (Department of Agriculture of Peninsular Malaysia, 2001). In the early 1990s, the Malaysian Agricultural Research and Development Institution (MARDI) conducted a Peninsular-wide collection trip of landraces or farmer varieties of cempedak (Dr. Abdul Rahman, MARDI, pers. comm.), as part of a larger collaboration . Sixty-nine varieties were selected based on desirable fruit characteristics, and five clonal trees of each variety were planted in the largest cempedak germplasm collection in Kluang, Johor. Another major cempedak germplasm in Peninsular Malaysia is in Bukit Tangga, Kedah. This germplasm has at least 13 varieties and five or more clonal trees of each variety, focusing on cultivars from Kedah and neighboring states. Those accessions are mostly duplicated in the Kluang germplasm. The Bukit Tangga germplasm also serves as a commercial orchard open to wholesalers and locals, who can pick the fruits themselves and pay for it at a cheaper price than the market (MYR 1.00-1.50/kg, compared to the market price of MYR 3.50-4.50/kg in the year 2013). The Department of Agriculture of Sabah also has an *Artocarpus* genebank in Ulu Dusun with accessions of cempedak, breadfruit, jackfruit, and terap. Based on my communication with germplasm managers and collecting trips, the germplasm collections and other field stations are not always well-maintained, due to insufficient interest, funding, or human resources. For example, the MARDI Kemaman field station, another major germplasm collection of various

fruits including cempedak, was reportedly overgrown with weeds. In both Kluang and Bukit Tangga, some of the trees in the germplasm have died but have not been replanted; most trees have lost their labels or were not labeled at all (pers. obsv.). Without proper documentation, much of the knowledge of the provenance of the trees will be lost when the original staff in charge of the germplasm leaves the institution.

1.6 Objectives and Hypotheses

Despite the importance of cempedak as a food source, little is known about the distribution of genetic diversity in cempedak and bangkong across its range. Current plant genetic research in Malaysia emphasizes genetic enhancement through traditional breeding and introgression for specific traits, and is focused more on annual crops such as rice and vegetables, or priority fruit crops such as mango, rambutan, and durian for which funding is more readily available (Nordin et al., 2007; Bhag Mal et al., 2011). Characterization and evaluation of germplasm collections prioritize morphological characters and agronomic traits, and only a few collections (durian, cocoa, sweet potato, and a few medicinal herbs) have been characterized by molecular markers (Nordin et al., 2007). For cempedak, while 50% of cempedak germplasm accessions have been characterized for morphological traits, none have been characterized by molecular markers or for biochemical traits (Nordin et al., 2007). My research aims to use microsatellite markers to quantify the genetic diversity of cempedak and bangkong in Malaysia and beyond, as well as to elucidate the relationship between cempedak and bangkong. I will test several hypotheses as follows:

Genetic diversity:

H1a. Cultivated cempedak was domesticated from bangkong and will harbor a subset of the genetic diversity that bangkong has. (Hypothesis based on Corner 1938)

H1b. Peninsular cempedak was domesticated from Borneo cempedak and will harbor a subset of the genetic diversity that bangkong has. (Hypothesis based on Primack 1985)

Rationale:

A genetic bottleneck, or founder effect, often occurs during the process of domestication, when humans select only a subset of individuals in a wild species to be cultivated, which leads to cultivated populations having a narrower base of genetic variation compared to their wild progenitors (Doebley et al., 2006; Olsen and Gross, 2008) Over time, selective propagation of superior individuals can lead to low genetic diversity in elite cultivars (Yamasaki et al., 2005; Hyten et al., 2006). As such, I expect that comparison of genetic diversity measures to show that that cultivated cempedak will harbor a subset of the genetic diversity that is found in its hypothesized progenitor - bangkong (H1a, based on Corner (1938), or Borneo cempedak (H1b, based on Primack (1938)).

Genetic/Geographic structure:

H2. Bangkong exhibits greater geographic structure than cultivated cempedak.

Rationale:

Natural populations of a species can become isolated and differentiated over time due to geographic distance, geographic or reproductive barriers, or limited dispersal abilities that reduce or prevent gene flow between populations. In contrast, human-mediated dispersal and transfer of

cultivated plants across geographic boundaries, in the form of fruits, seeds, or graft material, result in weak geographic structure. Hence, I hypothesize that analyses of genetic structure and isolation-by-distance will show greater geographic structure in bangkong compared to cultivated populations of cempedak.

Phylogenetic relationship:

H3. Cultivated cempedak was domesticated from bangkong in peninsular Malaysia and phylogenetic analysis will indicate if there was a single or multiple domestication events.

Rationale:

The similarity of bangkong and cempedak morphology and overlap in range indicate that bangkong is closely related to cempedak, while the bigger size and strong smell of cempedak fruits suggest that it was domesticated from the bangkong in Peninsular Malaysia (Corner 1938). However, Primack (1985) suggested the opposite - cempedak originated outside Peninsular Malaysia, while bangkong was derived from the cempedak after it arrived in Peninsular Malaysia. Multiple geographical origins is not uncommon in domesticated perennial fruit tree crops (Miller and Gross, 2011), including breadfruit (*Artocarpus altilis*) (Zerega et al., 2004); so both Corner (1938) and Primack (1985) may be at least partially correct. I will use phylogenetic analyses, based on genetic distances, to help shed light on the origin of cempedak.

Cultivar genotypes:

H4. Named cempedak cultivars sharing the same name also share the same genotype (true-to-type).

Rationale: As most cempedak cultivars are propagated asexually (e.g. grafting), there should be no change in the genome from the scion to the grafted plant. I will use a clonal assignment test to determine if cempedak trees with the same cultivar names share the same genotypes. The clonal assignment test will also allow for identification of unknown cultivars and detection of mislabeled cultivars.

2. Methods

2.1 Site selection and sampling

Fieldwork was conducted in Peninsular Malaysia during July to October 2013. Potential study sites covering a broad geographical range in the peninsula were identified based on herbarium records, from speaking to scientists and citizens who are familiar with the species, and searching for trees growing along roadsides and in home gardens while driving between sites. Not all geographical areas within peninsular Malaysia were sampled due to accessibility and time constraints, or difficulty in obtaining permits (Figure 2). Sites included private homesteads, mixed-fruit orchards, commercial orchards, germplasm and cultivar collections, roadsides, forest reserves, and secondary forests (Table 4). To increase geographic sampling, additional samples were obtained from herbarium vouchers; collections made by FRIM in Johor and Pahang in 2013; collections made by N. Zerega and E. Gardner in Sabah, Malaysia in summer 2013; collections made by N. Zerega in Malaysia (2002), Thailand (2012) and Singapore (2002); and collections made by E. Gardner from Miami, FL, USA (Table 5).

Trees were located with the help of field assistants, local guides or tree owners. In cultivated sites, individual trees were chosen to represent a range of cultivars based on suggestions of guides or landowners. Leaf tissue was collected from each sample tree and stored in silica for molecular analysis in the lab. Whenever possible, DBH (diameter at breast height), GPS coordinates, and digital photos of the whole tree, fruits and inflorescences were recorded for each tree. To obtain morphological data, two specimens of branches with twigs and leaves and two withering leaves were collected from each tree, labeled, and stored in an extra-large Ziploc bag at room temperature (Figure 3 and Figure 4). Branches were chosen based on what was accessible. Vegetative characteristics were visually assessed and recorded *ex situ* on the same day or within 1-2 days (Table 6). Whenever available, one fruit was collected per tree and

measured (length, circumference, shape, rind color, pulp color, fruit odor, pulp taste, and seed size; N = 12 cempedak, 3 bangkong). Owners and local guides were also interviewed to obtain information such as tree age, seasonality (fruiting frequency and fruiting months), fruit quality and yield, fruit color, vernacular name or cultivar, care given, pests, propagation method (e.g. seed, sapling, grafted), source or origin of tree, selling price, and profit (Appendix 1). Representative voucher specimens were collected for most cultivars and most sites and deposited at the herbaria at FRIM (KEP) and the Nancy Poole Rich Herbarium at the Chicago Botanic Garden (CHIC).

In Peninsular Malaysia, trees were determined to be bangkong based on reports of field assistants or local guides with knowledge of the area, and also if the trees were growing wild in or near forest reserves where there were herbarium records of bangkong. Otherwise, they were labeled as cempedak. All wild trees sampled in Peninsular Malaysia were labeled as bangkong. Collections made by FRIM (collection numbers beginning with FRI) were initially identified as wild cempedak, but relabeled as bangkong early on based on results from preliminary analyses and sequencing (Gardner, unpublished). In Sabah, fruits were often unavailable on specimens to determine if the trees were cempedak or bangkong. Additionally, there were no local guides who were familiar with individual trees to determine the ID to subspecies level. As bangkong has not been reported from Sabah, we considered all wild trees from Sabah to be "wild cempedak" (N. Zerega, pers. comm.). On the same note, we did not have any bangkong samples from outside Peninsular Malaysia. Lastly, in this study, we will refer to cempedak trees whose cultivation status we could not determine as "unknown cempedak".

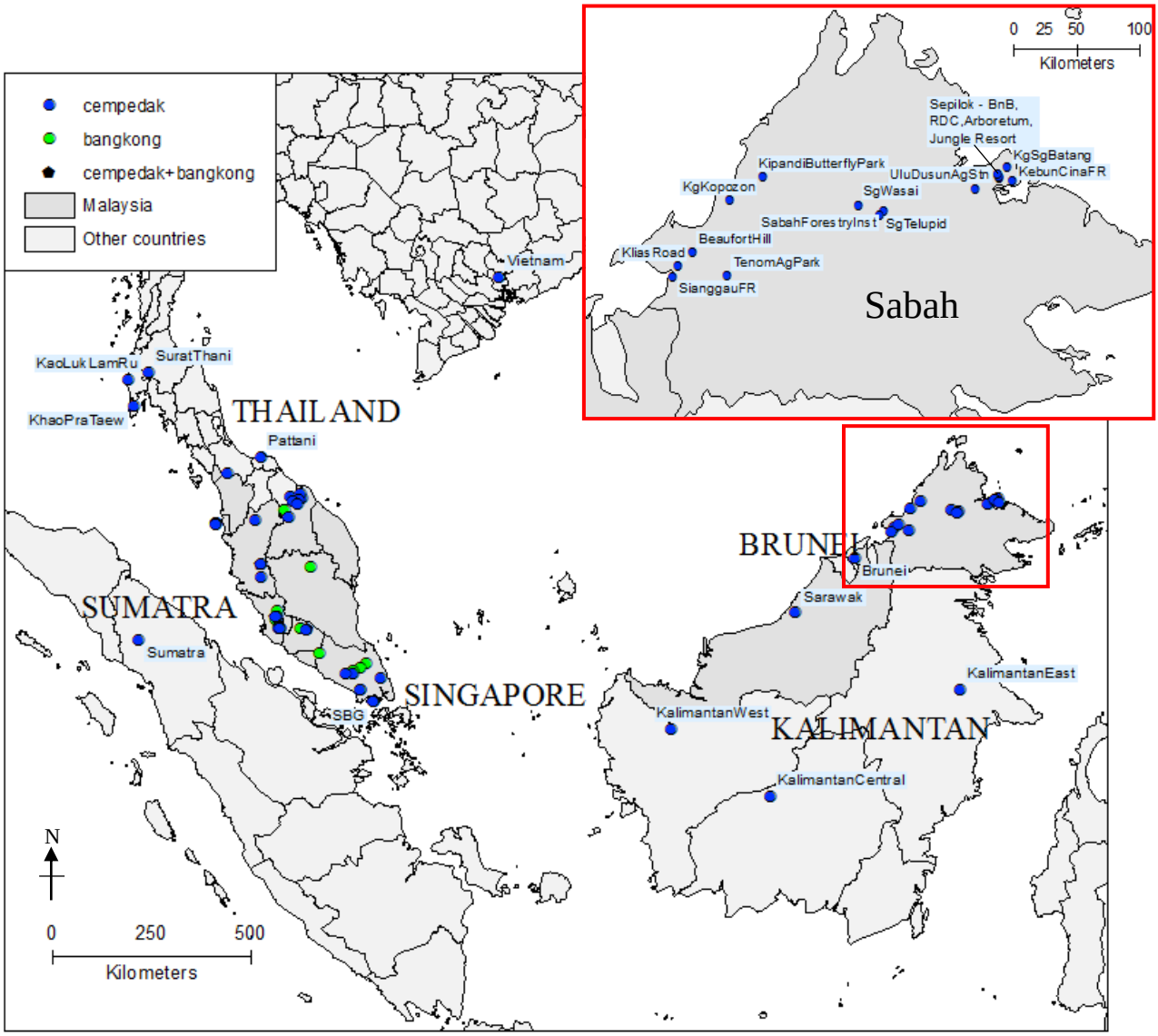


Figure 2a. Distribution of sites where cempedak (blue circles) or bangkong (green circles) were sampled in Malaysia and neighboring countries. Map inset shows sites in Sabah. Samples from Xishuangbanna Tropical Botanical Garden in China and Miami in Florida, USA are not shown on this map. See Figure 2b for details of sites in Peninsular Malaysia.

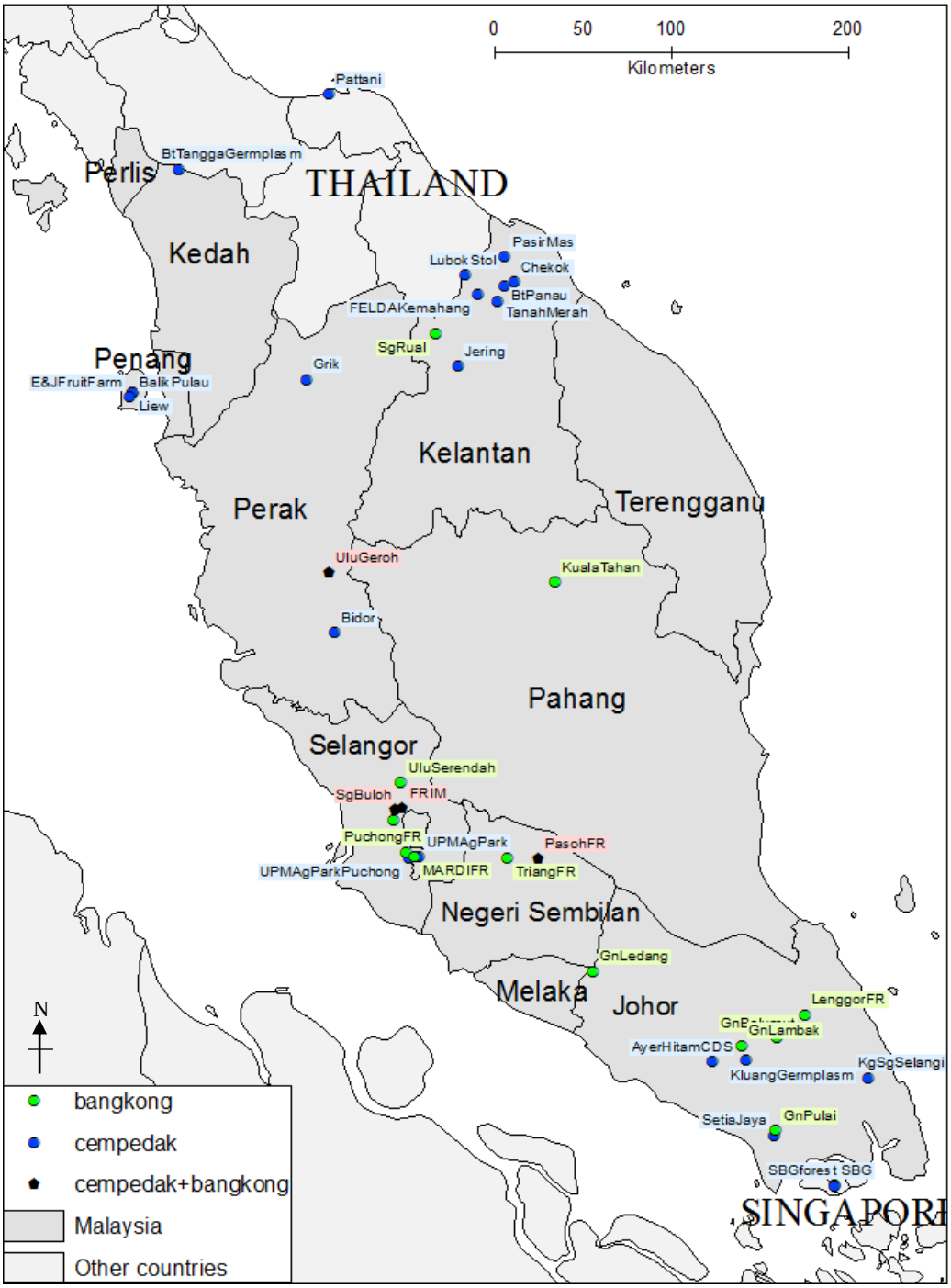


Figure 2b. Distribution of cempedak sites (blue circles and blue labels), bangkong sites (green circles and green labels), and mixed sites (black pentagon and pink labels) in Peninsular Malaysia.

Table 4. Location data for cempedak (N = 269 including 7 "nangka-cempedak" (jackfruit-cempedak) hybrids) and bangkong (N = 185) trees sampled from Peninsular Malaysia.
Representative herbarium vouchers were made for most sites and deposited at the Forest Research Institute of Malaysia Herbarium (KEP) in Selangor, Malaysia. Duplicate vouchers were deposited at the Nancy Poole Rich Herbarium at the Chicago Botanic Garden (CHIC). Numbers in site codes indicate the state/country where the site is located, and letters indicate different regions within the state. N = Number of samples analyzed from each site. Further details for each collection can be found in Appendix 2.
Site abbreviations: FR = forest reserve, DoA = Department of Agriculture, FELDA = Federal Land Development Authority, FRIM = Forest Research Institute of Malaysia, MARDI = Malaysian Agricultural Research and Development Institute.

Region of Peninsular Malaysia	State (District)	Site Name	Site Code	Lat. (N)	Long. (E)	Land-use description	N of cempedak	N of bangkong	Voucher no(s).
Northeast	Kelantan (Jeli)	Kampung Bukit Jering (MW429-431)	cult5d-Jering	5.483	101.904	Homestead	3	0	-
	Kelantan (Jeli)	Sungai Rual (MW1501-1506)	silv5d-SgRual	5.651	101.791	Jahai Orang Asli village. Forest edge/homestead	0	6	MW1501- 1506
	Kelantan (Pasir Mas)	Chekok (MW425-428)	cult5a-Chekok	5.912	102.190	Small mixed fruit orchard	4 hybrids	0	-
	Kelantan (Pasir Mas)	Pasir Mas (MW401-403)	cult5a-PasirMas	6.040	102.144	Homestead	2 + 1 hybrid	0	-
	Kelantan (Rantau Panjang)	Lubok Stol (MW441-443)	cult5a-LubokStol	5.946	101.940	Homestead; small communal mixed fruit orchard; roadside garden	3	0	MW441
	Kelantan (Tanah Merah)	Bukit Panau (MW410-424)	cult5b-BtPanau	5.895	102.141	Small mixed fruit orchard, not actively cared for	23 + 1 hybrid	0	-
	Kelantan (Tanah Merah)	FELDA Kemahang (MW603-640)	cult5c-FELDAKemahang	5.852	102.007	Commercial mixed fruit orchard	41	0	MW617
	Kelantan (Tanah Merah)	Tanah Merah (MW601-602)	cult5c-TanahMerah	5.811	102.109	Homestead, roadside/edge of oil palm estate	3	0	-

Region of Peninsular Malaysia	State (District)	Site Name	Site Code	Lat. (N)	Long. (E)	Land-use description	N of cempedak	N of bangkong	Voucher no(s).
Northwest	Kedah (Bukit Kayu Hitam)	MARDI Germplasm at Bukit Tangga (MW801-846)	cult6a-BtTanggaGermplasm	6.486	100.484	Germplasm collection	46	0	MW803, 808, 810, 814, 817, 820, 823, 827, 828, 831, 836, 837, 841, 843
	Penang (Balik Pulau)	Balik Pulau (MW701-730)	cult7a-BalikPulau	5.350	100.248	Mixed fruit orchard on hill (several owners)	12	0	-
	Penang (Balik Pulau)	Liew family orchard (MW711,713,715-730)	cult7b-Liew	5.350	100.250	Mixed fruit orchard on hill	18	0	MW724
	Penang (Balik Pulau)	E & J Fruit Farm (MW731-745)	cult7c-E&JFruitFarm	5.326	100.238	Mixed fruit orchard on hill	15	0	MW733, 742
Central-west	Perak (Grik)	Grik (MW851-855)	cult8a-Grik	5.417	101.133	Homestead	5	0	-
	Perak (Kampar)	Ulu Geroh (MW1101-1130, MW1151-1153)	cult8b-UluGeroh; silv8b-UluGeroh	4.433	101.250	Semai Orang Asli Village. Forest (bangkong), homestead (cempedak)	3	30	MW1151 (cempedak), MW1109, 1121, 1123 (bangkong)
	Perak (Batang Padang)	Bidor (MW501-514)	cult8c-Bidor	4.128	101.275	Small mixed fruit orchard; market	12	0	-
Southwest	Selangor (Hulu Selangor)	Ulu Serendah (MW1301-1312)	silv10a-UluSerendah	3.365	101.616	Forest edge near village	0	12	MW1301-1312
	Selangor (Kepong)	FRIM Kepong (MW101-119)	silv10b-FRIM-Kepong	3.239	101.624	Secondary forest	1	18	MW101,107, 113
	Selangor (Petaling)	Sungai Buloh (MW301-330, MW351-370)	silv10c-SgBuloh	3.231	101.586	Temuan Orang Asli Village. Secondary forest; cultivated trees around/in village	20	30	MW353, 364, 370 (cempedak), MW322, 329, 330 (bangkong)

Region of Peninsular Malaysia	State (District)	Site Name	Site Code	Lat. (N)	Long. (E)	Land-use description	N of cempedak	N of bangkong	Voucher no(s).
	Selangor (Petaling)	Kota Damansara Community Forest, Petaling Jaya (MW1313-1318)	silv10d-KotaDamansara	3.169	101.581	Forest park	0	6	MW1313-1318
	Selangor (Petaling)	Ayer Hitam FR, Puchong (MW1601-1633)	silv10e-PuchongFR	3.011	101.646	Forest reserve	0	33	MW1601-1633
	Selangor (Petaling)	MARDI FR, Serdang (MW201-219)	silv10f-MARDIFR	2.987	101.686	Forest park, rare fruits germplasm	0	19	MW205-219
	Selangor (Petaling)	Universiti Putra Malaysia (UPM) Agro-Park, Serdang (MW1401-1406)	cult10f-UPM	2.987	101.708	Cultivar collection for teaching and commercial purposes	5 + 1 hybrid	0	MW1401-1407
	Selangor (Petaling)	UPM Agro-Park, Puchong (MW1407)	cult10f-UPM-Puchong	2.983	101.658	Cultivar collection for teaching and commercial purposes	1	0	MW1407
Southwest	Negeri Sembilan (Jelebu)	Pasoh (FRIM field station) (MW1001-1015)	silv11a-FRIMPasoh	2.982	102.313	Palm oil estate (cempedak); forest research plot (bangkong)	1	15	MW1001 (cempedak), MW1002-1011, 1015 (bangkong)
South	Johor (Batu Pahat)	Ayer Hitam Commodity Development Station (DoA) (MW971-976)	cult12e-AyerHitamCDS	1.945	103.197	Registered cultivar (clone) collection	6	0	MW971-973
	Johor (Kluang)	Lenggor FR (MW1205-1206)	unk12b-Lenggor FR	2.183	103.667	Entrance to forest reserve/park area	0	2	MW1205-1206
	Johor (Kluang)	Gunung Belumut (MW1202-1204)	unk12c-Gn Belumut	2.065	103.527	Forest reserve; educational forest; park area	0	3	MW1202-1204
	Johor (Kluang)	Gunung Lambak (MW1201)	unk12c-Gn Lambak	2.025	103.345	Beside road to campground in forest	0	1	MW1201

Region of Peninsular Malaysia	State (District)	Site Name	Site Code	Lat. (N)	Long. (E)	Land-use description	N of cempedak	N of bangkong	Voucher no(s).
	Johor (Kluang)	MARDI Germplasm at Kluang (MW901-965; numbers not consecutive)	cult12d-KluangGermplasm	1.949	103.368	Germplasm collection	57 + 1 hybrid	0	MW901, 916, 924, 945
	Johor (Kota Tinggi)	Kampung Sungai Selangi (MW1207)	cult12f-KgSgSelangi	1.860	103.993	Kanaq Orang Asli Village. homestead	1	0	MW1207
	Johor (Kulaijaya)	Gunung Pulai (MW1208-1218);	silv12g-GnPulai	1.569	103.511	Disturbed forest	0	11	MW1208-1218
	Johor (Kulaijaya)	Kampung Setia Jaya (MW1219-1220)	cult12g-KgSetiaJaya;	1.592	103.522	Homestead/escaped	2	0	MW1219-1220

Table 5. Additional samples from other collectors and herbarium vouchers.
Collection numbers indicate collector name or sampling location, followed by extraction label in parentheses if not the same as collection number. Numbers are not always consecutive. Samples were determined to be putative jackfruit-cempedak hybrids based on morphological features. Abbreviations: NZ = N. Zerega, EG = E. Gardner, SBG = Singapore Botanic Garden, FR = forest reserve. Herbarium abbreviations follow the Index Herbariorum codes.

Coll. No.	Taxon	Location of collection	Site code	Country of collection	Voucher location
FRI 78710-78713	Bangkong	Kuala Tahan, primary forest, beside canopy walkway, Jerantut, Pahang	9-Kuala Tahan	Malaysia	KEP, CHIC
FRI 79724	Bangkong	Gunung Ledang FR, primary forest, trail, Muar, Johor	12a-Gn Ledang	Malaysia	KEP, CHIC
NZ220	Cempedak	Singapore Botanic Garden grounds	13-SBG	Singapore	CHIC
NZ222	Cempedak	Singapore Botanic Garden grounds	13-SBG	Singapore	CHIC
NZ227	Bangkong	Pasoh FR, in 50-ha research plot, Negeri Sembilan.	11a-Pasoh FR	Malaysia	CHIC
NZ236	Bangkong	Triang FR, secondary forest, Negeri Sembilan	11b-Triang FR	Malaysia	NY, CHIC
NZ260	Cempedak	Forest Research Centre, Sepilok Arboretum, Sandakan, Sabah	14d-Sepilok Arboretum	Malaysia	NY, CHIC
NZ466-467	Cempedak	Kao Luk Lam Ru National Park, in a patch of forest between the road and beach	4-Kao Luk Lam Ru	Thailand	CHIC

Coll. No.	Taxon	Location of collection	Site code	Country of collection	Voucher location
NZ476	Cempedak	Khao Pra Taew Wildlife Conservation Development and Extension Center. Growing near restaurant in park HQ	4-Khao Pra Taew	Thailand	CHIC
NZ485b	Cempedak	Peat-swamp forest, Pattani	4-Pattani	Thailand	CHIC
NZ509-511	Cempedak	Private yard, Surat Thani	4-Surat Thani	Thailand	CHIC
NZ616	Cempedak	Sepilok Rainforest Discovery Center, forest, Sandakan, Sabah	14b-Sepilok RDC	Malaysia	SAN, CHIC
NZ632-658	Cempedak; Nangka-cempedak hybrid cultivar	Ulu Dusun Agriculture Station, Sandakan, Sabah	14g-Ulu Dusun Ag Stn	Malaysia	SAN, CHIC
NZ689-691	Cempedak	Kebun Cina Forest Reserve, forest encompassing abandoned rubber plantation, Sandakan, Sabah	14f-Kebun Cina FR	Malaysia	SAN, CHIC
NZ703	Cempedak	Sepilok Bed and Breakfast, Sandakan, Sabah	14c-Sepilok BnB	Malaysia	SAN, CHIC
NZ710-711	Cempedak	Home garden in Kampong Sungai Batang, Sandakan	14a-Kg Sg Batang	Malaysia	SAN, CHIC
NZ712	Putative cempedak-jackfruit hybrid	Sepilok Jungle Resort, cultivated area, Sandakan, Sabah	14e-Sepilok Jungle Resort	Malaysia	SAN, CHIC
NZ725	Cempedak	Sabah Forestry Institute, cultivated area, Beluran, Sabah	14k-Sabah Forestry Inst	Malaysia	SAN, CHIC
NZ732	Cempedak	Along roadside near forested area, Sungai Wasai (Wasai River), Beluran, Sabah	14i-Sg Wasai	Malaysia	SAN, CHIC
NZ738	Cempedak	Along roadside near forested area, Sungai Telupid (Telupid River), Beluran, Sabah	14j-Sg Telupid	Malaysia	SAN, CHIC
NZ785-787, NZ792-793	Cempedak	Kampong Kopozon, village with cultivated and wild growing trees, Papar, Sabah	14h-Kg Kopozon	Malaysia	SAN, CHIC
NZ812	Putative cempedak-jackfruit hybrid	Kipandi Butterfly Park, along a managed forest trail, Kota Kinabalu, Sabah	14k-Kipandi Butterfly Park	Malaysia	SAN, CHIC
NZ836, 843, 845, 846	Cempedak	Beaufort Hill, disturbed forest, Beaufort, Sabah	14l-Beaufort Hill	Malaysia	SAN, CHIC
NZ850	Cempedak	Overgrown yard on road between Klias FR and Sianggau FR, Beaufort, Sabah	14l-Klias Road	Malaysia	SAN, CHIC
NZ854, 858-861, 863, 865, 869, 918	Cempedak; putative cempedak-jackfruit hybrid	Sianggau FR, disturbed forest, Beaufort, Sabah	14n-Sianggau FR	Malaysia	SAN, CHIC
NZ918	Cempedak	Tenom Agriculture Park, cultivated botanical garden, Tenom, Sabah	14m-Tenom Ag Park	Malaysia	SAN, CHIC

Coll. No.	Taxon	Location of collection	Site code	Country of collection	Voucher location
NZ941-943	Cempedak	Tenom Agricultural Park, forest edge near Artocarpus orchard, Tenom, Sabah	14m-Tenom Ag Park	Malaysia	SAN, CHIC
EG21	Putative cempedak-jackfruit hybrid	The Kampong, Hissar estate, Miami, Florida	1-Miami	USA	CHIC
EG55	Cempedak	Fairchild Tropical Botanic Garden, Rare Fruits Pavilion, Miami, Florida. Said to be from Malaysia	1-Miami	USA	CHIC
SBG3	Cempedak	Leaf material sent by mail from Singapore Botanic Garden grounds, unknown accession	13-SBG	Singapore	No voucher
Y. Q. Lee 36 (YQL36)	Cempedak	Dried herbarium leaf: Singapore Botanic Garden Rainforest, remnant primary forest	13-SBG forest	Singapore	SING, CHIC
CAS9	Cempedak	Dried herbarium leaf: Collected on 20 January 2013. No other information given	2-China	Singapore	HITBC
Gentry & La Frankie 66897 (GL66897)	Cempedak	Dried herbarium leaf: Pasoh FR, in 50-ha plot, Negeri Sembilan	11a-Pasoh FR	Malaysia	MO
Burley & Lee 246 (F.mus.int.1)	Cempedak	Dried herbarium leaf: Similajau FR, Sarawak	16-Sarawak	Malaysia	F
Berau 865 (Leiden865)	Cempedak	Dried herbarium leaf: Logged-over forest, Berau Inhutani area, East Kalimantan	17-Kalimantan East	Indonesia	L
J. Dransfield 6877 (Leiden6877)	Cempedak	Dried herbarium leaf: Ladan Hills FR, disturbed forest, Tutong District	15-Brunei	Brunei	L
Mogea 3665 (Kew3665)	Cempedak	Dried herbarium leaf: Tumbang Riang, primary forest, Bukit Raya and upper Katingan (Mendawai) River area, Tumbang Samba, Central Kalimantan	18-Kalimantan Central	Indonesia	K
Rahmat Si Boeea 5753 (NYBG1)	Cempedak	Dried herbarium leaf: Pargambiran, East Sumatra	19-Kalimantan West	Indonesia	NY
P. K. Loc 642 (NYBG2)	Cempedak	Dried herbarium leaf: Lai Thieu, Binh Duong	3-Vietnam	Vietnam	NY
Jong 383 (NYBG3)	Cempedak	Dried herbarium leaf: Trailside rubber garden, Ngira, West Kalimantan	20-Sumatra	Indonesia	NY

Table 6. Vegetative characteristics used to compare cempedak and bangkong based on Corner (1938). Sample size of measurements recorded varied depending on plant part measured. See Appendix 4 and 5 for more details.

Characteristic	Part of plant	Measurement recorded	N
Color	Leaf underside, leaf midrib, withering leaf (upper and underside)	Color based on mini RHS color chart (Flower Council of Holland, UK)	276 (cempedak) 185 (bangkong)
Hairiness	Leaf edges, leaf underside, leaf veins, midrib (upper side), petioles, nodes, twigs, stipules	Recorded as glabrous, hairy, or very hairy	60 - 277 (cempedak) 146 - 185 (bangkong)

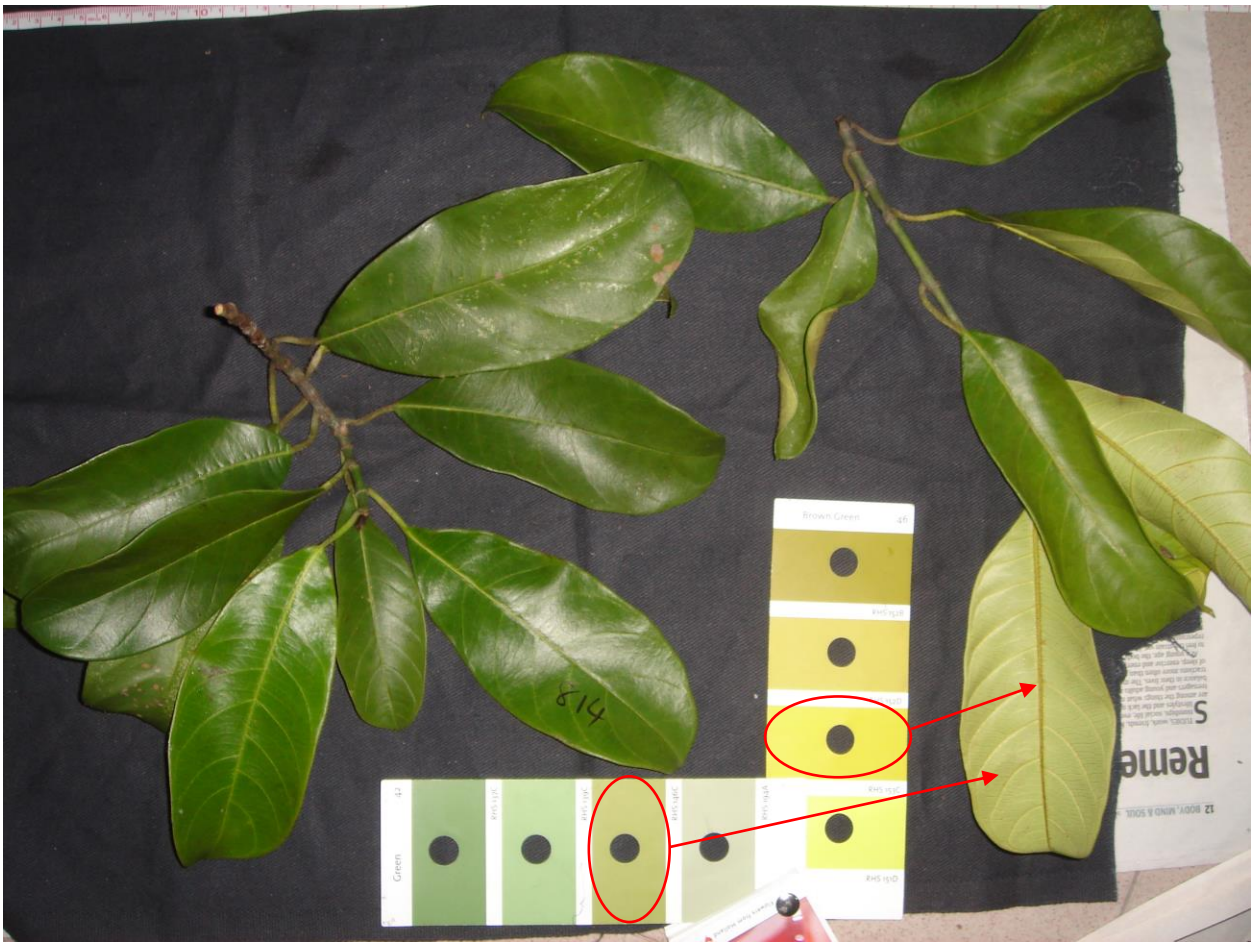


Figure 3. Example of two branches with leaves collected from an individual tree (MW814), with color chart. This image indicates how the leaf underside and midrib color were scored for this sample as 146c and 153c respectively. Hairiness of leaves, petioles, twigs, and stipules were also recorded from these branches.

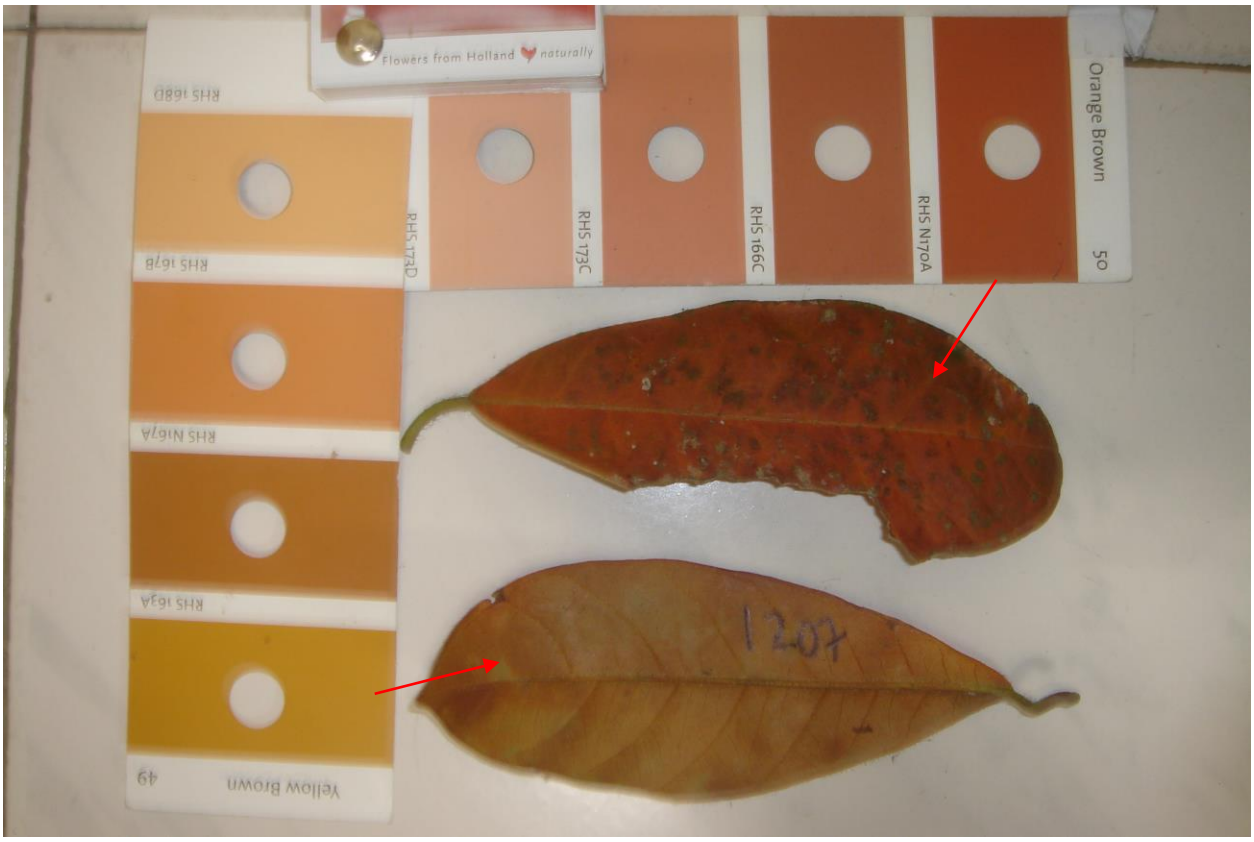


Figure 4. Sample of two withering leaves. The color of the top leaf (adaxial or upper side) was recorded as N170a, and the bottom leaf (abaxial or underside) was recorded as 163A.

2.2 Microsatellite amplification and genotyping

Leaf material collected in peninsular Malaysia and Sabah was dried on silica gel and shipped to the Chicago Botanic Garden where it was stored at -80°C. Total genomic DNA was extracted from $1cm^2$ dried leaf using either the Qiagen DNeasy kits (Qiagen, Valencia, CA) following standard protocol or a modified cetyltrimethyl ammonium bromide (CTAB) method (Doyle and Doyle, 1987). PCR was conducted using 10 primers (amplifying 12 loci) developed for *A. altilis* (Witherup et al., 2013). These primers have been shown to cross-amplify and exhibit variation in jackfruit (*A. heterophyllus)* and cempedak (*A. integer)* (Witherup et al., 2013; Zerega et al., in press). Forward primers were pre-labeled with WellRed D2 (black), D3 (green), or D4 (blue) fluorescent dye (Sigma-Proligo, St. Louis, Missouri, USA) (Table 7).

Initially, samples were amplified in multiplexed PCR reactions. PCR reactions contained varying amounts of 10μM forward and reverse primers based on color (detailed in Table 8), 5 μL 2x MyTaq Mix (Bioline, Taunton, Massachusetts, USA), 0.15 μl of 20 mg/ml BSA, 1 μL genomic DNA, and enough DNA-free H_20 to make up a 10 μL reaction. Initial denaturation was set at 94ºC for 2 min, followed by 34 cycles of 94ºC for 30 s, 52ºC for 1 min, 72ºC for 2 min, ending with a final extension at 72°C for 10 min.

PCR products were analyzed using a CEQ 8000 Genetic Analysis System (Beckman Coulter, Brea, California, USA). Before loading, 14.4 μl of GenomeLab 400 bp internal size standard ladder (Beckman Coulter, Brea, CA, USA) was added to 1.5 ml of HiDi formamide (Azco Biotech., San Diego, CA, USA), and 30 μl of this mixture was added to each well containing 1.0 – 1.5 μl of multiplexed PCR product. The large difference in fluorescence intensity between the black (low intensity) and blue (high intensity) dyes caused difficulties in scoring alleles. Even after adjusting the ratios of blue and black primers, multiplexed PCR product still produced noisy results, which were difficult to interpret. Subsequently, samples

were amplified in single-primer PCR reactions, but PCR products were multiplexed in the CEQ 8000. PCR reactions and analysis of PCR products were conducted using the same methods described above, but the amount of primers in the PCR and the amount of PCR product loaded on the CEQ 8000 varied based on dye color (see Table 8 for details). Samples that were difficult to score or displayed ambiguous results were rerun using the single-primer PCR method, multiplexed in the CEQ 8000, and rescored.

Allele sizes were manually scored using the CEQ 8000 software 9.0. Allele sizes were rounded to integers to fit the repeat pattern of each microsatellite locus (all primers except 145 were dinucleotide repeats). Cempedak (*Artocarpus integer*) is considered functionally diploid, but primers 196 and 156 showed more than two alleles in some cases. When more than two strong peaks were observed, I omitted the rarest allele (allele size 271 for primer 196) or alleles that showed the weakest signal (allele size 278 for primer 156).

Each locus was tested for potential null alleles and deviations from Hardy Weinberg equilibrium using Micro-Checker 2.2.3 (Van Oosterhout et al., 2004) for each population. Only results for eight wild populations with sample sizes of at least 10 were considered because cultivated populations were not expected to follow Hardy Weinberg assumptions. Primers 140, 105, and 145 showed evidence of null alleles for five out of eight wild populations. To determine if those primers should be included in further analyses, null allele frequencies were calculated for each locus for each wild population and the total wild population using FreeNA (Chapuis and Estoup., 2007). A null allele frequency value of > 0.20 can cause significant underestimation of population differentiation, and this threshold value has been used in other studies to exclude primers that may significantly affect heterozygosity (Chapuis and Estoup., 2007; Minn et al., 2014; Muzzalupo et al., 2014).

Table 7. Microsatellite primers used in this study. Detailed information for each primer can be found in Witherup et al. (2013).

Multiplex combination for PCR and CEQ	**Primer**	**Color and WellRed #**	**Allele size range in cempedak**	**Notes**
#1	MAA54	Blue (D4)	160-172 194-206	Primer amplified at two loci.
	MAA178	Green (D3)	211-261 249-291	Primer amplified at two loci.
	MAA196	Black (D2)	250-278	Primer amplified at two loci in *A. atilis* and *A. heterophyllus*, but did not reliably amplify the second locus in cempedak. This primer sometimes showed more than two alleles per individual in which scoring was done as described in the methods.
#2	MAA140	Blue (D4)	121-161	
	MAA156	Green (D3)	263-285	This primer sometimes showed more than two alleles per individual.
	MAA105	Black (D2)	245-275	
#3 This combination works only for multiplexing in CEQ 8000 (see notes)	MAA182	Blue (D4)	158-178	Did not amplify when multiplexed in PCR with MAA26 and MAA122, but amplified when duplexed in PCR with either MAA122 or MAA145.
	MAA26	Black (D2)	260-282	Did not amplify when multiplexed in PCR with MAA182 and MAA122, but amplified when duplexed in PCR with either MAA122 or MAA145.
	MAA122	Green (D3)	237-287	
#4	MAA145	Green (D3)	257-295	Primer was usually run by itself.

Table 8. Amounts of 10 µM primer used in multiplexed and single-primer PCR reactions and amounts of PCR product loaded in the CEQ 8000 Genetic Analysis System. The ratio of primers was adjusted based on dye color because the fluorescence intensity of the blue dye was much higher than the intensity of the black dye.

	Amount of 10 µM primer (forward/reverse) added to each 10-µL PCR reaction	**Amount of PCR product added to each well in the CEQ**
Multiplexed PCR		
Blue + Green	0.10 µL blue + 0.25 µL green	1.0-1.50 µL
Green + Black	0.125 µL green + 0.25 µL black	1.0-1.50 µL
Blue + Green + Black	0.10 µL blue + 0.25 µL green + 0.50 µL black	1.0-1.50 µL
Single-primer PCR		
Blue	0.10 - 0.125 µL blue *dilute to 1/2 concentration before loading in CEQ	Single-primer runs: 1.0-1.50 µL Multiple-primer ('multiplex') runs: 0.5 µL diluted blue + 0.5 µL green + 1.0 µL black
	0.25 µL blue *dilute to 1/3 concentration before loading in CEQ	
Green	0.25 µL green	
Black	0.25 µL black	

2.3 Genetic diversity

Diversity statistics including average number of alleles per locus (Na), number of effective alleles (Ne), observed and expected heterozygosity (Ho and He), fixation index (F) and number of private alleles (Pa) were calculated using GenAlEx v.6.5 for each locus and population (Peakall and Smouse, 2012).Allelic richness, a statistic which corrects for bias in sampling sizes, was calculated using FSTAT ver. 2.9.3.2 (Goudet, 2001). To test for statistical significance in comparisons of diversity statistics between cempedak and bangkong or between Peninsular and Borneo cempedak, Student's t-tests were performed in R v.3.0.2 statistical software (R Development Core Team, 2014).

2.3 Geographic patterns in genetic structure and differentiation

Genetic structure was assessed in STRUCTURE v.2.3.4 (Pritchard et al., 2000). Twenty independent runs per k were carried out with a burn-in period of 10,000 and 10,000 MCMC iterations for k = 1-20. The most likely value of genetic clusters (K) was identified using

Structure Harvester implementing the Evanno method (Evanno et al., 2005; Earl and VonHoldt, 2012). Principle coordinate analyses (PCoA) were conducted using the "vegan" package in R v.3.0.2 (R Development Core Team, 2014). Calculations of genetic distances were conducted in GenAlEx v.6.5 (Peakall and Smouse, 2012) following the method of Smouse and Peakall (1999).

AMOVA and Mantel tests of matrix correspondence between pairwise-population Fst or Nei's genetic distance (Da) and geographic distance were conducted in GenAlEx v.6.5 (Peakall and Smouse, 2012). Geographic distances between sites (in km) were calculated from GPS coordinates (mean center for sites) based on a modified Harvesine formula developed by Sinnott (1984), as implemented in GenAlEx v.6.5 (Peakall and Smouse, 2012).

2.4 Phylogenetic analyses

Nei's genetic distance, Da (Nei et al., 1983) and Fst distances with sample size bias correction (Latter 1972) were computed for cempedak and bangkong populations ($N \geq 4$ only) with 1000 bootstrap replicates, and the resulting matrix used to construct neighbor-joining and UPGMA trees in POPTREE2 (Takezaki et al., 2010). The resulting dendograms were edited in FigTree v. 1.4.2 (Rambaut, 2014). Individual-pairwise genetic distances across 12 loci were computed in GenAlEx v.6.5 (Peakall and Smouse, 2012) following the method of Smouse and Peakall (1999). Phylogenetic trees were generated from the resulting distance matrix using the neighbor-joining and UPGMA algorithms in PHYLIP (Felsenstein, 2005). The dendograms were visualized and edited in FigTree v. 1.4.2 (Rambaut, 2014). Jackfruit samples from Midanahall, India (N=22) were used as the outgroup for all phylogenetic analyses (courtesy of T. Melhem).

2.5 Analysis of cempedak cultivars

Clonal assignment was conducted in GENODIVE (Meirmans and Van Tienderen, 2004) using the step-wise mutation model. Sixty-nine samples with missing data were removed prior to analysis, resulting in a reduced sample size of 476. Thresholds (referring to the maximum distance allowed between two individuals to be considered clonemates sharing the same genotype) ranging from 0 to10 were tested. Higher thresholds allow more "wiggle-room" for individuals with scoring errors and mutations to be considered clones, but may also underestimate clonal diversity. A threshold of 4-5 was selected, because it maximized the number of individuals with cultivar names corresponding to clonal groups without introducing too many errors (i.e., unexpected clonal relationships or lumping of clonal groups with different cultivar names). Structure analysis (Pritchard et al., 2000) was also carried out among individuals with cultivar names, with a burn-in period of 10,000 and 10,000 MCMC iterations for k = 1-30, and 20 independent runs per k.

3. Results

3.1 Genetic diversity

3.1.1 Diversity across loci

The 12 microsatellite loci were polymorphic across all 545 individuals analyzed and detected a total of 150 alleles. MAA178b was the most polymorphic (Na = 19), while MAA54a was the least polymorphic (Na = 5), with an average of 12.5 ± 1.433. Observed heterozygosity (Ho) ranged from 0.024 (MAA54b) to 0.628 (MAA122), while expected heterozygosity ranged from 0.028 (MAA54b) to 0.805 (MAA122). All loci showed homozygote excess ranging from F = 0.048 (MAA156) to F = 0.495 (MAA105), indicating inbreeding or undetected null alleles.

Based on exact tests in Microchecker, loci 140, 145, and 105 showed evidence of null alleles in five out of 10 wild populations. Null allele frequencies of all three loci for the total wild population were higher than other loci, but were lower than the threshold of 0.20 above which population differentiation is significantly underestimated (Table 9; Chapuis and Estoup 2007; Muzzalupo et al. 2014; Minn et al. 2014). Null allele frequencies of higher than 0.20 were found in only one out of 10 wild populations for loci 140 and 145, and none for locus 105 (not shown). Because these loci did not consistently show high frequency of null alleles across populations, they were included in subsequent analyses.

Table 9. Measures of genetic diversity for each locus across all samples.

Locus	N	Na	Ne	Ho	He	F	Null allele frequency
54a	532	5	1.55	0.336	0.354	0.050	0.004
54b	531	6	1.03	0.024	0.028	0.126	0.001
178a	533	16	3.61	0.508	0.723	0.297	0.021
178b	534	19	4.25	0.517	0.764	0.324	0.049
196	532	9	2.62	0.479	0.618	0.224	0.050
140	537	18	3.92	0.609	0.745	0.182	0.127
156	542	8	2.73	0.603	0.634	0.048	0.019
105	533	11	2.07	0.261	0.516	0.495	0.103
182	525	11	1.56	0.248	0.360	0.312	0.074
26	523	12	1.71	0.304	0.416	0.269	0.057
122	524	18	5.14	0.628	0.805	0.220	0.069
145	538	17	1.41	0.178	0.293	0.390	0.132
Mean	532	12.5	2.633	0.391	0.521	0.245	0.059
SE	1.7	1.4	0.380	0.056	0.068	0.038	0.013

N = sample size; Na = no. of different alleles; Ne = no. of effective alleles; Ho = Observed heterozygosity; He = Expected heterozygosity; F = Fixation Index, [1 – (Ho/He)]

3.1.2 Diversity of cempedak and bangkong

Genetic diversity of the total cempedak and total bangkong population did not differ significantly for all measures except for allelic richness (Ar), which was higher in bangkong (Table 10). Genetic diversity of individual populations of cultivated cempedak and bangkong were also compared (Table 11, Table 12). Results of Student's t-test showed that, compared to cultivated cempedak populations, bangkong populations had significantly higher number of effective alleles (Ne), allelic richness (Ar), Shannon's Information Index (I), and expected heterozygosity (He) (Table 12); but other diversity statistics did not differ significantly between the two groups (Table 12). The total cempedak and total bangkong populations harbored 31 and 23 private alleles respectively.

Table 10. Summary of genetic diversity statistics and paired Student's t-test results comparing genetic diversity across 12 loci between the total cempedak and bangkong populations. Mean values and standard errors across 12 loci are reported for all diversity statistics.

Diversity statistics	Cempedak		Bangkong		t	df	p
	Mean	SE	Mean	SE			
N	***335.6***	1.2	***183.8***	0.9	-	-	-
Na	***10.3***	1.2	***9.7***	1.4	0.89	11	0.319
Ne	***2.1***	0.2	***2.8***	0.5	-1.66	11	0.130
Ar	***9.1***	1.1	***9.6***	1.4	-0.81	11	0.045**
Pa	***2.6***	0.6	***1.9***	0.5	0.89	11	0.394
I	***0.967***	0.136	***1.174***	0.182	-1.98	11	0.087*
Ho	***0.371***	0.066	***0.408***	0.058	-0.61	11	0.593
He	***0.447***	0.063	***0.525***	0.073	-1.49	11	0.179
F	***0.220***	0.055	***0.196***	0.045	0.39	11	0.784

*** p <0.01, ** p <0.05, * p<0.10
N = sample size; Na = no. of different alleles; Ne = no. of effective alleles; Ar = allelic richness; Pa = no. of private alleles; I = Shannon's Information Index; Ho = Observed heterozygosity; He = Expected heterozygosity; F = Fixation Index, [1 – (Ho/He)].

Table 11. Measures of genetic diversity by taxon, population and cultivation status across 12 loci. Only populations with N ≥ 4 are shown, but see Appendix 3 for complete data. Numbers in population code indicate state, and letters indicate locality within states (details of each population are in Table 4). Mean values of diversity measures are given as overall values for each category, except for N and P, which was summed. Pa totals (*) include private alleles from populations not shown in this table because their sample sizes were less than 4.

Site name	N	Na	Ne	Ar	Pa	I	Ho	He	F
Cempedak: Cultivated									
Bukit Panau, Kelantan	**14**	3.1	1.9	2.1	0	0.737	0.401	0.424	0.141
FELDA Kemahang, Kelantan	**37**	5.8	2.4	2.3	6	0.989	0.399	0.471	0.117
Bukit Tangga Germplasm, Kedah	**46**	3.8	1.9	2.5	0	0.723	0.404	0.378	0.007
Balik Pulau, Penang	**12**	3.4	2.0	2.2	0	0.788	0.355	0.423	0.194
Liew orchard, Penang	**18**	2.3	1.5	2.4	1	0.449	0.380	0.277	-0.159
E&J Fruit Farm, Penang	**15**	2.7	1.7	2.6	0	0.600	0.361	0.342	-0.015
Grik, Perak	**5**	2.5	2.2	2.6	1	0.680	0.517	0.398	-0.324
Bidor, Perak	**13**	2.7	1.6	2.4	0	0.517	0.305	0.289	0.058
Sungai Buloh, Selangor	**20**	3.9	2.0	1.8	0	0.782	0.334	0.409	0.197
UPM Agro-Park, Selangor	**5**	2.3	1.8	2.2	0	0.606	0.400	0.363	-0.047
Kluang Germplasm, Selangor	**53**	6.2	2.3	2.4	2	1.021	0.412	0.497	0.213
Ayer Hitam Commodity Dev. Station, Johor	**6**	2.7	2.1	2.5	0	0.694	0.347	0.395	0.139
Ulu Dusun Ag. Research Station, Sabah	**23**	3.8	1.8	2.2	2	0.742	0.385	0.393	0.100
Tenom Agricultural Park, Sabah	**4**	1.9	1.5	2.4	0	0.432	0.271	0.266	0.013
Bangkong: Wild									

Site name	N	Na	Ne	Ar	Pa	I	Ho	He	F
Sungai Rual, Kelantan	**6**	3.6	2.6	2.3	3	0.981	0.465	0.534	0.124
Ulu Geroh, Perak	**30**	4.7	2.6	2.4	3	1.029	0.517	0.533	0.001
Ulu Serendah, Selangor	**4**	2.9	2.3	2.7	0	0.826	0.458	0.471	0.003
FRIM Kepong, Selangor	**12**	4.1	2.7	2.3	2	0.966	0.422	0.483	0.171
Sungai Buloh, Selangor	**18**	3.7	2.2	3.1	0	0.860	0.433	0.465	0.049
Kota Damansara, Selangor	**29**	4.8	2.4	2.0	1	0.894	0.365	0.440	0.167
Puchong (Ayer Hitam) FR, Selangor	**6**	2.0	1.8	2.4	0	0.463	0.346	0.272	-0.280
MARDI FR, Selangor	**33**	4.4	2.3	2.4	2	0.846	0.352	0.425	0.127
Kuala Tahan, Pahang	**19**	5.2	2.6	2.2	5	1.000	0.426	0.477	0.090
Pasoh FR, Negeri Sembilan	**17**	4.1	2.4	2.5	0	0.828	0.373	0.408	0.128
Gunung Pulai, Johor	**11**	3.0	2.0	2.3	0	0.723	0.342	0.408	0.303
Cempedak: Wild/unknown cultivation status									
Beaufort Hill, Sabah (**Wild**)	**4**	1.7	1.4	2.8	0	0.328	0.208	0.211	0.002
Kampong Kopozon, Sabah (**Unknown**)	**5**	1.9	1.4	2.3	0	0.371	0.200	0.218	0.067
Sianggau FR (**Unknown**)	**7**	1.8	1.4	2.6	0	0.357	0.214	0.222	0.021
Overall									
Cempedak: Cultivated	***271***	***3.4***	***1.9***	***2.3***	***13****	***0.697***	***0.377***	***0.380***	***0.045***
	SE	*0.6*	*0.1*	*0.1*	-	*0.087*	*0.029*	*0.034*	*0.075*
Bangkong: Wild	***185***	***3.8***	***2.4***	***2.4***	***17****	***0.856***	***0.409***	***0.447***	***0.080***
	SE	*0.3*	*0.1*	*0.1*	-	*0.048*	*0.017*	*0.022*	*0.044*
Cempedak: Wild/unknown cultivation status	***16***	***1.8***	***1.4***	***2.5***	***1****	***0.352***	***0.208***	***0.217***	***0.030***
	SE	*0.1*	*0.0*	*0.1*	-	*0.013*	*0.004*	*0.003*	*0.019*

N = sample size; Na = no. of different alleles; Ne = no. of effective alleles; Ar = allelic richness; Pa = no. of private alleles; I = Shannon's Information Index; Ho = Observed heterozygosity; He = Expected heterozygosity; F = Fixation Index, [1 – (Ho/He)].

Table 12. Summary of Student's t-test comparison of diversity statistics between cultivated cempedak and bangkong.

Diversity statistic	t	df	Estimated mean value		p
			Cempedak	Bangkong	
Na	-1.08	22.9	3.4	3.8	0.290
Ne	-3.79	21.2	1.9	2.4	0.001****
Ar	-3.03	21.0	2.1	2.4	0.006****
I	-2.38	22.7	0.697	0.856	0.026**
Ho	-1.4	22.4	0.377	0.409	0.176
He	-2.34	22.0	0.380	0.447	0.029**
F	-0.59	21.7	0.045	0.080	0.564

**** p <0.005, *** p <0.01, ** p <0.05, * p<0.10

N = sample size; Na = no. of different alleles; Ne = no. of effective alleles; Ar = allelic richness; I = Shannon's Information Index; Ho = Observed heterozygosity; He = Expected heterozygosity; F = Fixation Index, [1 – (Ho/He)].

3.1.3 Diversity of Peninsular and Borneo cempedak

Genetic diversity of the total Borneo cempedak population was significantly lower compared to that of the total Peninsular cempedak population, as measured by number of different alleles (Na), number of effective alleles (Ne), number of private alleles (Pa), Shannon's Information Index (I); and nearly significantly lower allelic richness (Ar), observed heterozygosity (Ho) and expected heterozygosity (He) (Table 13). Borneo and Peninsular cempedak did not differ significantly in fixation index (F), which is a measure of inbreeding. The result was likely affected by the large difference in sampling sizes (more than twice as many cempedak trees were sampled from the Peninsula compared to Borneo). Peninsular cempedak and Borneo cempedak populations harbored 43 and 17 private alleles respectively (Table 13).

Table 13. Summary of genetic diversity statistics and paired Student's t-test results comparing genetic diversity across 12 loci between the Peninsular and Borneo cempedak populations. Mean values and standard errors across 12 loci are reported for all diversity statistics.

Diversity	**Peninsula**		**Borneo**				
statistics	**Mean**	SE	**Mean**	SE	**t**	**df**	**p**
N	159.4	1.3	62.6	0.5	-	-	-
Na	8.4	1.1	6.1	0.6	3.68	11	0.002****
Ne	2.1	0.2	1.8	0.2	2.79	11	0.015**
Ar	7.0	0.8	6.2	0.6	2.01	11	0.069*
Pa	3.7	0.6	1.3	0.3	3.68	11	0.002****
I	0.946	0.143	0.790	0.105	2.12	11	0.039**
Ho	0.375	0.067	0.313	0.055	1.56	11	0.097*
He	0.450	0.067	0.386	0.054	1.77	11	0.078*
F	0.171	0.060	0.237	0.062	-1.07	11	0.207

**** $p<0.005$, *** $p<0.01$, ** $p<0.05$, * $p<0.10$

N = sample size; Na = no. of different alleles; Ne = no. of effective alleles; Ar = allelic richness; Pa = no. of private alleles; I = Shannon's Information Index; Ho = Observed heterozygosity; He = Expected heterozygosity; F = Fixation Index, [1 – (Ho/He)].

3.2 Geographic patterns in genetic structure and differentiation

3.2.1 Structure analysis

Structure analysis was conducted on all samples (N=546) with all 12 loci. The Evanno method (2005) as implemented in Structure Harvester (Earl and VonHoldt, 2012) identified the most likely number of genetic clusters as k = 2, corresponding to predetermined cempedak and bangkong groupings. Individuals identified in the field as cempedak and bangkong sorted out into two distinct clusters ("KC" and "KB", respectively) in the inferred population structure, with a few exceptions (Figure 5). Individuals with ≥ 60% probability of membership to one group were assigned to that group, while the remaining 29 were considered admixed individuals (Appendix 4). Fourteen cempedak samples clustered with the bangkong (KB) group (hereafter referred to as "KB-cempedak"), while 18 bangkong samples clustered with the cempedak (KC) group (hereafter referred to as "KC-bangkong") (Table 14 and Table 15). Among the 14 cempedak-jackfruit hybrids, eight clustered with the KB group and the rest were admixed. These hybrids include 10 individuals of the Malaysian 'nangka-cempedak' cultivar, hybrids reported by local guides and landowners, and suspected hybrids based on morphological characters. All wild cempedak clustered with the KC group.

To assess genetic structure within cempedak and bangkong, structure analyses were conducted separately for homogenous cempedak and homogenous bangkong groups (≥ 85% probability of membership to the group). Three KB-cempedak and 12 KC-bangkong samples that did not cluster with the expected group were excluded. The best k following Evanno's method (2005) was 18 for cempedak but this value had a large error bar, so results for the second best k (k=4) was presented instead (Figure 6 and Figure 7). In bangkong, the best k following Evanno's method (2005) was k=4 (Figure 9).

Within the homogenous KC (cempedak) group, Borneo samples formed a largely distinct cluster (D, purple), except for several individuals from Tenom Agricultural Park in cluster A (blue) and Ulu Dusun Agricultural Research Station in cluster B (red) (Figure 7 and Figure 8). Genetic structure within Peninsular Malaysia was not clearly defined by geographical region, as each state had individuals assigned to different clusters (Figure 7 and Figure 8). Both germplasms comprised individuals in clusters A-C, while the Johor germplasm (Kluang) also had individuals in cluster D (Figure 7), suggesting that all major gene pools of cempedak in Malaysia were represented in the Kluang germplasm.

Within the homogenous KB (bangkong) group, the Puchong (Ayer Hitam FR) population formed a distinct cluster (D, yellow) (Figure 9 and Figure 10). Kelantan and Pahang, the two more northern populations, largely comprised cluster A (blue) (Figure 10). Perak samples were a mix between cluster A (blue) and C (gray) (Figure 9 and Figure 10). FRIM Kepong, MARDI FR, and Pasoh consisted of primarily cluster C (gray) with some cluster B (orange) (Figure 9 and Figure 10). Sungai Buloh and Kota Damansara populations were predominantly cluster B (orange) (Figure 9 and Figure 10). There was only one individual from Triang FR, and it was admixed, while the two individuals from Johor were assigned to cluster B (orange) (Figure 9 and Figure 10).

To further investigate whether geographic patterns could be detected by Structure-assigned groupings, neighbor-joining and UPGMA analyses were conducted on "K-populations". "K-populations" were designated by, which consisted of individuals assigned to KC1-4 and KB1-4 with ≥ 70% probability of membership to the group, and geographical location. Cempedak and bangkong formed clades that were sister to each other (Figure 11 and Figure 12). Within cempedak, the K-populations (KC1-KC4) sorted out cleanly, with KC2 being

basal to the other KC's (Figure 11 and Figure 12). Within bangkong, the K-populations (KB1-KB4) were not as well-resolved, as indicated by mixed clades and branching patterns that differ among trees (Figure 11 and Figure 12). In both cempedak and bangkong, there was no clear geographical pattern, as K-populations from the same location did not cluster together, but instead sorted out according to their Structure-assigned clusters (K).

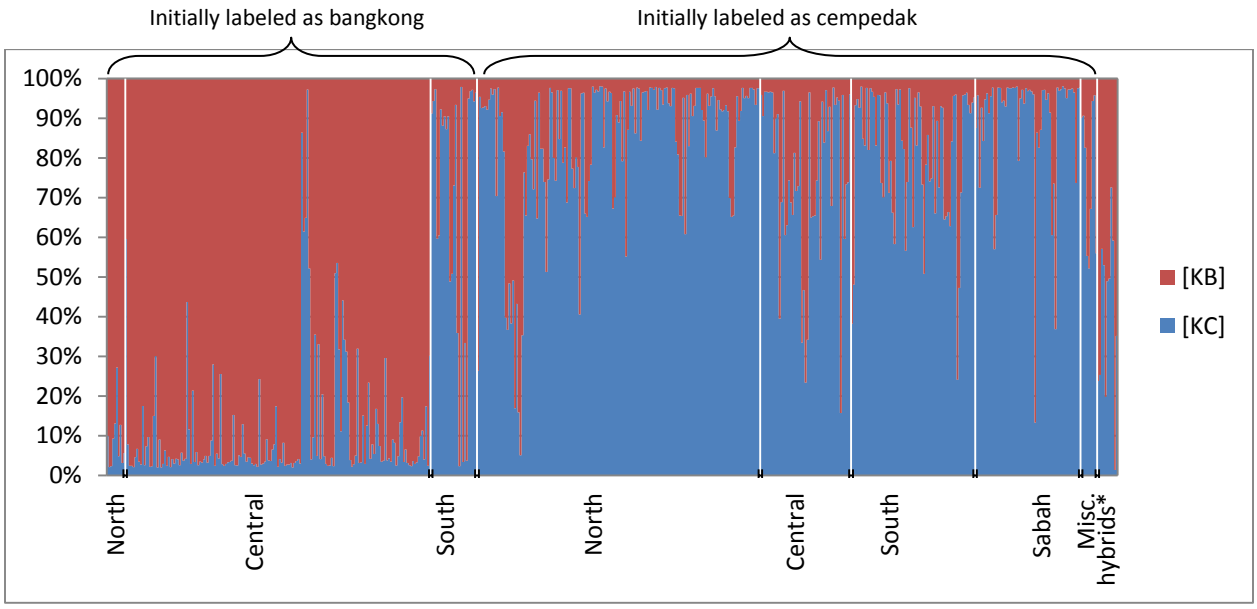


Figure 5. Inferred genetic structure (k=2) for all samples (N=546), with blue representing the cempedak group ("KC") and red representing the bangkong group ("KB"). Samples were sorted by geographical location on Peninsular Malaysia, followed by Sabah, "Misc." (includes individuals from Sarawak, Brunei, Indonesia, Vietnam, China, and Miami), and hybrids. Structure assignments agreed for the most part with initial determinations of whether samples were cempedak or bangkong.

Table 14. Cempedak samples assigned to the bangkong ("KB") group with ≥ 60% probability of membership.

Coll. no.	Site	Probability of membership to cluster		Notes
		KC	KB	
MW113	FRIM Kepong	0.335	0.665	Said to be cultivated, part of an orchard at the forest edge. May actually be bangkong.
MW356	Sungai Buloh	0.234	0.766	Said to be cempedak planted by villagers, growing near the forest where bangkong was found. May actually be bangkong.
MW357	Sungai Buloh	0.342	0.658	Same as MW356.
MW512	Bidor	0.395	0.605	From cempedak fruit in Bidor Market. Possibly a jackfruit-cempedak hybrid.
MW605	FELDAKemahang	0.398	0.602	Trees in this orchard were grown from seed sourced from Thailand. MW625 and MW627 had shoots that were covered in short white hairs which were unlike regular cempedak.
MW609	FELDAKemahang	0.367	0.633	
MW617	FELDAKemahang	0.383	0.617	
MW619	FELDAKemahang	0.169	0.831	
MW625	FELDAKemahang	0.159	0.841	
MW627	FELDAKemahang	0.051	0.949	
MW634	FELDAKemahang	0.353	0.647	
MW937b	MARDIKluang	0.241	0.759	Cultivar unknown.
NZ485b	Pattani, Thailand	0.334	0.666	From a peat swamp forest.
NZ511	Surat Thani, Thailand	0.037	0.963	Growing in a private yard.

Table 15. Bangkong samples assigned to the cempedak ("KC") group with ≥ 60% probability of membership.

Coll. no.	Site	Probability of membership to cluster		Notes
		KC	KB	
MW1101	Ulu Geroh	0.596	0.404	Samples from forest near Orang Asli village.
MW1201	Gn Lambak	0.979	0.021	Growing beside a road to a campground.
MW1204	Gn Belumut	0.934	0.066	Next to main entrance arch sign (Hutan Lipur Gn Belumut) by road.
MW1206	Lenggor FR	0.731	0.269	Park area of Lenggor Forest Reserve, near shelter huts (pondok). Young fruits had sweet smell.
MW1208	Gn Pulai	0.912	0.088	Trees were growing next to a road on a hill. Some trees sampled were young saplings covered in soft white hairs. MW1215 had both lobed and entire leaves, and some leaves were a very dark green-olive color unlike regular bangkong.
MW1209	Gn Pulai	0.944	0.056	
MW1210	Gn Pulai	0.974	0.026	
MW1211	Gn Pulai	0.597	0.403	
MW1213	Gn Pulai	0.605	0.395	
MW1214	Gn Pulai	0.923	0.077	
MW1215	Gn Pulai	0.882	0.118	
MW1216	Gn Pulai	0.905	0.095	
MW1217	Gn Pulai	0.873	0.127	
MW1218	Gn Pulai	0.905	0.095	
MW1301	Ulu Serendah	0.864	0.136	Samples from forest near Orang Asli village.
MW1302	Ulu Serendah	0.615	0.385	
MW1303	Ulu Serendah	0.649	0.351	
MW1304	Ulu Serendah	0.973	0.027	

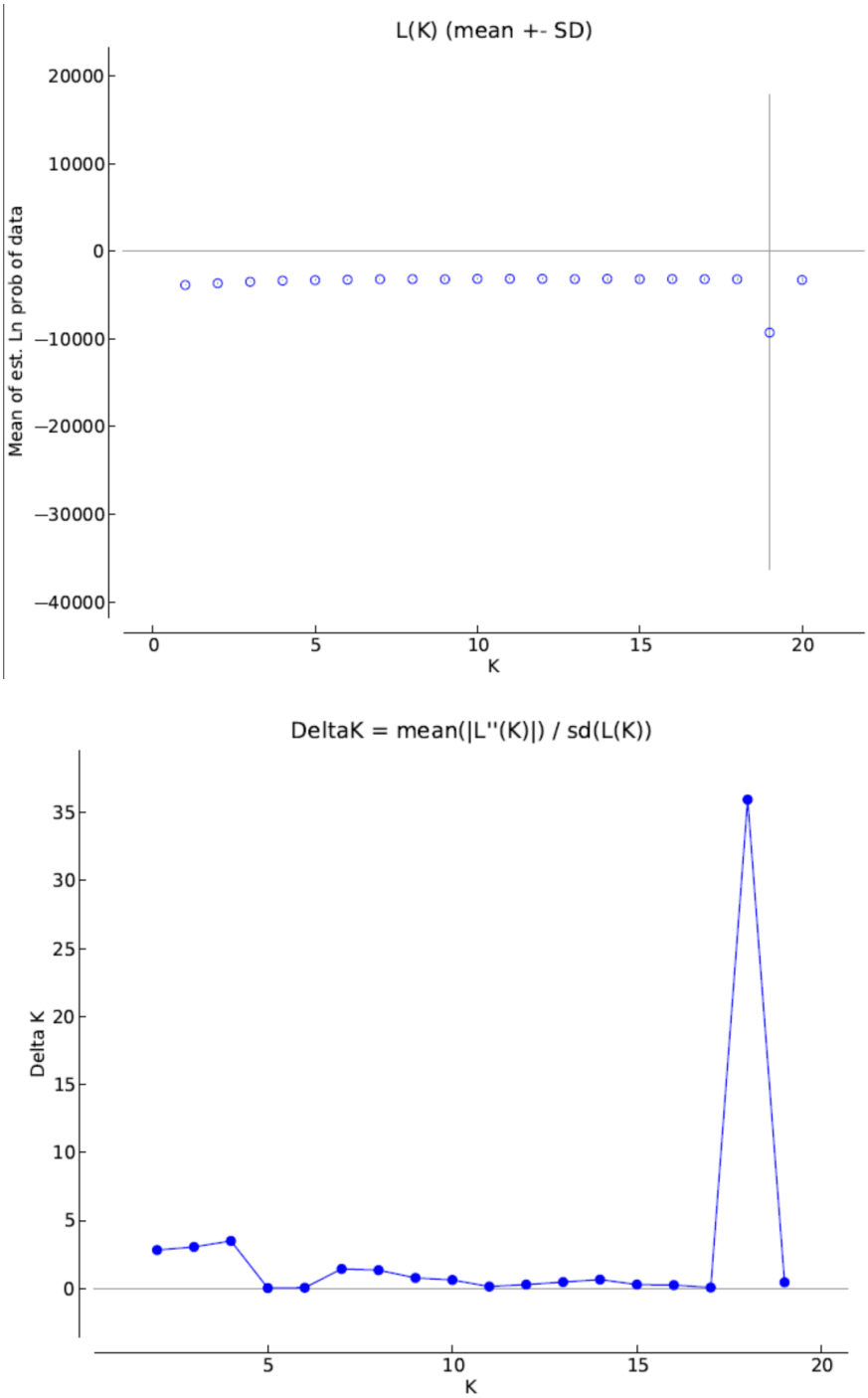


Figure 6. Determination of most likely k for the homogenous cempedak group based on Evanno's method (2005). At k=18, the mean of Ln(k) showed an extremely large standard deviation, which may lead to biased estimations of DeltaK. As such, the second most likely value of k=4 was selected instead.

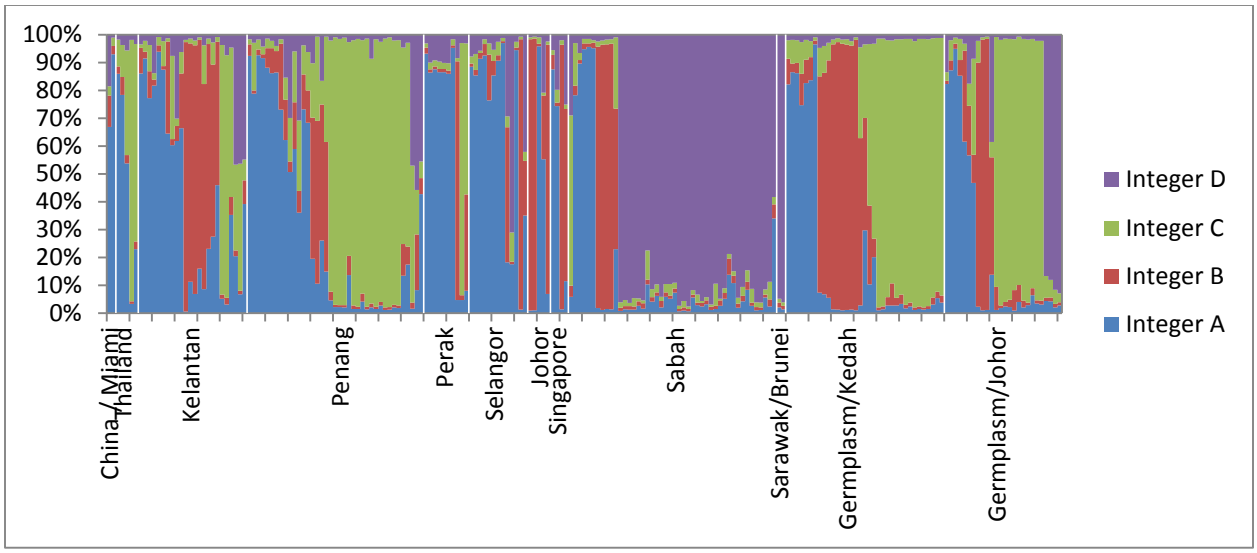


Figure 7. Genetic structure (k=4) within the homogenous cempedak ("KC") group (N = 211). Samples are sorted geographically by country/state, roughly from north to south, except for germplasm samples which are shown to the right.

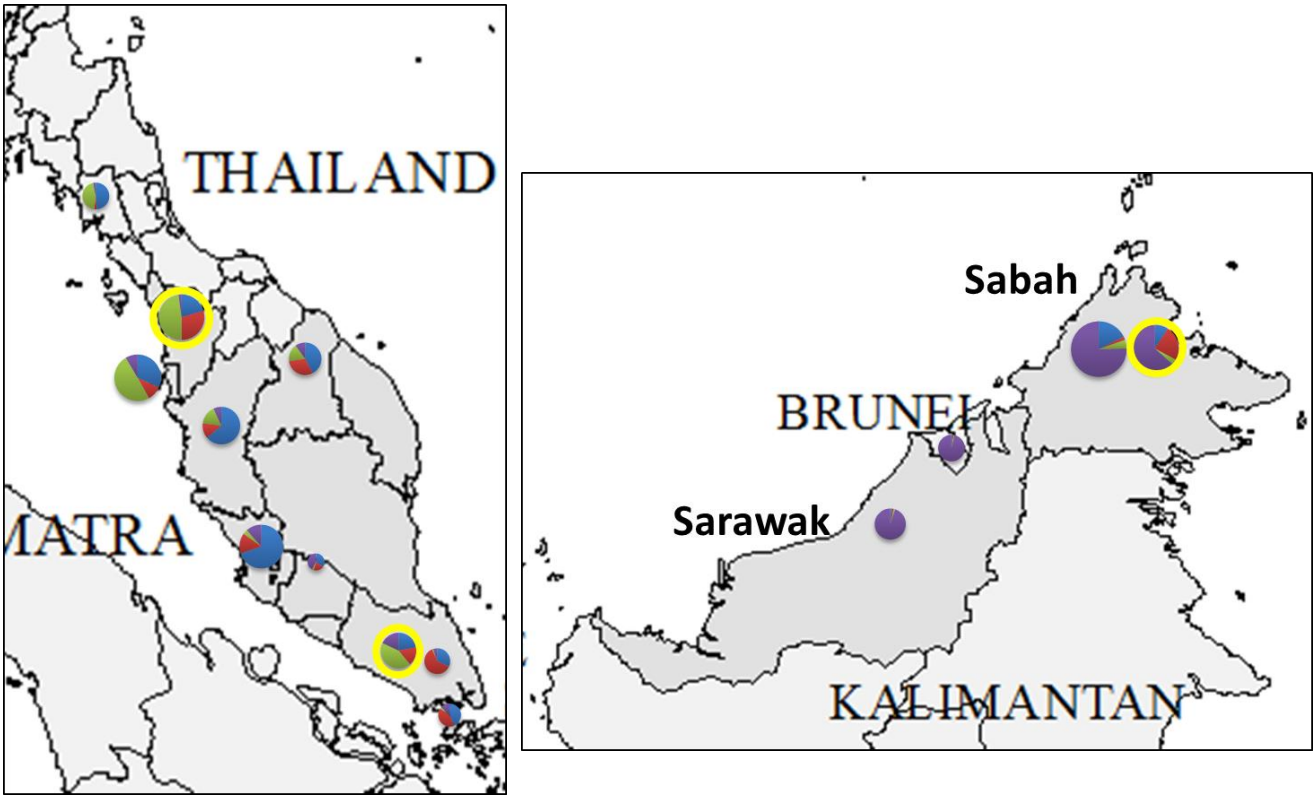


Figure 8. Structure results (K=4) of homogenous cempedak ("KC") group, mapped by country/states, with Peninsular Malaysia on the left, and Borneo on the right. The sizes of the pie charts represent sample sizes. Maps are not to scale.

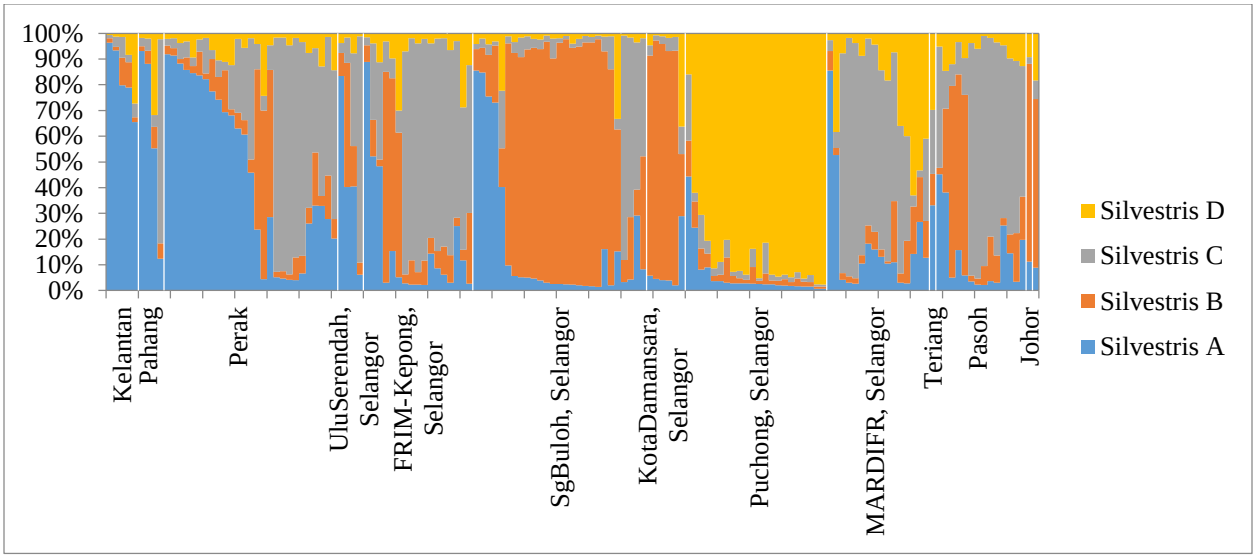


Figure 9. Genetic structure (k=4) within the homogenous bangkong ("KB") group (N=145). Samples are sorted geographically by state, roughly from north to south. The Puchong population formed a distinct cluster (D, yellow).

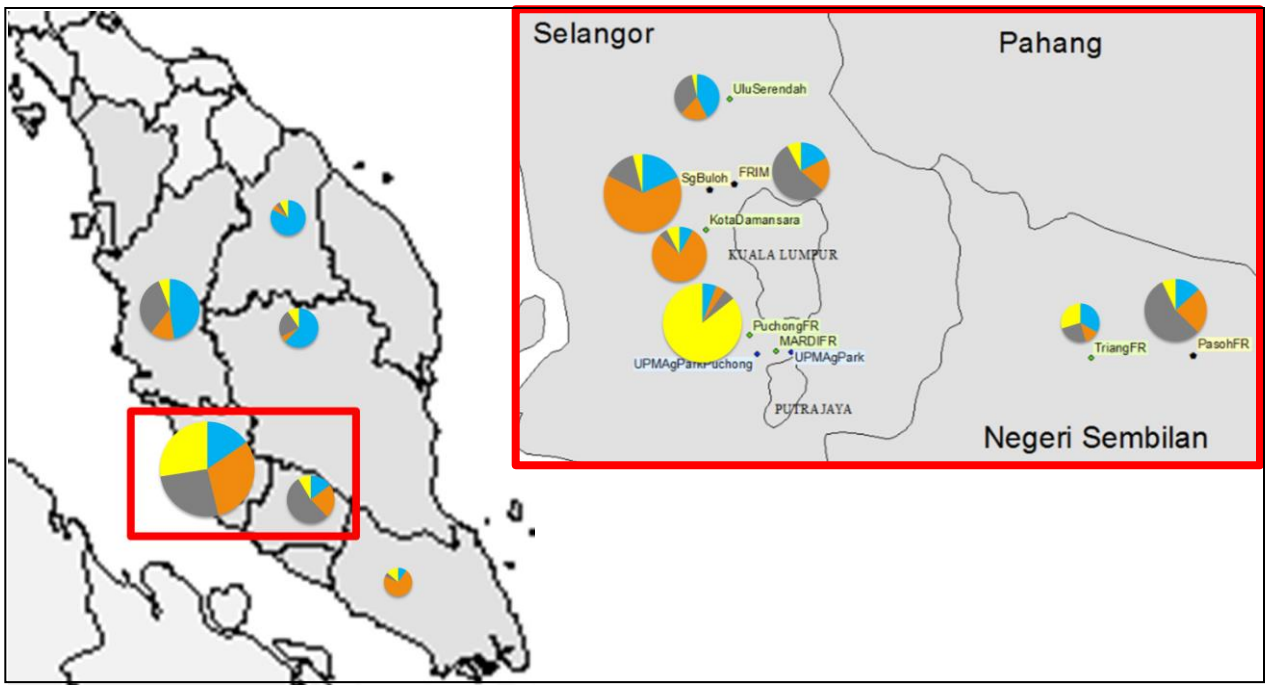
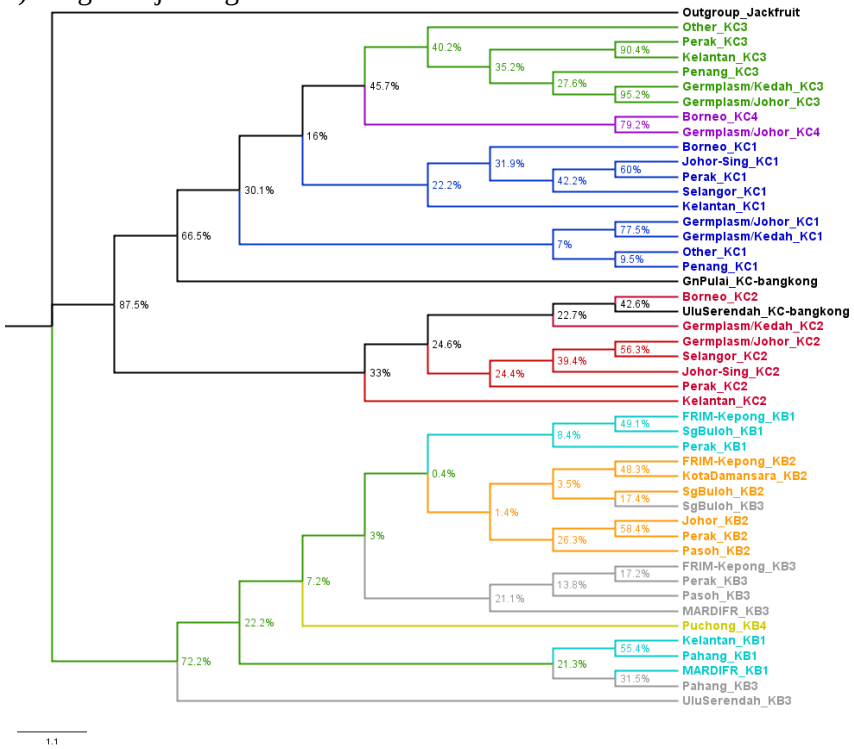


Figure 10. Structure results (K=4) of homogenous bangkong ("KB") group, mapped by states in Peninsular Malaysia. Map inset shows a close-up of bangkong populations in the states of Selangor and Negeri Sembilan. The sizes of the pie charts represent sample sizes. Maps are not to scale.

a) Neighbor-joining



b) UPGMA

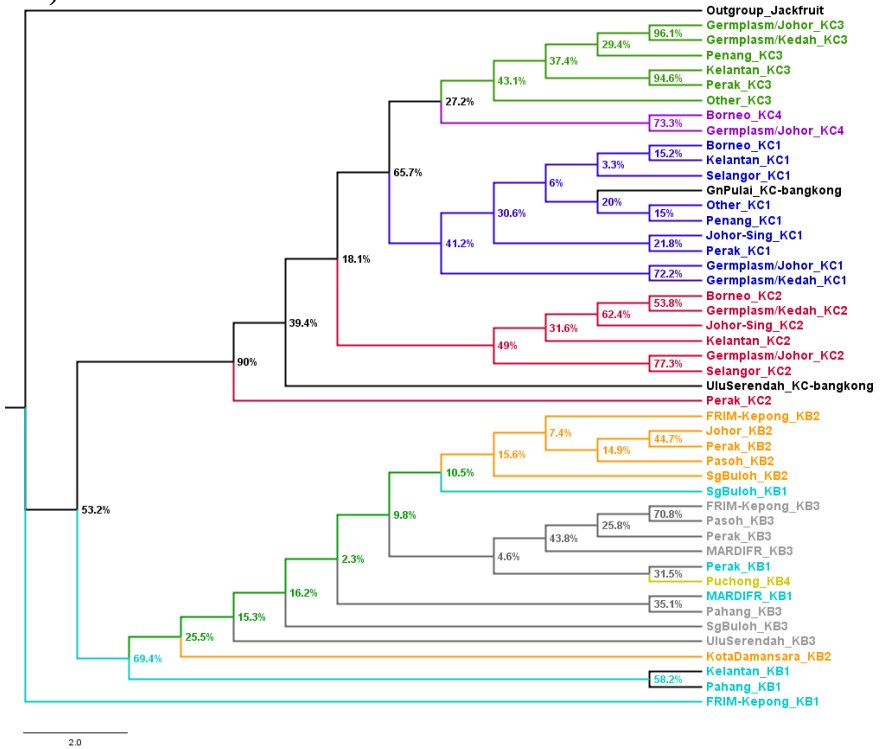
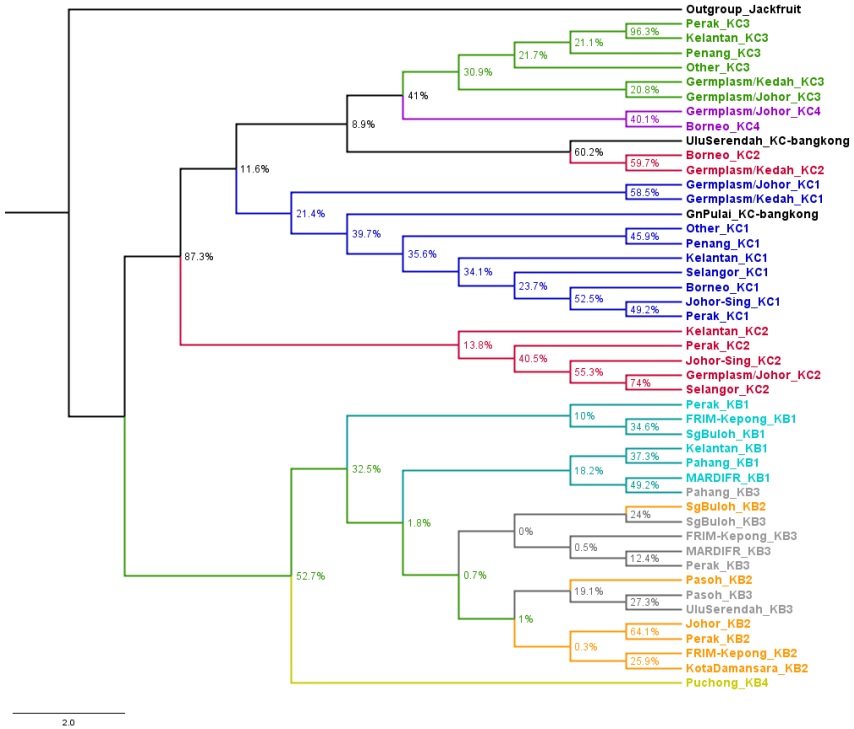


Figure 11. a) Neighbor-joining and b) UPGMA trees based on pairwise Nei's genetic distance (Da) between K-populations, color-coded according to their Structure-assigned clusters (K) within the homogenous cempedak group - KC1-KB4 (Dark Blue, Red, Green, Purple) – or homogenous bangkong group - KB1-KB4 (Light Blue, Orange, Gray, Yellow). The outgroup (Jackfruit or *Artocarpus heterophyllus*) and two bangkong populations that were assigned to the cempedak group in Structure are in Black. Bootstrap values are indicated at nodes.

a) Neighbor-joining



b) UPGMA

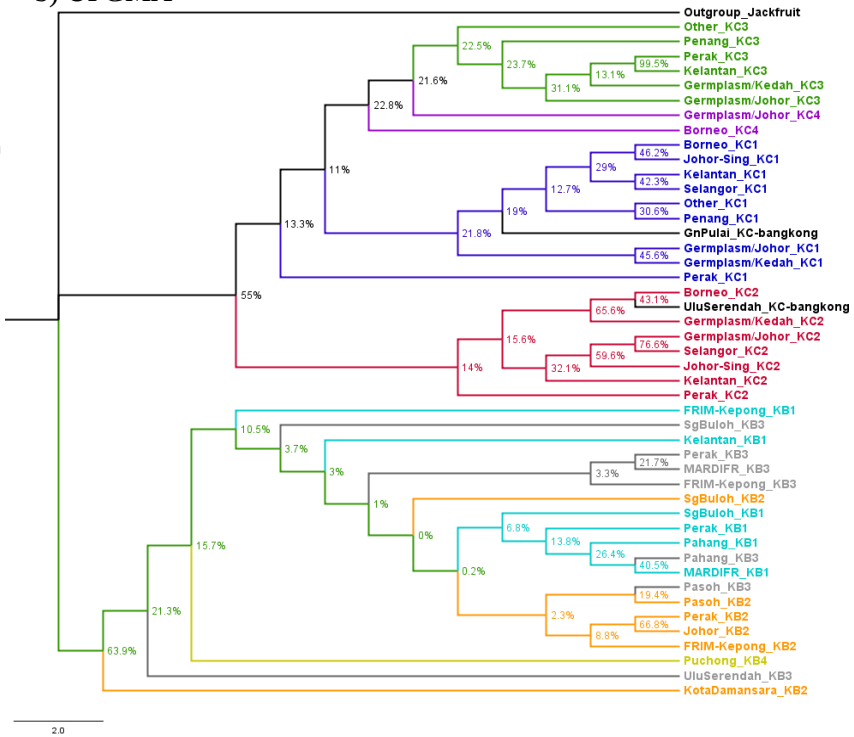


Figure 12. a) Neighbor-joining and b) UPGMA trees based on pairwise Fst values between K-populations, color-coded according to their Structure-assigned clusters (K) within the homogenous cempedak group - KC1-KB4 (Dark Blue, Red, Green, Purple) – or homogenous bangkong group - KB1-KB4 (Light Blue, Orange, Gray, Yellow). The outgroup (Jackfruit or *Artocarpus heterophyllus*) and two bangkong populations that were assigned to the cempedak group in Structure are in Black. Bootstrap values are indicated at nodes.

3.2.2 Population Differentiation

AMOVA was conducted for individual cempedak and bangkong populations with minimum sample sizes of four. Results of the AMOVA showed that 78% of the variance was accounted for within populations, 9% among populations, and 13% among categories of cultivation status (Table 16). Principal coordinate analysis (PCoA) of pairwise-population Fst values from the AMOVA revealed differentiation between bangkong and cultivated cempedak populations, although there was some overlap in 95% confidence-interval ellipses (Figure 13). In particular, one bangkong population (Gunung Pulai, Johor) clustered with the cultivated populations outside the confidence ellipse (Figure 13). This was in line with Structure analysis where individuals from Gunung Pulai clustered with the cempedak group (Figure 5). The cempedak populations with unknown cultivation status clustered with the cultivated cempedak populations, whereas the wild cempedak population (Beaufort Hill, Sabah) was isolated from other populations (Figure 13).

AMOVA was also conducted to assess how genetic diversity of cempedak and bangkong is partitioned across broader geographic regions. Cempedak and bangkong samples were sorted into eight sub-categories based on cultivation status and geographic region. AMOVA results showed that 86% of the variance was accounted for within categories, 11% among categories, and 3% among geographical regions (Table 17). Pairwise comparison of Fst values based on Wright's (1978) guidelines showed that wild Peninsula bangkong and wild Borneo cempedak were moderate to greatly differentiated from all other categories (Table 18). Little to no differentiation was found among cultivated cempedak across geographical regions, nor among cultivated and unknown cempedak within the same region (Table 18). However, there was moderate differentiation between unknown Borneo cempedak with cultivated and unknown Peninsular cempedak (Table 18).

Principle coordinate analysis (PCoA) of all samples (N = 545) showed that bangkong and cempedak individuals clustered apart from each other, with overlaps in confidence intervals (Figure 14). The first two PCA axes explained 71.3% of the overall genetic variation. Cultivated cempedak from Peninsular Malaysia and Borneo had very similar confidence interval ellipses, and wild Borneo cempedak was enclosed within the confidence interval ellipse of cultivated Borneo cempedak (Figure 14). The clustering of wild Borneo cempedak with cultivated Borneo cempedak in this plot is in contrast with the earlier PCoA of population-pairwise Fst, where the wild cempedak population from Sabah was isolated from other populations (Figure 13 and Figure 14). Cempedak-jackfruit hybrids formed a gradient between the bangkong and cempedak clusters (Figure 14).

Table 16. Summary of results for AMOVA of individual cempedak and bangkong populations with $N \geq 4$, as listed in Table 11. Cultivation categories (or cultivation/taxon) were defined for each population as "cultivated cempedak", "bangkong", or "wild/unknown cempedak", as listed in Table 11.

Source	df	SS	MS	Est. Var.	%
Among cultivation categories	2	246.615	123.307	0.455	13%
Among populations	25	318.756	12.750	0.304	9%
Within populations	916	2467.346	2.694	2.694	78%
Total	943	3032.717		3.453	100%

df = degrees of freedom, SS = sum-of-squares, MS = mean square, Est. Var. = estimated variance

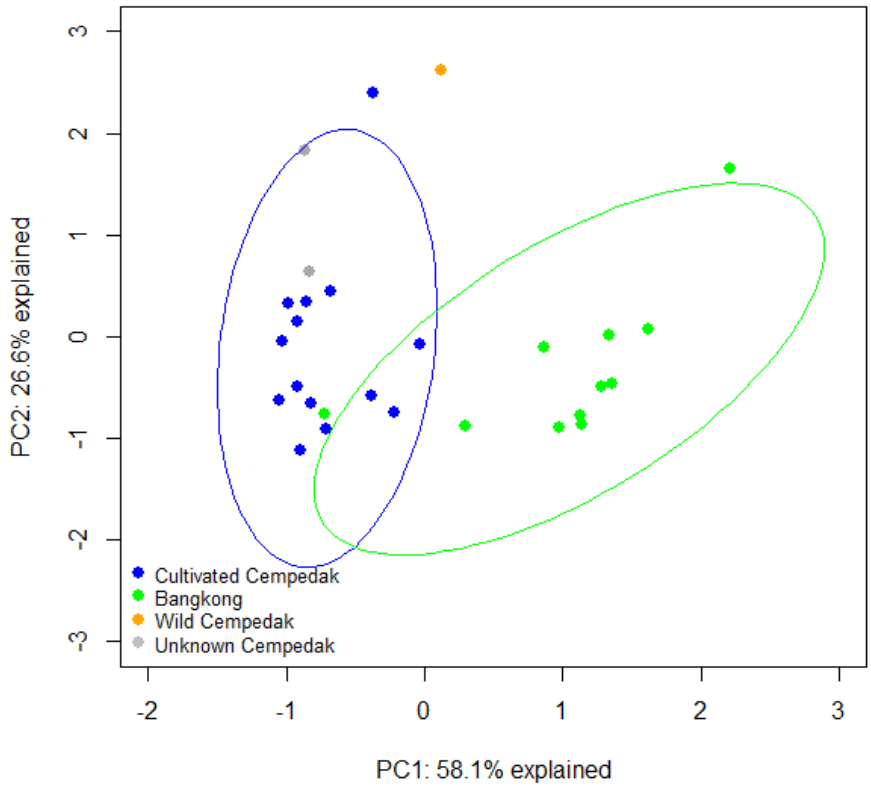


Figure 13. Principal coordinate analysis (PCoA) of pairwise-population Fst values. Cultivated cempedak and bangkong 95% interval ellipses are shown in blue and green.

Table 17. Summary of results for AMOVA of cempedak and bangkong at the regional level. Cempedak and bangkong samples were sorted into eight sub-categories based on cultivation status and geographic region. The sub-categories are: cultivated cempedak from the Peninsula, Borneo, and other regions; cempedak of unknown cultivation status from the Peninsula, Borneo, and other regions; wild cempedak from Borneo; and bangkong from the Peninsula.

Source	df	SS	MS	Est. Var.	%
Among geographical regions	2	70.370	35.185	0.065	2%
Among sub-categories	5	228.795	45.759	0.388	11%
Within sub-categories	1054	3106.173	2.947	2.947	87%
Total	1061	3405.337		3.400	100%

df = degrees of freedom, SS = sum-of-squares, MS = mean square, Est. Var. = estimated variance

Table 18. Pairwise Fst values across geographic regions and cultivation status. Geographic regions are Peninsular Malaysia ("Penin."), Borneo, and "Other" (comprising Sumatra, Thailand, Vietnam, China, and USA). Cultivation/taxon status are categorized as cultivated cempedak ("cult. cemp."), bangkong, wild cempedak ("wild cemp."), and cempedak of unknown cultivation status ("unk. cemp."). Fst values were interpreted based on Wright's guidelines (1978), i.e. 0 - 0.05, little differentiation (no color); 0.05 – 0.15, moderate differentiation (light yellow); 0.15 – 0.25, great differentiation (darker yellow); 0.25, very great differentiation (none - there were no values >0.25). All values were significant based on statistical testing by random permutation (999 permutations) unless indicated by (NS) following the number.

	Cult. cemp. Penin.	Bangkong Penin.	Unk. cemp. Penin.	Cult. cemp. Borneo	Wild cemp. Borneo	Unk. cemp Borneo	Cult. cemp Other
Cult. cemp. Penin.	-						
Bangkong Penin.	0.133	-					
Unk. cemp. Penin.	0.030	0.087	-				
Cult. cemp. Borneo	0.025	0.185	0.048	-			
Wild cemp. Borneo	0.109	0.226	0.120	0.067	-		
Unk. cemp Borneo	0.058	0.215	0.120	0.037	0.074	-	
Cult. cemp Other	0.014 (NS)	0.152	0.039	0.020 (NS)	0.130	0.098	-
Unk. cemp Other	0.000 (NS)	0.158	0.034 (NS)	0.009 (NS)	0.138	0.065	0.005 (NS)

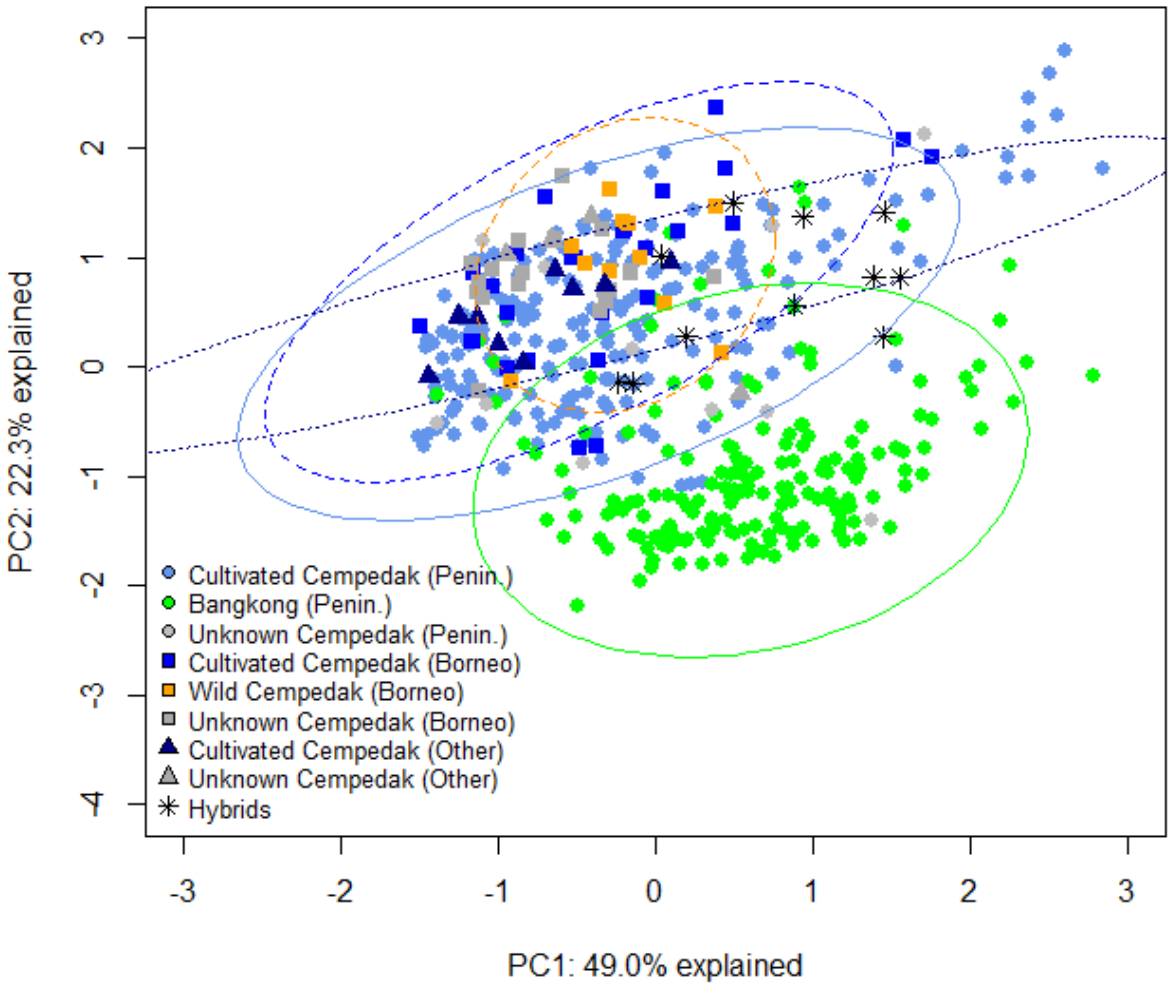


Figure 14. Principle coordinate analysis (PCoA) of all cempedak and bangkong individuals (N=545), color-coded by cultivation status (blue: cultivated cempedak; green: bangkong; red: wild cempedak; black: cempedak of unknown cultivation status), while shapes represent different geographic regions (circle: Peninsular Malaysia; square: Borneo; triangle: Other, i.e. all other regions). Putative hybrids from all regions are represented by asterisks (*). Ellipses representing 95% confidence intervals are shown for cultivated cempedak (light blue/solid line: Peninsular, medium blue/dashed line: Borneo, dark blue/dotted line: Other) Borneo wild cempedak (orange/dashed line), and bangkong (green/solid line).

3.2.3 Isolation-by-distance

Mantel tests of isolation-by-distance among bangkong populations were significant based on Nei's genetic distance ($p = 0.013$, $r^2 = 0.3256$) but not significant based on Fst ($p = 0.142$) (Table 19). Mantel tests of isolation-by-distance among cultivated cempedak populations were not significant for both genetic distance measures (Table 19). For both measures, bangkong populations showed greater increase in genetic distance with geographic distance compared to cultivated cempedak populations (Table 19).

Table 19. Mantel tests of matrix correspondence between pairwise-population geographic distance (km), and genetic distance (Fst or Nei's genetic distance, D).

Population type	**Genetic distance**	**Equation**	**R^2**	**P-value**
Cultivated cempedak	Fst	$y = -7 \times 10^{-6} x + 0.0788$	0.0183	0.373
	D	$y = -6 \times 10^{-6} x + 0.115$	0.0048	0.450
Bangkong	Fst	$y = 0.0001x + 0.076$	0.1135	0.142
	D	$y = 0.0005x + 0.127$	0.3256	0.013**

3.3 Phylogenetic analysis of cempedak and bangkong

Neighbor-joining and UPGMA trees based on population-pairwise Nei's genetic distance (Da) values showed that most bangkong populations cluster more with one another than with cempedak populations (Figure 15; UPGMA tree not shown because it was nearly identical to the neighbor-joining tree). Two bangkong populations (Gunung Pulai and Ulu Serendah) fell into the cempedak "clade" with low bootstrap support, in line with Structure results where individuals from those two populations clustered with the cempedak group (Figure 15, Appendix 4). In addition, the bangkong population from Ulu Serendah is basal to all the cempedak populations in the neighbor-joining trees (Figure 15 and Figure 16).

Neighbor-joining and UPGMA trees based on population-pairwise Fst values were in line with the PCoA plots presented earlier (Figures 13-14, Figures 16-17). Most bangkong populations clustered more with one another than with cempedak populations. Interestingly, the wild cempedak population (Beaufort Hill, Sabah) sorted out as a sister group to cempedak and bangkong clades in the UPGMA tree (Figure 17), corresponding with the first PCoA plot where the Beaufort Hill population was isolated from other populations (Figure 13). In contrast, the wild cempedak population (Beaufort Hill, Sabah) fell into the cempedak "clade" in the neighbor-joining tree (Figure 16), corresponding with the second PCoA plot based on genetic distance between individuals, where wild cempedak individuals clustered with cultivated cempedak individuals (Figure 14).

Neighbor-joining and UPGMA trees based on individual-pairwise genetic distances show the same trends as the population-pairwise trees (Figure 18 and Figure 19). Bangkong (green) forms a distinct group, but several bangkong individuals are interspersed within the cempedak groups (Figure 18 and Figure 19). Wild cempedak (in orange) and cempedak of unknown cultivation status (in red) are interspersed throughout the cultivated cempedak groups (in black)

(Figure 18 and Figure 19). Putative cempedak-jackfruit individuals (in purple) were more closely clustered with the jackfruit outgroup, although there were a few putative hybrid individuals that sorted out further away from the jackfruit outgroup (Figure 18 and Figure 19).

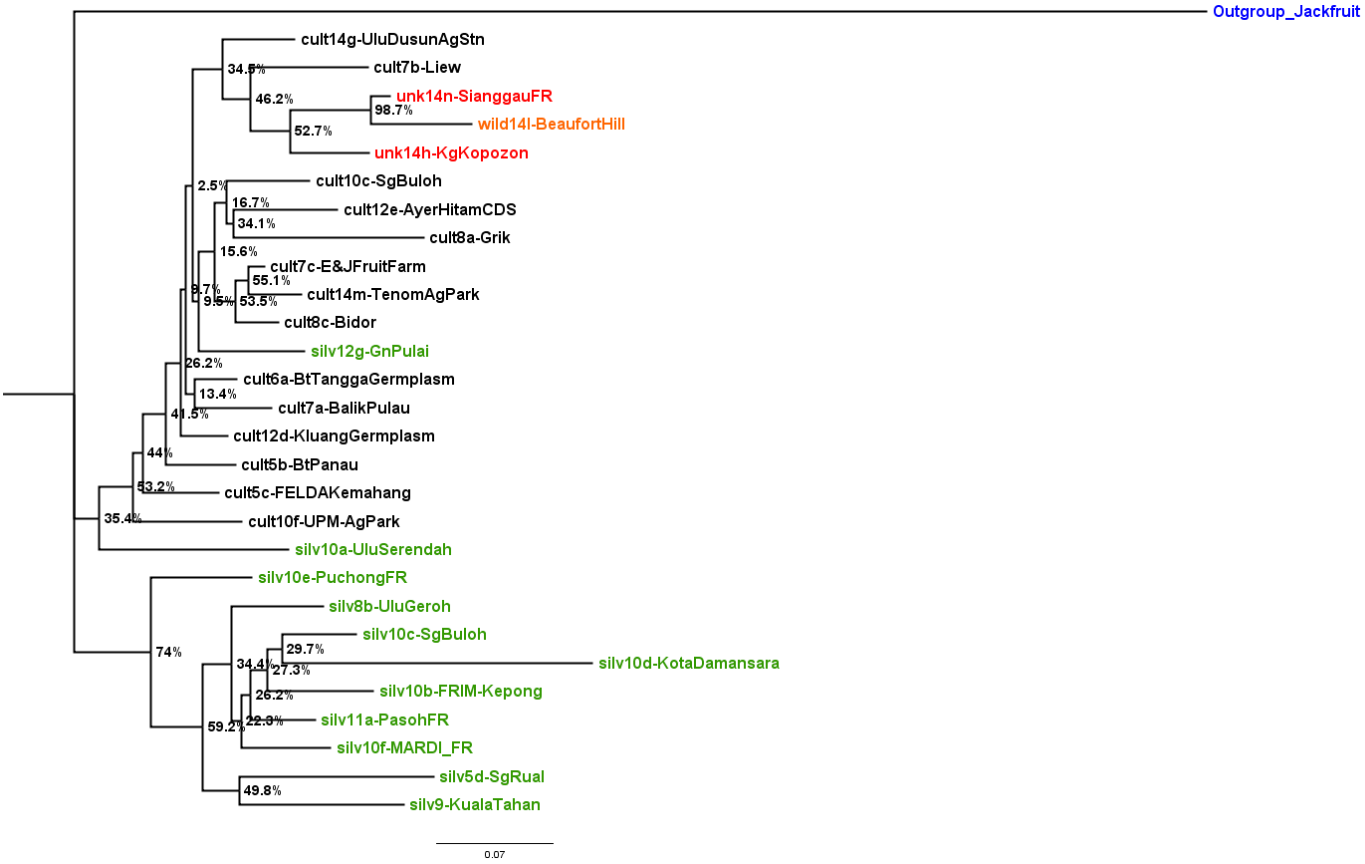


Figure 15. Neighbor-joining tree based on pairwise Nei's genetic distance (Da) between populations. Bootstrap values are indicated at nodes. Color codes: Blue: Outgroup (Jackfruit or *Artocarpus heterophyllus*), Black: Cultivated cempedak, Green: Bangkong, Orange: Wild cempedak, Red: Cempedak of unknown cultivation status. See Table 4 for explanation of population codes.

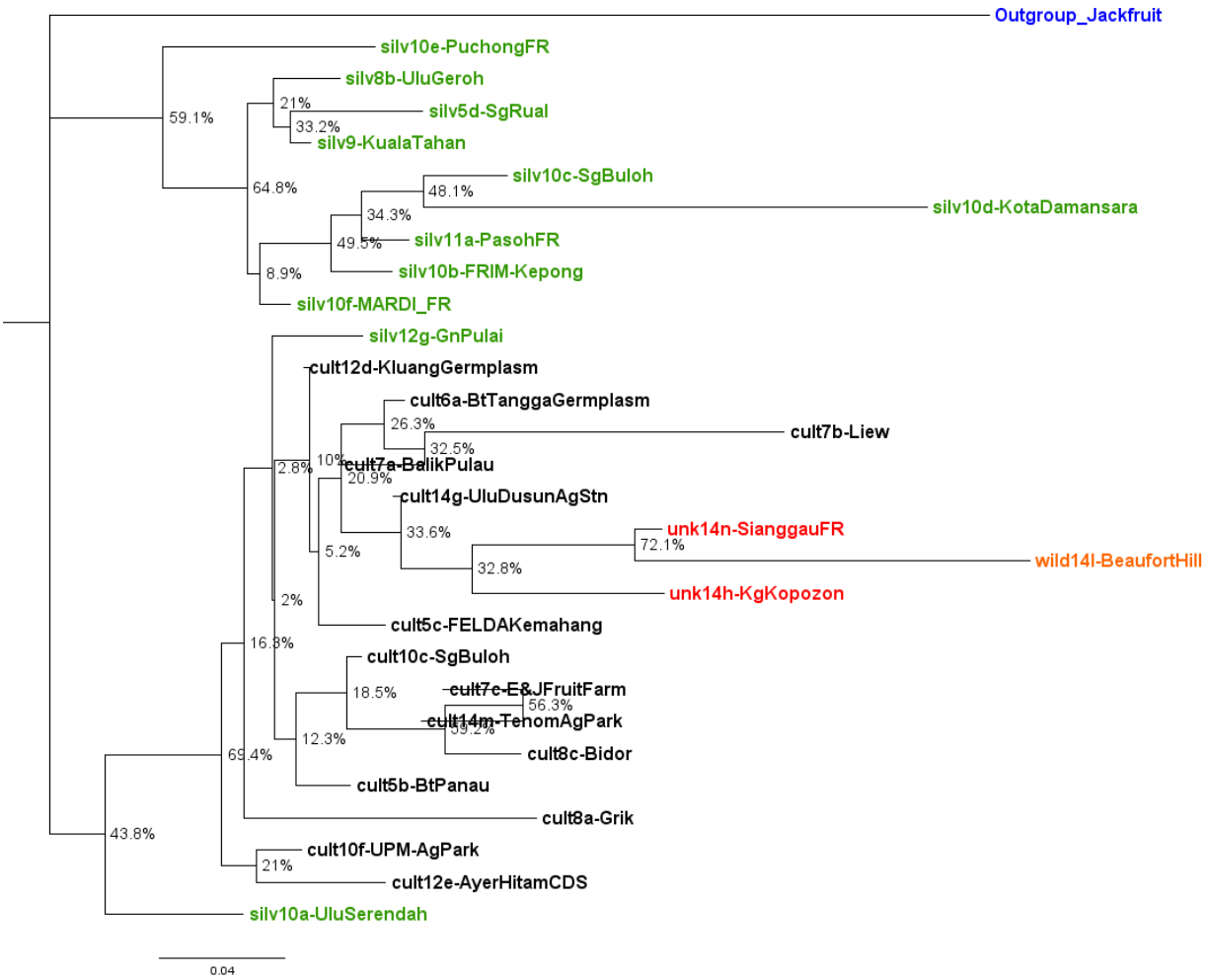


Figure 16. Neighbor-joining tree based on pairwise Fst values between populations. Bootstrap values are indicated at nodes. Color codes: Blue: Outgroup (Jackfruit or A. heterophyllus), Black: Cultivated cempedak, Green: Bangkong, Orange: Wild cempedak, Red: Cempedak of unknown cultivation status. See Table 4 for explanation of population codes.

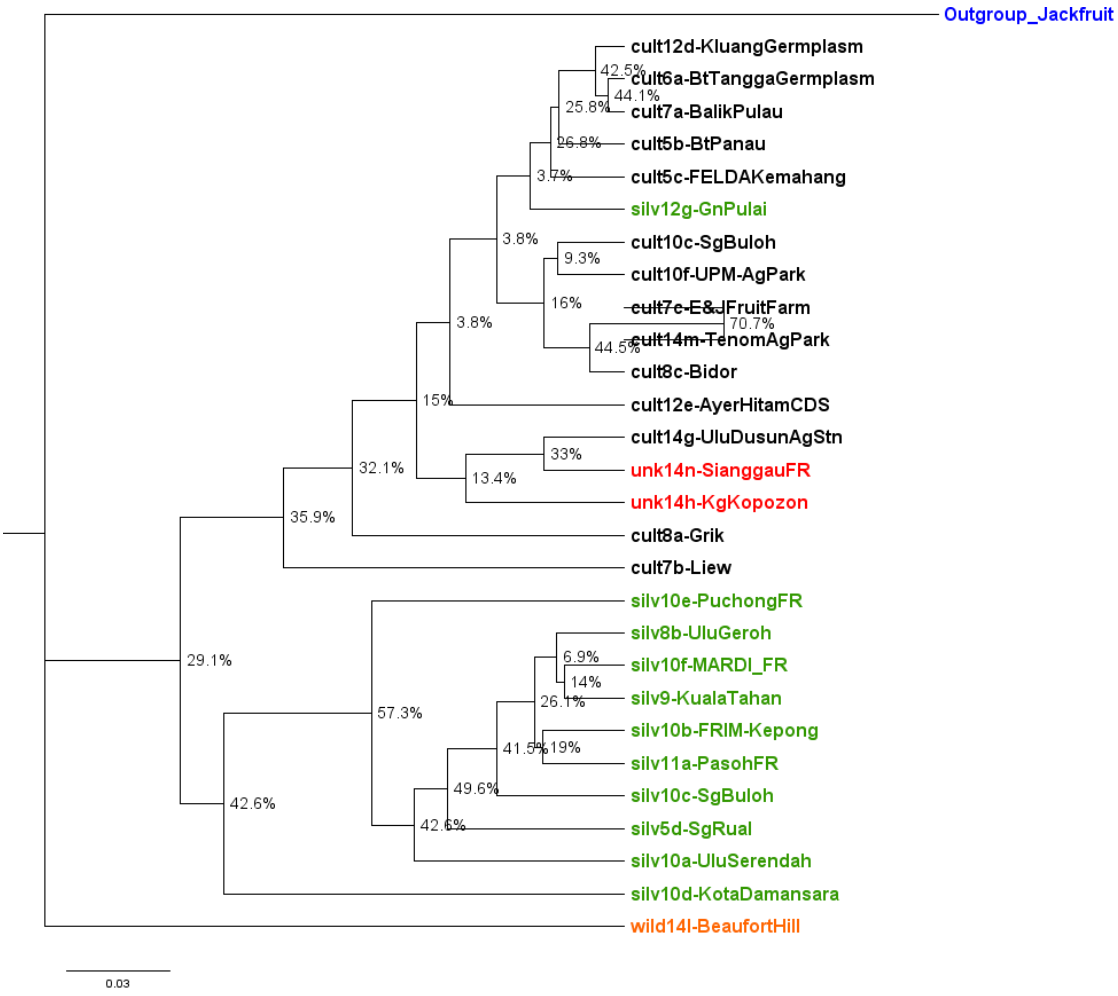


Figure 17. UPGMA tree based on pairwise Fst values between populations. Bootstrap values are indicated at nodes. Color codes: Blue: Outgroup (Jackfruit or A. heterophyllus), Black: Cultivated cempedak, Green: Bangkong, Orange: Wild cempedak, Red: Cempedak of unknown cultivation status. See Table 4 for explanation of population codes.

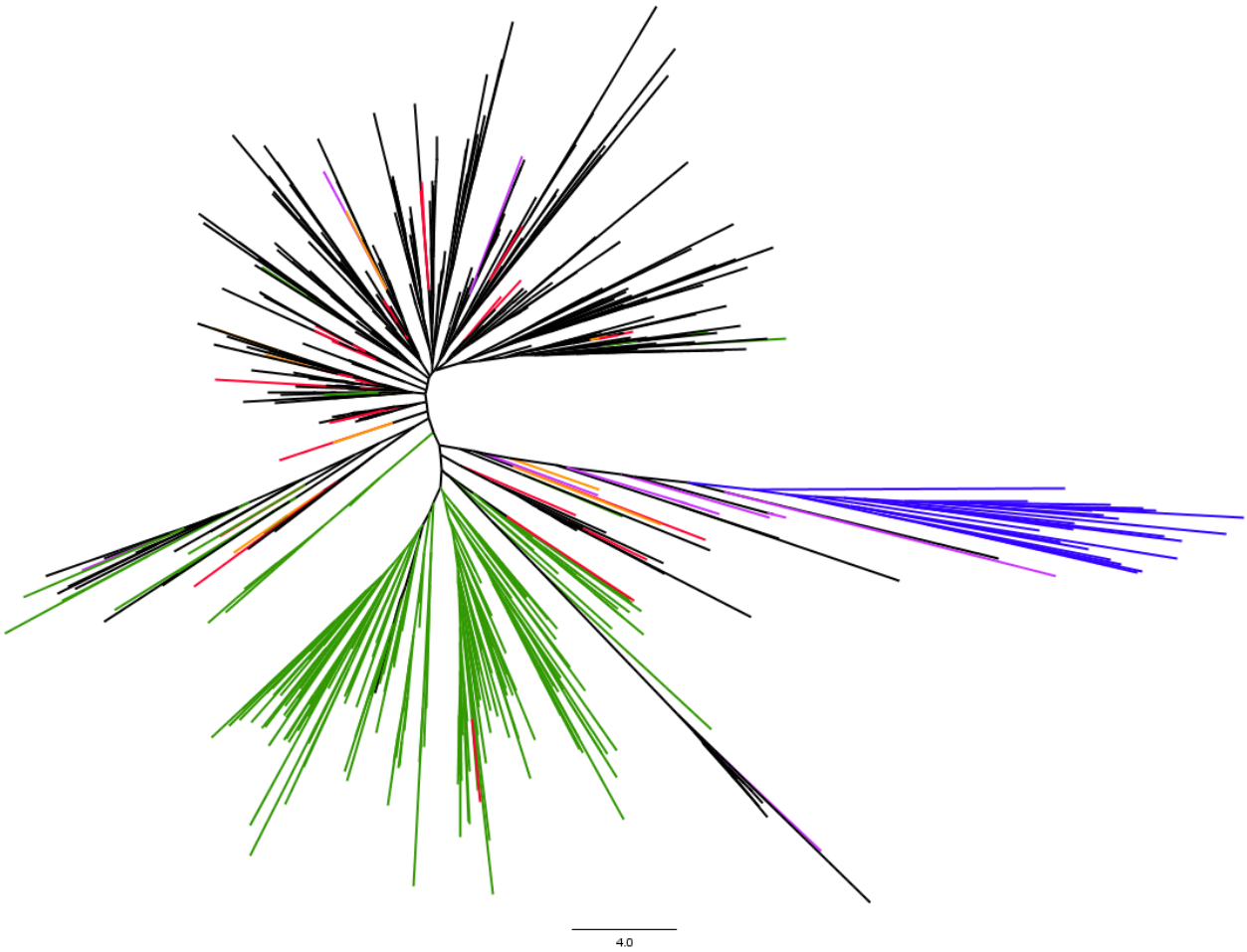


Figure 18. Radial neighbor-joining tree based on pairwise genetic distances between individuals. Color codes: Blue: Outgroup (Jackfruit or *A. heterophyllus*), Black: Cultivated cempedak, Green: Bangkong, Orange: Wild cempedak, Red: Cempedak of unknown cultivation status, Purple: Putative cempedak-jackfruit hybrids.

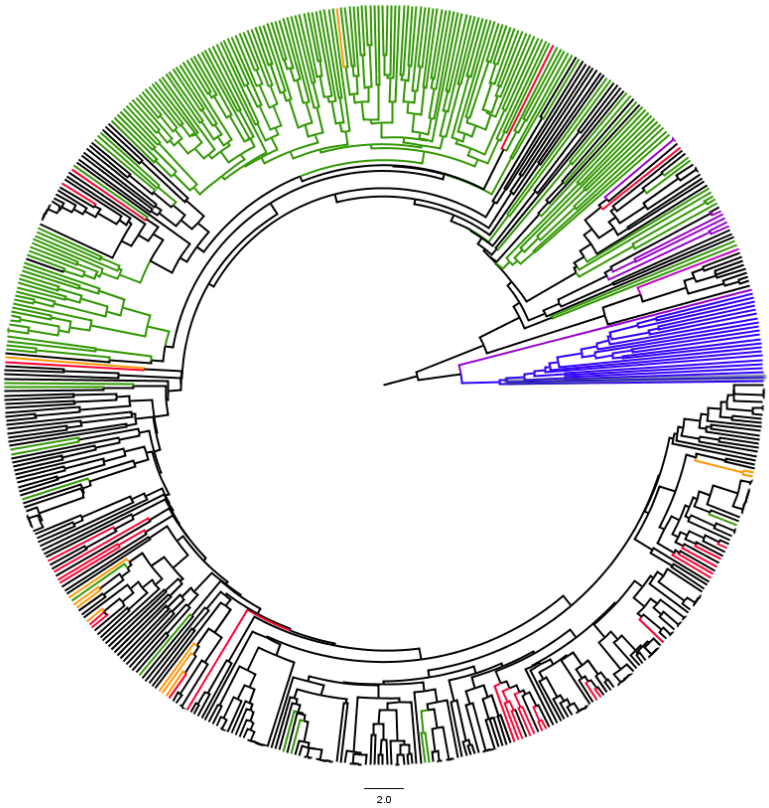


Figure 19. Polar UPGMA tree based on pairwise genetic distances between individuals. Color codes: Blue: Outgroup (Jackfruit or *A. heterophyllus*), Black: Cultivated cempedak, Green: Bangkong, Orange: Wild cempedak, Red: Cempedak of unknown cultivation status, Purple: Putative cempedak-jackfruit hybrids.

3.4 Analysis of cempedak cultivars

3.4.1 Cultivar names, clonal assignment and structure

Twenty-five different cultivar names were recorded across 101 cultivated cempedak samples in this study. This number is not representative of the range of cultivars sampled, because cultivar names were not recorded for many samples (e.g., trees without tags at the Kluang germplasm). To identify clonal relationships between samples, a clonal assignment analysis was conducted on all cempedak and bangkong samples (N=476 after removing samples with missing data) in GENODIVE (Meirmans and Van Tienderen, 2004). One hundred and twenty-one individuals (25%) shared genotypes with at least one other individual (Figure 20; Appendix 4). Only three cultivar names (durian-cempedak, Kuala Nerang, and Sik) were each associated with a single genotype (true-to-type), while seven cultivar names were singletons, and all other cultivar names were represented by more than one genotype (Table 20). The jackfruit-cempedak hybrid cultivar, known locally as "nangka-cempedak", was the least predictive of genotype, as none of the cultivar samples shared the same genotype.

Six bangkong samples shared a genotype with one other bangkong or cempedak individual (Figure 20). Four of those bangkong samples were from the same forest (Ayer Hitam FR, Selangor), with MW1601 and MW1605 sharing one genotype, and MW1625 and MW1626 sharing another genotype. The other two bangkong individuals were sampled from different localities in Johor. One individual (MW1205, Lenggor FR) shared the same genotype as cultivars of CH27, while the other (MW1201, Gunung Lambak in Johor) shared the same genotype as a cultivated cempedak collected in another state without a cultivar name or provenance information (MW367, Sungai Buloh in Selangor).

Structure analysis of 116 samples (101 cempedak samples with recorded cultivar names + 15 samples from E&J Fruit Farm), revealed genetic structure within the same cultivar names

(Figure 21). The most likely k was 21 based on the Evanno (2005) method. Many cultivar groups sorted out into distinct genetic clusters; whereas several cultivar names showed a high percentage of admixture (Berseri*, Bukit Selambau 1 &2, E&J Orange) (Figure 21). With the exception of one nangka-cempedak individual (NZ632 from Ulu Dusun Agriculture Research Station, Sabah) and one putative hybrid (NZ812 from Kipandi Butterfly Park), the nangka-cempedak cultivars/putative hybrids formed a distinct and fairly homogenous cluster. At a coarser resolution of k=2 (the most likely k under 10), Sri Harum 2, CH27, CH28, nangka-cempedak, and putative hybrids clustered into one homogenous group, while all the other samples clustered into another homogenous group (Figure 21). "E&J Orange" was included as a putative "cultivar" in the Structure analysis, based on the owner's reports that the trees produced bright-orange fruits that were distinct from other cempedak cultivars on the market. The orchard owner also said that there were two types of trees in his orchard, based on the shapes of the leaves. However, Structure analysis showed that most of the E&J samples had a high degree of admixture (Figure 21). The "Red Balik Pulau" cultivar (MW506) had a similar genetic composition to some "E&J Orange" individuals, and could have been sourced from the E&J Fruit Farm located in Balik Pulau (Figure 21).

Some specimens labeled with the same cultivar name had leaves that looked different (MW837-838 looked different from MW839-840 although all four were labeled as Berseri; MW807 looked different from MW808 although both were labeled as Jeneri). Clonal assignment partially agreed with the leaf morphological differences in Berseri specimens, as MW837-848 shared one genotype while MW839 shared the same genotype as Bayas, and MW840 had a unique genotype (Table 20; Appendix 4). As for Jeneri, MW807 and MW808 were assigned the same genotype as MW809 (labeled as Bukit Selambau 1) and MW901 (cultivar name not

recorded) at threshold = 8; and their genetic composition also looked similar in the Structure analysis at k=18 (Figure 20). This suggests that MW809 was mislabeled, and MW807, 808, and 809 likely belong to the same cultivar stock.

Table 20. Cultivar names and associated clonal assignment results. Shared genotypes were assigned a number ("genotype no."). Asterisks (*) indicate that I observed morphological differences between specimens of the same cultivar name. Samples with missing microsatellite data were excluded from the clonal assignment analysis. Samples with unique genotypes refer to samples that did not share genotypes with any other individual.

Cultivar name	**Sites**	**No. of samples with cultivar name**	**No. of samples in clonal assignment test**	**Primary genotype no.**	**Other genotype no.**	**No. of samples with unique genotypes**	**No. of genotypes associated with cultivar name**
Liew Balik Pulau	Liew orchard	18	17	Genotype 6 (n=5)	Genotype 11 (n=3), Genotype 15 (n=2), Genotype 34 (n=2)	4	8
Durian-cempedak	Bidor, Bidor market, UPM Agro-Park	9	7	Genotype 1 (n=7)		0	**1**
Nangka-cempedak (not including four samples that were suspected hybrids based on morphology)	Chekok, UPM Agro-Park, Kluang Germplasm, Ulu Dusun	10	6	None	-	6	6
CH30	Bt. Tangga Germplasm, UPM Agro-Park	6	6	Genotype 7 (n=5)	Genotype 8 (n=1)	0	2
CH29	Bt. Tangga Germplasm	5	5	Genotype 12 (n=3)	Genotype 4 (n=2)	0	2
Pendang	Bt. Tangga Germplasm	5	5	Genotype 10 (n=2)	Genotype 21 (n=1)	2	4
Berseri*	Bt. Tangga Germplasm	4	4	Genotype 10 (n=2)	Genotype 13 (n=1)	1	3
Bayas	Bt. Tangga Germplasm	4	4	Genotype 13 (n=2)	-	2	3
Bukit Manik	Bt. Tangga Germplasm	4	4	Genotype 17 (n=3)	-	1	2
Bukit Selambau 1	Bt. Tangga Germplasm	4	4	Genotype 16 (n=3)	-	1	2
Kuala Nerang	Bt. Tangga Germplasm	4	4	Genotype 3 (n=4)	-	0	**1**
CH27	Ayer Hitam Commodity Dev. Stn.	3	3	Genotype 2 (n=1)	-	2	2
CH35	Ayer Hitam Commodity Dev. Stn., UPM Agro-Park	3	3	Genotype 8 (n=2)	Genotype 5 (n=1)	None	2

Cultivar name	Sites	No. of samples with cultivar name	No. of samples in clonal assignment test	Primary genotype no.	Other genotype no.	No. of samples with unique genotypes	No. of genotypes associated with cultivar name
Bukit Selambau 2	Bt. Tangga Germplasm	3	3	Genotype 3 (n=1)	-	2	3
CH34	Ayer Hitam Commodity Dev. Stn., UPM Agro-Park	2	2	Genotype 5 (n=1)	-	1	2
Bantai	Bt. Tangga Germplasm	2	2	None	-	2	2
Jeneri*	Bt. Tangga Germplasm	2	2	None	-	2	2
Perak Champion	Bt. Tangga Germplasm	2	2	Genotype 14 (n=1)	-	1	2
Sik	Bt. Tangga Germplasm	2	2	Genotype 21 (n=2)	-	0	**1**
CH26 (Paya Jaras)	Kluang Germplasm	1	1	Genotype 31 (n=1)	-	0	1
CH28	UPM Agro-Park	1	1	None	-	1	1
CH33	UPM Agro-Park	1	1	None	-	1	1
Red Ah Chai	Bidor	1	1	Genotype 19 (n=1)	-	0	1
Red Balik Pulau	Bidor	1	1	None	-	1	1
Sri Harum 2	Kluang Germplasm	1	1	None	-	1	1

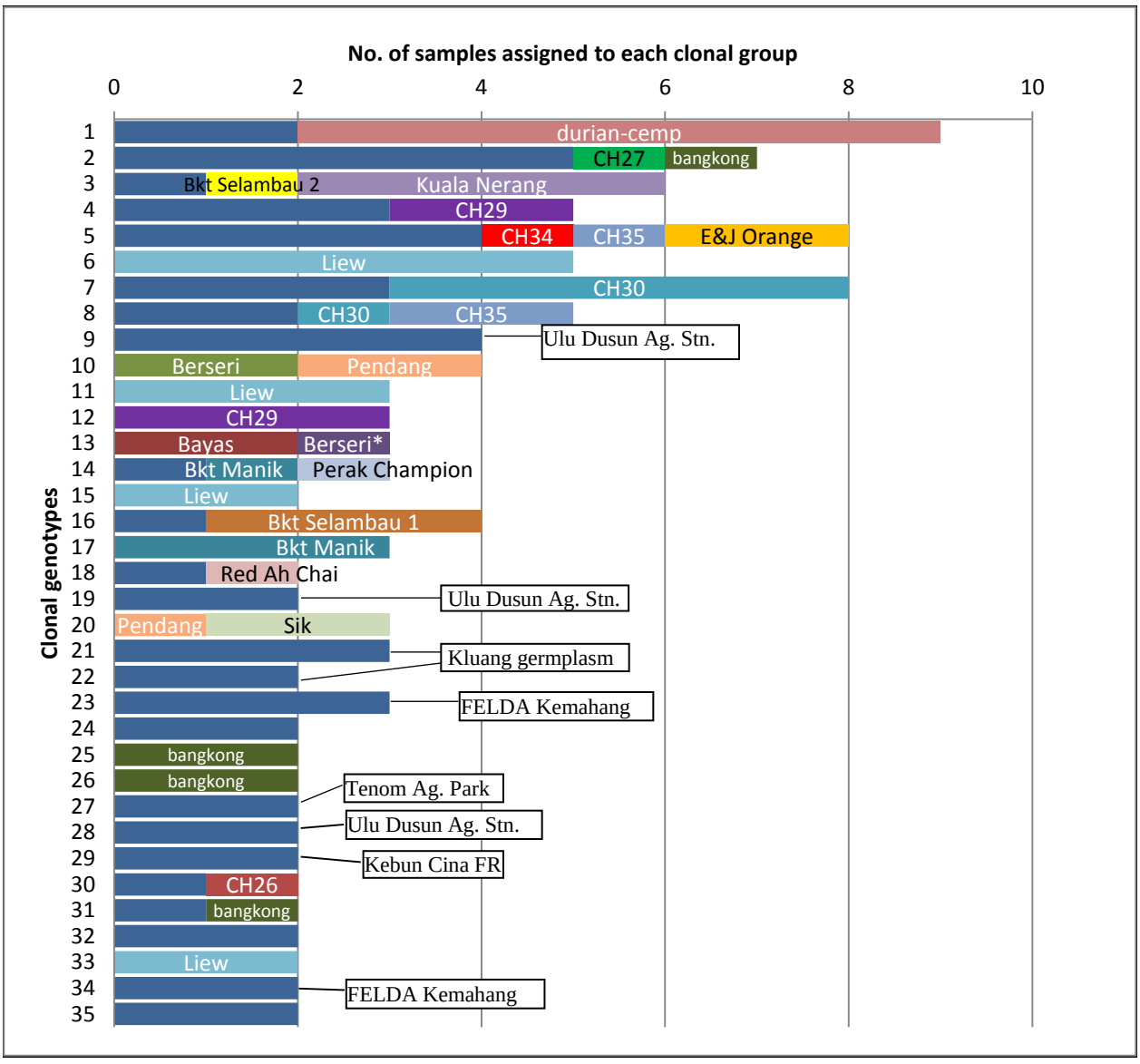


Figure 20. Clonal assignment in Genodive (threshold = 4) detected 121 individuals with shared genotypes, listed here by their "genotype no.". Unlabeled sections of the bar chart represent cempedak samples without recorded cultivar names. Eight clonal groups consisted of only cempedak individuals from the same sampling location, indicated here by site names given in white text boxes.

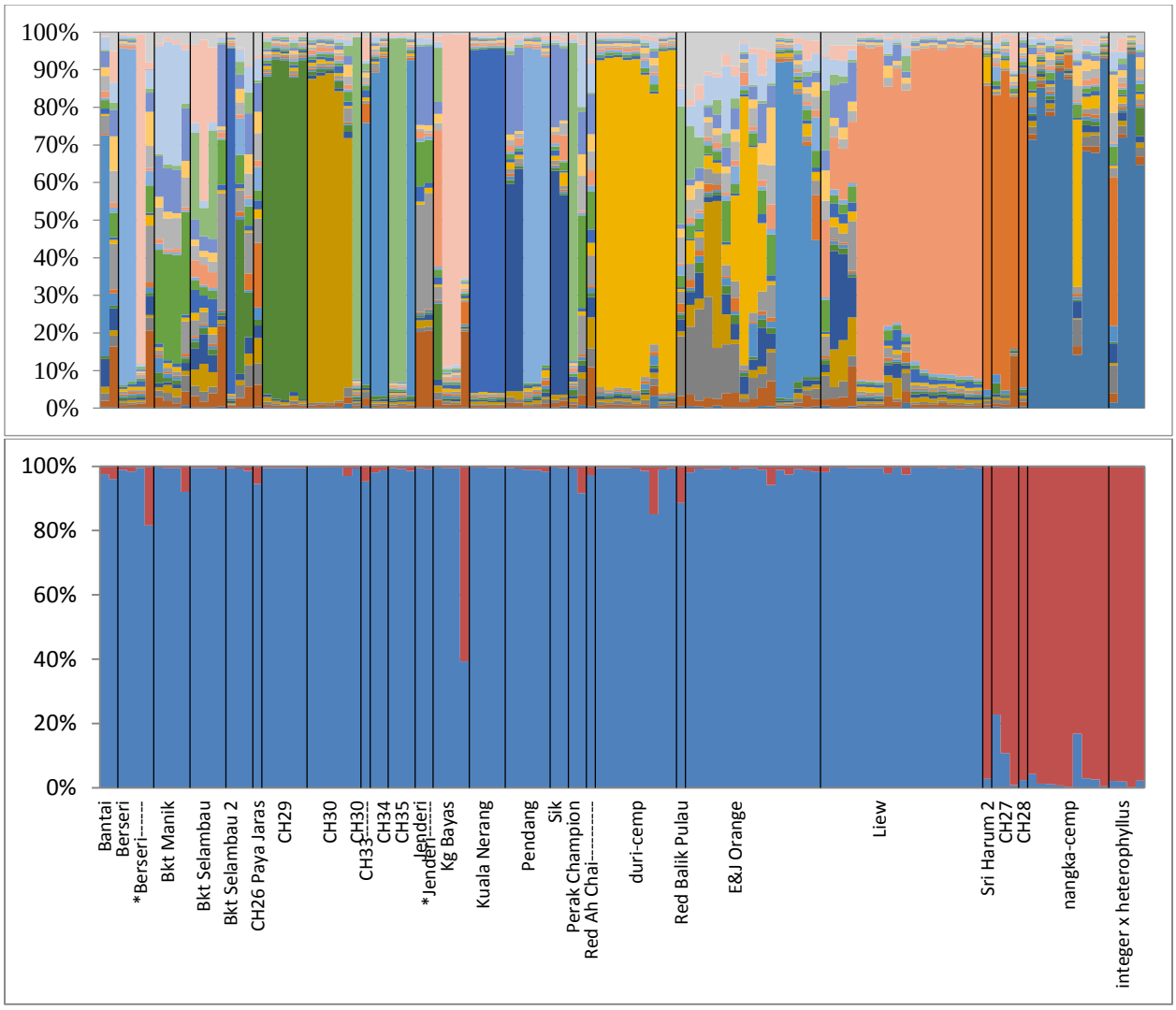


Figure 21. Genetic structure among all samples with recorded cultivar names (top: k=21, bottom: k=2; N =116). Asterisks (*) indicate that I observed morphological differences between specimens of the same cultivar name.

3.4.2 Identification of unknowns and detection of errors

The clonal assignment and Structure analysis of individuals with and without cultivar names enabled identification of some unknowns and detection of possible errors (Table 21 and Table 23). Among the 113 samples sharing the same genotype with at least one other individual, 53 were cempedak samples without recorded cultivar names (Figure 20, Appendix 4). Among the 53 cempedak without recorded cultivar names, 19 individuals shared genotypes with known cultivars (Table 21), and 23 individuals shared genotypes with only individuals from the same sampling location (Figure 20, Table 22). Six bangkong samples shared the same genotype with another cempedak or bangkong samples (Figure 20, Appendix 4). Four bangkong samples were from the same sampling location (Table 22), while two other bangkong samples each shared the same genotype with one or more other cempedak samples (Table 23). Mislabeling, or misidentification of cultivar names associated with trees in germplasms or agricultural collections, seemed to be the most prevalent problem among cempedak samples (Table 23).

Table 21. Cultivar names inferred for unnamed cempedak samples through inference from shared genotypes.

Genotype no.	Coll.	Site	Inferred cultivar name
1	MW351	Sungai Buloh	Durian-cempedak
	NZ918	Tenom Agricultural Park	Durian-cempedak
2	MW353	Sungai Buloh	CH27
	MW354	Sungai Buloh	CH27
	MW403*	Pasir Mas	CH27
	NZ647	Ulu Dusun Agriculture Research Station	CH27
	MW937	Kluang Germplasm	CH27
3	MW903	Kluang Germplasm	Kuala Nerang
4	NZ653	Ulu Dusun Agriculture Research Station	CH29
	NZ654	Ulu Dusun Agriculture Research Station	CH29
	NZ655	Ulu Dusun Agriculture Research Station	CH29
7	MW911	Kluang Germplasm	CH30
	MW905	Kluang Germplasm	CH30
	MW939	Kluang Germplasm	CH30
8	MW917	Kluang Germplasm	CH35
	MW918	Kluang Germplasm	CH35
16	MW909	Kluang Germplasm	Bkt Selambau 1
18	MW421*	Bukit Panau	Red Ah Chai
30	MW922	Kluang Germplasm	CH26

* MW403 was a sapling bought from a local nursery.

** MW421 shared the same genotype as a sample named Red Ah Chai (MW501) from the Bidor orchard. The owners of the two orchards are friends and could have likely shared plant material with one another, or obtained the cultivar material from the same source.

Table 22. Site-specific genotypes. Each genotype listed here is shared by only individuals from the same sampling location.

Genotype no.	Site	Individuals sharing this genotype
9	Ulu Dusun Agriculture Research Station	NZ648, NZ649, NZ656, NZ657
19	Ulu Dusun Agriculture Research Station	NZ643, NZ644
21	Kluang Germplasm	MW906, MW934, MW935
22	Kluang Germplasm	MW914, MW924
23	FELDA Kemahang	MW620, MW622, MW629
27	Tenom Agricultural Park	NZ942, NZ943
28	Ulu Dusun Agriculture Research Station	NZ636, NZ637
29	Kebun Cina FR	NZ689, NZ690
34	FELDA Kemahang	MW624, MW630
25	Puchong (Ayer Hitam) FR	MW1601, MW1605 (both bangkong)
26	Puchong (Ayer Hitam) FR	MW1625, MW1626 (both bangkong)

Table 23. Detection of errors in cempedak/bangkong samples and unexpected results based on clonal assignment test.

Unexpected results refer to cases where samples from very different sites or different states share the same genotype. Some cultivars in the Bt. Tangga Germplasm were duplicated in the Kluang Germplasm. Kluang Germplasm accession numbers (beginning with MART) associated with cultivar names were obtained from documentation provided by Puan Noor Baiti, research officer at the Kluang field station.

Genotype no.	Collection no.	Taxon/cultivar	Site	Details
2	MW1205	bangkong	Lenggor FR	Unexpected/misindentified – a bangkong which shared the same genotype as a CH27 cultivar and six other cempedak. This tree was growing near the entrance to the Lenggor FR in Johor. Structure analysis indicated that this individual was admixed, with 50% probability of membership to both cempedak and bangkong genetic groups (Appendix 4). This tree may be an escaped cultivar.
3	MW817	Bkt Selambau 2	Bt. Tangga Germplasm	Mislabeled – same genotype as Kuala Nerang cultivars.
	MW903	-	Kluang Germplasm	Inferred as Kuala Nerang cultivar, which is accession MART25-C in the Kluang Germplasm
5	MW1402	CH34	UPM Agro-Park	Mislabeled – two different cultivars shared the same genotype. Two other CH35 individuals belong to genotype no. 8. Actual cultivar name could not be determined.
	MW1401	CH35	UPM Agro-Park	
	MW358	-	Sungai Buloh	Planted by villager, unknown cultivar
	MW601	-	Tanah Merah	Homestead, sapling bought from a truck
	NZ941	-	Tenom Ag. Park	Agricultural park, unknown cultivar
	MW742, MW743	E&J Orange	E&J Fruit Farm	Unexpected but possible. Insufficient information to make conclusions.
	MW921	-	Kluang Germplasm	Actual cultivar name could not be determined.
8	MW917, MW918	-	Kluang Germplasm	Inferred unknown - Same genotype as two CH35 individuals, which is accession MART41-C in the Kluang Germplasm.
	MW1407	CH30	UPM Agro-Park (Puchong branch)	Mislabeled – same genotype as two CH35 individuals.
10	MW835, MW836	Pendang	Bt. Tangga Germplasm	Mislabeled – could be Berseri or Pendang.
	MW837, MW838	Berseri	Bt. Tangga Germplasm	Mislabeled – could be Berseri or Pendang.
13	MW839	Berseri*	Bt. Tangga Germplasm	Mislabeled – same genotype as Bayas cultivars.
14	MW822	Bkt Manik	Bt. Tangga Germplasm	Mislabeled – did not share same genotype as other Bukit Manik cultivars, instead sharing the same genotype as Perak Champion (MW820).
	MW945	-	Kluang Germplasm	Inferred as Perak Champion, which is accession MART19-C in the Kluang Germplasm.
20	MW833	Pendang	Bt. Tangga Germplasm	Mislabeled – could be Pendang or Sik.
	MW831, MW830	Sik	Bt. Tangga Germplasm	Mislabeled – could be Pendang or Sik.
24	MW638	-	FELDA Kemahang	Trees from orchards in different states.
	MW708	-	Balik Pulau	Trees from orchards in different states.

Genotype no.	Collection no.	Taxon/cultivar	Site	Details
30	MW922	-	Kluang Germplasm	Based on GPS coordinates, MW922 was far apart from MW953.
	MW953	CH26 Paya Jaras	Kluang Germplasm	This tree was tagged as MART53-C (CH26) in the germplasm.
31	MW1201	bangkong	Gunung Lambak	Unexpected/misidentified – bangkong shared same genotype as MW367, a cultivated cempedak from Sungai Buloh. In Structure analysis, MW1201 clustered with the KC/cempedak group. This tree was growing beside a road to a campground in Gunung Lambak FR, and was initially identified as silvestris based on its location in a forest reserve, but based on molecular data, it may be an escaped cempedak or planted by campground management.
32	MW608	-	FELDA Kemahang	Trees from orchards in different states.
	MW714	-	Balik Pulau	Trees from orchards in different states.
35	MW420	-	BktPanau	Unexpected, because Bukit Panau is an orchard in Kelantan, while Sianggau FR is a disturbed forest in Sabah.
	NZ863	-	Sianggau FR	
21	MW906, MW934, MW935	-	Kluang Germplasm	All three trees had different GPS coordinates. Based on GPS coordinates, MW934 and 935 are likely to be from the same accession, but not MW906. Based on GPS coordinates, MW934 and 935 may each have been sampled twice because they had the same GPS coordinates as MW947 and 948. However, MW947 and 948 did not share the same genotype as MW934 and 935.
22	MW914, MW924	-	Kluang Germplasm	The two trees had different GPS coordinates, but are too far apart to be from the same accession.

4. Discussion

4.1 Genetic diversity of cempedak and bangkong

Although all measures of genetic diversity except for allelic richness (Ar) did not differ significantly between the total cultivated cempedak and total bangkong populations, comparison of individual cempedak and bangkong populations showed that four measures of genetic diversity (Ne, Ar, I, and He) were significantly higher in bangkong than cempedak (Table 10, Table 12). Considering that the sample size of bangkong was much smaller than cultivated cempedak, the statistically significant results suggest that wild bangkong populations are indeed more genetically diverse than cultivated cempedak populations. However, the diversity in cultivated cempedak was not a mere subset of bangkong diversity, as cultivated cempedak populations contained 31 private alleles not found in bangkong populations.

The total Borneo cempedak population had significantly lower diversity (Na, Ne, Pa, and I) than the total Peninsular cempedak population, but not significantly higher inbreeding (Table 13). Again, the diversity in Peninsular cempedak was not subset of Borneo cempedak diversity, as Peninsular cempedak harbored 43 alleles not found in Borneo populations (Table 13).

Consequently, these data do not point directly to the bangkong or Borneo cempedak populations in this study as being representative of ancestral populations from which cempedak was domesticated, nor does it point to a genetic bottleneck in cultivated cempedak due to founder effects. At the same time, bangkong and Borneo cempedak populations in this study were limited in sample size or geographic distribution (all bangkong samples were from Peninsular Malaysia, and most of them from secondary forests), which may have limited the amount of genetic diversity observed in this study. Moreover, habitat loss may have reduced the diversity levels of current bangkong populations relative to historical populations. Lastly, it is not unusual for new diversity to arise due to mutation followed by selection in a cultivated species like cempedak.

For these reasons, we cannot rule out the possibility that cempedak was domesticated from bangkong or from Borneo.

A review by Miller and Gross (2011) found that domestication bottlenecks may be less severe in perennial crops compared to annual crops. Miller and Gross (2011) suggested that the smaller impact of domestication bottleneck on perennial crops was due to their long juvenile phase and clonal propagation, which reduce the number of reproductive cycles that selection can act on; additionally, genetic variation is maintained as perennial crops tend to outcross and hybridize easily with other species. In five microsatellite studies comparing genetic diversity in perennial crops and their wild relatives, they found that 94.1-111.4% of the expected heterozygosity (He) from the wild species was retained in the domesticated species ((Miller and Gross, 2011). In this study, 86.7% of the expected heterozygosity (He) of bangkong was found in cultivated cempedak (He of total bangkong population = 0.525; He of total cultivated cempedak = 0.455), which is lower than the range reported by the five studies reviewed by Miller and Gross (2011). Interestingly, those studies also reported higher He values compared to cempedak and bangkong (range of He of wild species: 0.68-0.721; range of He of domesticated species: 0.64-0.814). This may be due to ascertainment bias because the microsatellite primers used in this study were developed from breadfruit (*A. altilis*) (Hutter et al., 1998). These primers may capture a lower amount of genetic variation in congeneric species such as cempedak (see Table 24 for comparison of He values). However, we should still be cautious when comparing diversity measures across different species, because those measures can be affected by many other factors such as breadth of sampling locations, species differences, marker differences and so on.

Table 24. Comparison of expected heterozygosity (He) among four *Artocarpus* taxa across 12 microsatellite loci. Values for breadfruit and jackfruit are from (Witherup et al., 2013).

Locus	Breadfruit (N=79)	Jackfruit (N=426)	Cempedak (N=347)	Bangkong (N=187)
26	ND	0.776	0.354	0.48
54a	0.783	0.454	0.429	0.15
54b	0.313	0.729	0.021	0.022
105	ND	0.616	0.411	0.611
122	0.698	0.588	0.719	0.761
140	0.695	0.747	0.696	0.787
145	0.659	0.662	0.245	0.34
156	0.533	0.504	0.568	0.613
178a	0.793	0.185	0.612	0.553
178b	0.729	0.574	0.562	0.87
182	0.536	0.609	0.182	0.594
196a	ND	0.652	0.602	0.521
Mean	***0.638***	***0.591***	***0.450***	***0.525***
SE	0.042	0.044	0.060	0.069

4.2 Geographic Patterns in Genetic Structure & Differentiation

Structure analyses of all samples sorted individuals into cempedak ("KC") or bangkong ("KB") groups, and also detected individuals that may have been misidentified or that represent unique genetic composition. In particular, individuals from Gunung Pulai initially identified as bangkong may have been misidentified, as other analyses have found that the Gunung Pulai population deviates from other bangkong populations (PCoA in Figure 13; phylogenetic trees in Figures 15-17). This population may consist of cempedak that has escaped cultivation, and may even be a wild progenitor of cultivated cempedak in Peninsular Malaysia. The Gunung Pulai population also has the highest level of inbreeding among the bangkong populations, even the ones with smaller sample sizes (fixation index = 0.303; Table 11). However, other analyses such as sequencing or a systematic study of morphological characters would be required to make any conclusions about this population.

Both cempedak and bangkong exhibited weak geographic structure within Peninsular Malaysia, with gene flow occurring among populations from different states (Figures 7-10). However, Borneo cempedak formed a distinct genetic cluster, suggesting that cempedak may exhibit geographic structure at a broader, regional level (Figure 7 and Figure 8). The South China Sea is a geographical barrier to gene flow between the Peninsula and Borneo, which may reduce the trade or movement of cempedak propagation material between the two regions, but not completely prevent it. For example, Tenom Agricultural Park and Ulu Dusun Agriculture Research Station in Sabah contains several individuals that were assigned to the Peninsular group, likely a result of cultivar transfer from the Peninsula (Figure 7). On the other hand, the Kluang germplasm in Johor contains individuals assigned to the Borneo group (Figure 7). Indeed, germplasm records show that there were two cultivars from Sarawak and one cultivar from West Kalimantan in the Kluang collection. In addition, the Kluang germplasm – the biggest

cempedak germplasm collection in Malaysia - also has individuals in the three Peninsular subgroups (cluster A-C in Figure 7), indicating that this germplasm represented all the major gene pools in Malaysia found in this study. Nevertheless, additional gene pools may be found in other Malaysian states that were not covered in this study (i.e. Perlis, Terengganu, Negeri Sembilan, Melaka, and Sarawak). Germplasm collections in Peninsular Malaysia should also aim to include more cultivars or landraces from Sabah and Sarawak to increase the genetic diversity conserved in those *ex situ* facilities. The three Peninsular subgroups may represent groupings of cultivars (Figure 21), but do not correspond to the geographical locations of the source of cultivar propagation material (Table 25). This lack of geographical pattern in the Peninsular subgroups suggests that cempedak growers are actively exchanging propagation material across borders, and perhaps crossing trees of different provenances to create new varieties.

In bangkong, geographical isolation between bangkong populations was perhaps too recent for genetic divergence, or gene flow is still occurring between the populations. In Selangor, the FRIM-Kepong, Sungai Buloh, and Kota Damansara populations are relatively close to each other, while the MARDI Forest Reserve is relatively close to the Puchong (Ayer Hitam) Forest Reserve; however, these bangkong populations did not cluster according to their geographical proximity (Figure 10). Interestingly, the Puchong (Ayer Hitam) population stood out as a distinct genetic cluster (Figure 9 and Figure 10). The Ayer Hitam Forest Reserve in Puchong, Selangor, is a remnant forest about four square miles in size, surrounded by urban and residential areas. This population may have been isolated from other bangkong populations in Selangor for a longer period; however, the lack of information prevents us from making any conclusions.

AMOVA results showed that at both population and regional levels, most of the genetic variation was accounted for within rather than among cultivation/taxon categories (Table 16 and Table 17). This suggests that gene flow is occurring among wild and cultivated populations across regions, or that bangkong and cempedak have been separated only recently and not enough time has gone by to see the effects of isolation. However, pairwise-comparison of Fst values and PCoA reveal observable separation between bangkong and cempedak (Table 18, Figures 13-14). Wild Borneo cempedak - specifically the Beaufort Hill population in Sabah - may be differentiated from cultivated cempedak populations (Figure 13), but PCoA of individual-pairwise genetic distances did not show this pattern (Figure 14), suggesting that there is gene flow between wild and cultivated cempedak in Borneo. Gene flow between cultivated populations and their wild relatives is a common occurrence, especially where cultivated and wild species overlap in range, which is the case for cempedak and bangkong (Ellstrand et al., 1999). Studies using neutral genetic markers have found that natural populations of long-lived species exhibit high levels of variation within populations and weak population structure, which was what we found for bangkong in this study (Miller and Gross, 2011). However, Mantel tests of isolation-by-distance showed that genetic differentiation among bangkong populations was significantly correlated with geographic distances, but not for cempedak populations (Table 19). This suggests that geographic distance may play a stronger role in shaping the genetic structure within bangkong than was detected by the Structure analysis.

Table 25. Cempedak cultivars with known geographic sources and corresponding structure clusters. Source of cultivar was obtained from germplasm records or interviews, and is color-coded by state (lighter to darker shade from north to south). Structure clusters are also color-coded according to the colors in Figure 7 (cluster A = blue, Cluster B = red, Cluster C = green).

Cultivar names	N	Sampling location	Source of cultivar	Structure cluster
Bukit Manik	3	Bt. Tangga Germplasm, Kedah	Perlis	A
Bayas	4	Bt. Tangga Germplasm, Kedah	Langkawi (belongs to state of Kedah but is geographically closer to Perlis)	A
duri-cemp	6	Bidor, Perak	Bidor, Perak	A
Red Balik Pulau	1	Bidor, Perak	Balik Pulau, Penang	B
Perak Champion	1	Bt. Tangga Germplasm, Kedah	Temoh, Perak	B
CH35	2	DoA Ayer Hitam, Johor	Sri Temoh, Perak	B
CH29	5	Bt. Tangga Germplasm, Kedah	Kuang, Selangor	B
Bkt Selambau	4	Bt. Tangga Germplasm, Kedah	Sg. Petani, Kedah	B/C
Bkt Selambau 2	2	Bt. Tangga Germplasm, Kedah	Sg. Petani, Kedah	B/C
Kuala Nerang	4	Bt. Tangga Germplasm, Kedah	Kuala Nerang, Kedah	C
Pendang	2	Bt. Tangga Germplasm, Kedah	Pendang, Kedah	C
Jeneri	1	Bt. Tangga Germplasm, Kedah	Sik, Kedah	C
Jeneri*	1	Bt. Tangga Germplasm, Kedah	Sik, Kedah	C
Sik	2	Bt. Tangga Germplasm, Kedah	Sik, Kedah	C
Liew	14	Balik Pulau, Penang	Balik Pulau, Penang	C
Red Ah Chai	1	Bidor, Perak	Temoh, Perak	C
CH30	5	Bt. Tangga Germplasm, Kedah	Kuang, Selangor	C

.

4.3 Phylogenetic Relationships

The neighbor-joining and UPGMA trees showed that at both population and individual levels, bangkong formed a separate cluster from cempedak, with some degree of intermixing between the two taxa (Figures 15-19). In the population-level neighbor-joining analyses (Figure 15 and Figure 16), one bangkong population (Ulu Serendah) is basal to all the cempedak populations, suggesting that this population may be a progenitor population for cempedak, which would be in line with Corner (1938)'s hypothesis that cempedak was derived from bangkong. Nonetheless, neighbor-joining and UPGMA analyses are limited in their ability to infer ancestry of species, so we should be careful when making such interpretations based on those types of analyses.

In contrast, phylogenetic trees based on genetic sequences have greater power to answer questions of evolutionary origins. Maximum-likelihood trees (RAxML with the GTRGAMMA model) was constructed for cempedak and bangkong using sequences of the ITS and trnH-psbA regions (Figure 22 and Figure 23, courtesy of E. Gardner, unpublished). The ITS tree shows that cempedak (*Artocarpus integer*) and bangkong (*Artocarpus integer* var. *silvestris)* together formed a well-supported clade that was sister to jackfruit (*A. heterophyllus)* and a cempedak-jackfruit hybrid (NZ712), but the lack of resolution within the cempedak+bangkong clade make it unclear whether cempedak or bangkong has a more basal position (Figure 22). This may be partly because the gene regions sampled were limited and did not capture the divergence of cempedak and bangkong, and partly due to the limited number of individuals sequenced and/or limited geographical distribution of individuals.

The trnh-psbA tree, however, shows that bangkong (in dark green) and several outgroups (blue) are sister to a well-supported clade consisting of all cempedak and KC-bangkong (i.e. bangkong which had been assigned to the cempedak group in Structure analyses) (Figure 23).

Within the cempedak/KC-bangkong clade, Sabah cempedak individuals (in red) are sister to a clade containing all Peninsular and Thai individuals and only one Borneo cempedak (Figure 23). This suggests that the break between cempedak and bangkong may have occurred prior to the domestication of cempedak, and that Sabah cempedak may represent the progenitor population to Peninsular cempedak, while bangkong evolved separately. This interpretation agrees with Primack's hypothesis that cempedak evolved outside Peninsular Malaysia and was brought into the Peninsula, but conflicts with his hypothesis that bangkong was derived from cempedak in Peninsular Malaysia.

It is evident from this discussion that different phylogenetic methods and different molecular markers/gene regions each tell a different story, suggesting that cempedak has multiple origins or multiple domestication events, which is common for domesticated fruit trees (Miller and Gross, 2011). Future studies investigating the origins of cempedak should incorporate sampling of multiple gene regions, a larger sample size, as well as broader geographical sampling of cempedak and bangkong throughout their ranges, especially from Borneo, Sumatra, the Riau Islands (located between Peninsular Malaysia, Sumatra, and Borneo), and the islands to the east of Borneo (Sulawesi to Papua New Guinea).

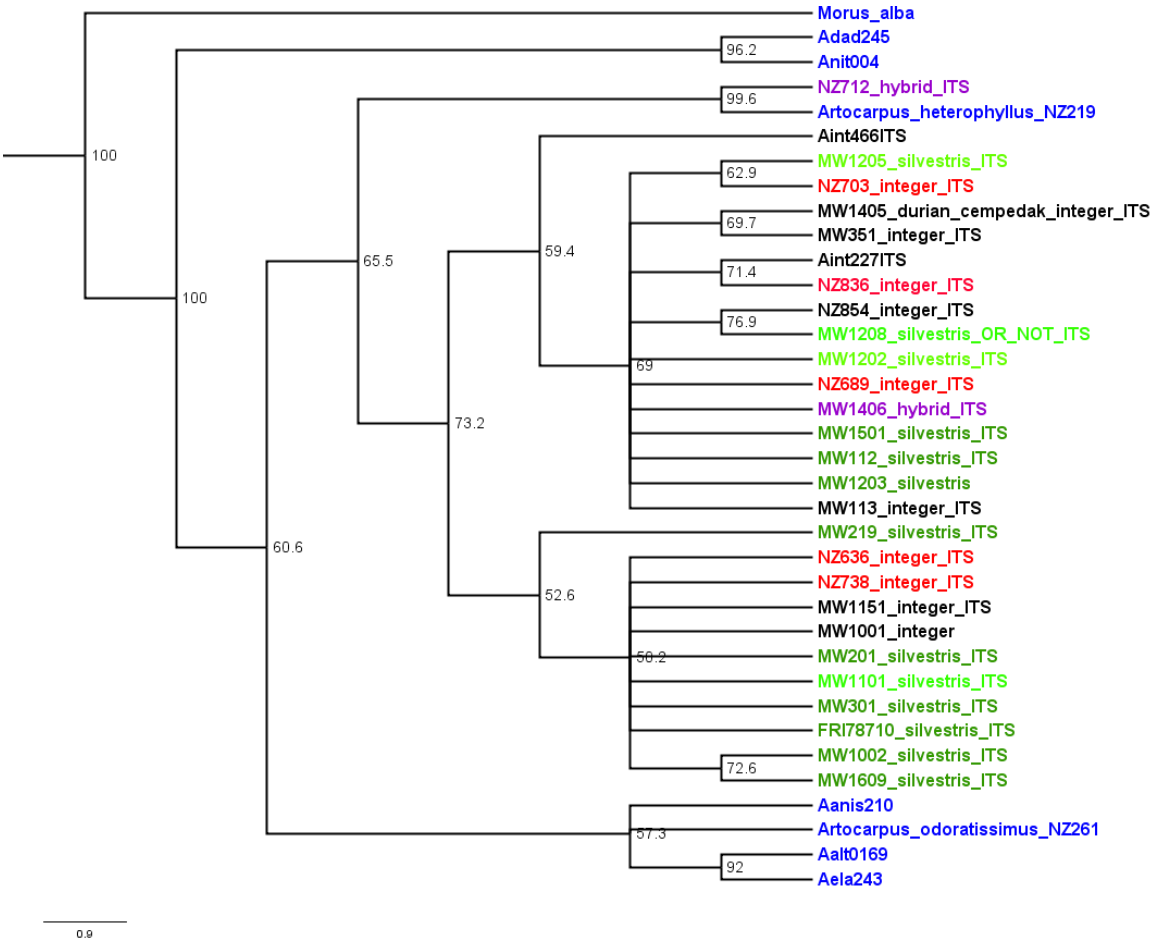


Figure 22. Majority-rule (50%) consensus tree (RAxML with the GTRGAMMA model) generated from ITS region sequences, with bootstrap values generated from 1000 replicates (Gardner, unpublished). Color codes: blue (outgroup), green (bangkong); bright green ("KC-bangkong", i.e. bangkong that was assigned to cempedak group in Structure); red (cempedak from Sabah); purple (cempedak-jackfruit hybrid (NZ712) or nangka-cempedak (MW1406). Congeneric outgroups shown on the tree are *A. heterophyllus*, *A. dadah (Adad)*, *A. nitidus (Anit)*, *A. odoratissimus*, *A. anisophyllus (Aanis)*, *A. altilis (Aalt)*, *A. elasticus (Aela)*.

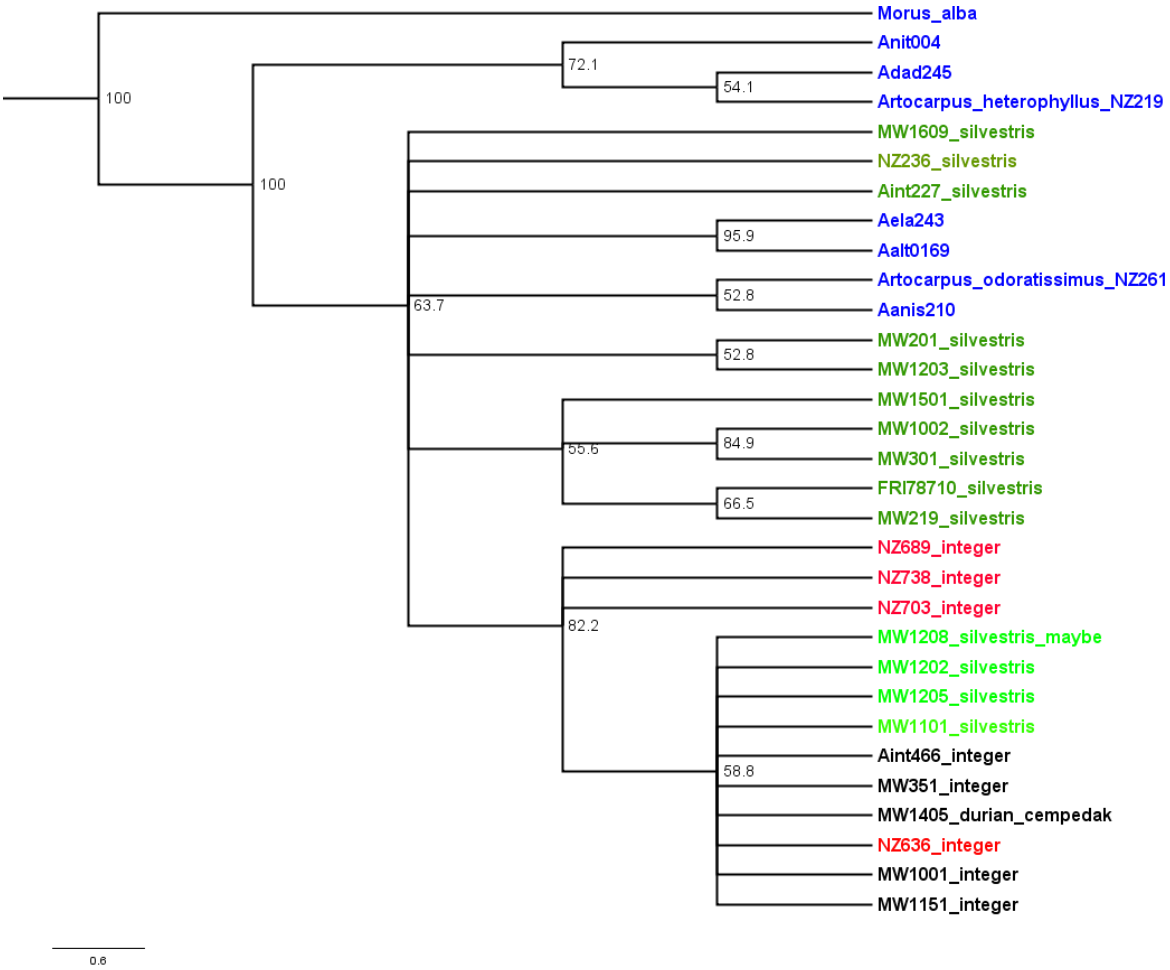


Figure 23. Majority-rule (50%) consensus tree (RAxML with the GTRGAMMA model) generated from trnH-psbA region sequences, with bootstrap values generated from 1000 replicates (Gardner, unpublished). Branches are presented as cladograms for better visualization of the tree. Cempedak are indicated as "integer" (*Artocarpus integer)* and bangkong as "silvestris" (*Artocarpus integer* var. *silvestris).* Color codes: blue (outgroup), green (bangkong); bright green ("KC-bangkong", i.e. bangkong that was assigned to cempedak group in Structure); red (cempedak from Sabah). Congeneric outgroups shown on the tree are *A. heterophyllus, A. dadah (Adad), A. nitidus (Anit), A. odoratissimus, A. anisophyllus (Aanis), A. altilis (Aalt), A. elasticus (Aela).*

4.4 Cultivar Genotypes

4.4.1 Cultivar Names, Clonal Analysis and Genetic Structure

The discrepancies in genotype-heritability within cultivar names are likely related to the propagation method, which may be a combination of seed propagation and grafting. In particular, the lack of shared genotypes among samples labeled as nangka-cempedak may be indicative of multiple hybridization events between cempedak and jackfruit. Individuals that share the same multilocus genotype may not necessarily be clonally reproduced, as identical genotypes can be due to insufficient discriminative power of the molecular markers (Halkett et al., 2005). This may explain why the clonal assignment test detected bangkong individuals that shared the same genotypes in Ayer Hitam FR, or bangkong individuals that shared the same genotype as cempedak individuals (Table 22 and Table 23). Nevertheless, it is also possible that scoring and PCR errors or the presence of null alleles in some loci may have led to unexpected results (Table 9). We do not know the likelihood of scoring errors in this study, but we recommend that future studies using the same microsatellite markers should calculate the error rate by re-genotyping and re-scoring a random subset of samples (Selkoe and Toonen, 2006).

4.4.2 Identification of unknowns and detection of errors

Clonal assignment and Structure analysis can be useful for identifying the cultivar names of individuals lacking provenance information, and also for detecting labeling errors in germplasm and agricultural collections (Table 21 and Table 23). Other sources of errors may be genotyping or PCR errors, presence of null alleles, as discussed in the preceeding section. Three trees in the UPM Agro-Park teaching collection may have been mislabeled, where two trees respectively recorded as CH34 and CH35 were assigned the same genotype; and another tree recorded as CH30 shared the same genotypes as two other CH35 individuals from Ayer Hitam Commodity Development Station (CDS) (Table 23). The trees I sampled at the Ayer Hitam CDS

were unlikely to have been mislabeled because they had been specifically identified and marked with cultivar names by the station officers (Miss Lem, Ayer Hitam CDS, pers. comm.). At least ten trees in the Bt. Tangga germplasm may have been mislabeled (Table 23). The labeling errors in the Bt. Tangga germplasm was largely caused by trees not having tags or identifiers – identification of cultivar names depended on a single staff member, who learned the cultivar names by working with an ex-staff member who was involved in the collecting trips. At the Kluang germplasm, we had documentation for each cultivar accession in the germplasm; however, the majority of the trees had lost their identifier tags and so we could not match the provenance or cultivar information to the individual trees sampled. Only three samples could be confidently matched wtih their accession number and cultivar name (MW912 = MART67-C (Nangka-cempedak); MW953 = MART53-C (CH26); MW965 = MART65-C (Sri Harum 2)). We also had a schematic map showing the layout of the germplasm collection; however, the large size of the germplasm collection and the death of many trees complicated the navigation process. GPS coordinates were recorded for each tree sampled; however, the GPS coordinates may have been inaccurate because they did not match the schematic map or patterns found in molecular data (see 'Details' in Table 23).

Molecular characterization of germplasm collections have limited utility without reliable identifiers (i.e. tree tags) linking molecular data to individual accessions. Hence, for future studies of germplasm collections lacking identifiers, we recommend that researchers attach new tags to the trees sampled, so that molecular data can be linked to individual trees in the field.

5. Conclusion

Cempedak is a locally important fruit crop in Malaysia but has garnered little attention for research. Its wild relative and putative wild progenitor, bangkong, is a potential source of valuable genetic resources for improvement of the cempedak crop. This is the first study characterizing the genetic diversity of cempedak and bangkong across Peninsular and East Malaysia. We found that cempedak and bangkong harbored distinct pools of genetic diversity, and the amount of genetic diversity in cultivated cempedak is comparable and unique compared to the diversity in bangkong, although results strongly suggest that bangkong is more genetically diverse. We also found that Peninsular cempedak harbored unique and greater allelic diversity compared to Borneo cempedak, but this result was likely biased by the disparity in sampling sizes between the two regions. Both cempedak and bangkong showed weak geographic structure in Peninsular Malaysia, but Borneo cempedak comprises a gene pool distinct from Peninsular cempedak. Genetic differentiation and geographic distances between bangkong populations were also correlated. Although we were unable to obtain a conclusive answer to whether bangkong was the wild progenitor of cempedak, we found evidence to support both Corner's and Primack's hypotheses about the origin of cempedak. Microsatellite-based neighbor-joining trees indicate that the bangkong population in Ulu Serendah may be a progenitor population, in contrast with the maximum likelihood tree based on trnH-psbA sequences, which suggest that bangkong and cempedak evolved separately and that Borneo cempedak may be the progenitor to Peninsular cempedak. Lastly, we found that samples with the same cultivar names were not always true-to-type, which may be due to multiple reasons such as mislabeling, hybridization, and scoring errors.

This study provides a baseline of genetic diversity of cempedak and bangkong across Peninsular Malaysia and for cempedak in Sabah, which will be useful for future studies comparing changes in diversity over time. We also provide molecular characterization of two cempedak germplasm collections in Kluang and Bukit Tangga, and found that they represented adequate coverage of the major gene pools detected in this study. However, efforts should be made to ensure that trees in the germplasm collections can be readily and reliably identified and linked to accession data. Furthermore, *ex situ* conservation of bangkong and wild cempedak from Borneo should be given attention because these wild populations may hold key genes for crop improvement, and yet are the most threatened by forest fragmentation.

We hope that the information generated from this study will be useful for future research and continued conservation of genetic resources of tropical fruit crops and their wild relatives in Malaysia and beyond. Future work that incorporates morphological data, sequencing data, and additional geographic sampling of bangkong as well as cultivated and wild cempedak across their ranges will likely yield more interesting and valuable insights into the origin and domestication history of cempedak.

References

ABDUL RAHMAN, M., S. ANUAR KARSANI, I. OTHMAN, P. SHAFINAZ ABDUL RAHMAN, AND O. HAJI HASHIM. 2002. Galactose-binding lectin from the seeds of champedak (*Artocarpus integer*): sequences of its subunits and interactions with human serum O-glycosylated glycoproteins. *Biochemical and Biophysical Research Communications* 295: 1007–1013.

ALTIERI, M., AND L. MERRICK. 1987. In situ conservation of crop genetic resources through maintenance of traditional farming systems. *Economic Botany* 41: 86–96.

AZIZ, N.A.A., AND M.A. ZABIDI. 2011. Partial Substitution of Wheat Flour with Chempedak (Artocarpus integer) Seed Flour in Bread. *In* V. R. Preedy, R. R. Watson, and V. B. Patel [eds.], Flour and Breads and Their Fortification in Health and Disease Prevention, 524. Academic Press, San Diego, CA.

BHAG MAL, V. RAO, AND R. ARORA. 2011. Conservation and Sustainable Use of Tropical Fruit Species Diversity: Bioversity's Efforts in Asia, the Pacific and Oceania. *Indian Journal of Plant Genetic Resources* 24: 1–22.

BOONLAKSIRI, C., W. OONANANT, P. KONGSAEREE, P. KITTAKOOP, M. TANTICHAROEN, AND Y. THEBTARANONTH. 2000. An antimalarial stilbene from *Artocarpus integer*. *Phytochemistry* 54: 415–417.

CENTRAL INTELLIGENCE AGENCY. Malaysia. In The World Factbook. Available at: https://www.cia.gov/library/publications/the-world-factbook/geos/br.html [Accessed November 12, 2014].

CHAPMAN, A.D. 2009. Numbers of Living Species in Australia and the World. Toowoomba, Australia.

CHAPUIS, M.P., AND A. ESTOUP. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24: 621–631.

CHAUDHURI, S. 2005. Genetic erosion of agrobiodiversity in India and intellectual property rights: interplay and some key issues. *Patentmatics* 5: 1–10.

CONVENTION ON BIOLOGICAL DIVERSITY (CBD). 2014. Malaysia overview page. Available at: http://www.cbd.int/countries/?country=my [Accessed October 30, 2014].

CORNER, J.H. 1938. The Jack and the Chempedak. *In* Notes on the Systematy and Distribution of Malayan Phanerograms, II, 56–81.

DEPARTMENT OF AGRICULTURE OF PENINSULAR MALAYSIA. 2001. Technology Package: Cempedak (in Malay). 1st ed. Department of Agriculture of Peninsular Malaysia.

DOEBLEY, J., B. GAUT, AND B. SMITH. 2006. The molecular genetics of crop domestication. *Cell* 127: 1309–1321.

DOYLE, J., AND J.L. DOYLE. 1987. Genomic plant DNA preparation from fresh tissue-CTAB method. *Phytochemical Bulletin* 19: 11–15.

DYER, G.A., A. LOPEZ-FELDMAN, A. YUNEZ-NAUDE, AND J.E. TAYLOR. 2014. Genetic erosion in maize's center of origin. *Proceedings of the National Academy of Sciences* 111: 14094–14099.

EARL, D.A., AND B.M. VONHOLDT. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.

ELLSTRAND, N.C., H.C. PRENTICE, AND J.F. HANCOCK. 1999. Gene flow and introgression from domesticatd plants into their wild relatives. *Annual Review of Ecology and Systematics* 30: 539–563.

EVANNO, G., S. REGNAUT, AND J. GOUDET. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611–2620.

FAO. 2012. Energy-Smart Food at FAO: An Overview. Rome, Italy.

FAO. 1991. Food balance sheets: 1984-1986 average. Rome, Italy.

FAO. 2010. Second Report on the State of the World's Plant Genetic Resources for Food and Agriculture. Rome.

FELSENSTEIN, J. 2005. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

FORESTRY DEPARTMENT OF PENINSULAR MALAYSIA. 2013. Annual Report.

GEPTS, P. 2006. Plant genetic resources conservation and utilization: the accomplishments and future of a societal insurance policy. *Crop Science* 46: 2278–2292.

HALKETT, F., J.-C. SIMON, AND F. BALLOUX. 2005. Tackling the population genetics of clonal and partially clonal organisms. *Trends in ecology & evolution* 20: 194–201.

HAMMER, K., AND G. LAGHETTI. 2005. Genetic Erosion – Examples from Italy. *Genetic Resources and Crop Evolution* 52: 629–634.

HANSEN, M.C., P. V POTAPOV, R. MOORE, M. HANCHER, S.A. TURUBANOVA, A. TYUKAVINA, D. THAU, ET AL. 2013. High-resolution global maps of 21st-century forest cover change. *Science* 342: 850–3.

HANSEN, M.C., S. V STEHMAN, P. V POTAPOV, T.R. LOVELAND, J.R.G. TOWNSHEND, R.S. DEFRIES, K.W. PITTMAN, ET AL. 2008. Humid tropical forest clearing from 2000 to 2005 quantified by using multitemporal and multiresolution remotely sensed data. *Proceedings of the National Academy of Sciences of the United States of America* 105: 9439–44.

HARLAN, J.R. 1976. Genetic Resources in Wild Relatives of Crops. *Crop Science* 16: 329–333.

HEAL, G., B. WALKER, S. LEVIN, K. ARROW, P. DASGUPTA, G. DAILY, P. EHRLICH, ET AL. 2004. Genetic diversity and interdependent crop choices in agriculture. *Resource and Energy Economics* 26: 175–184.

HOYT, E. 1988. Conserving the Wild Relatives of Crops. Rome, Italy and Gland, Switzerland.

HUTTER, C.M., M.D. SCHUG, AND C.F. AQUADRO. 1998. Microsatellite variation in Drosophila melanogaster and Drosophila simulans: a reciprocal test of the ascertainment bias hypothesis. *Molecular Biology and Evolution* 15: 1620–1636.

HYTEN, D.L., Q. SONG, Y. ZHU, I.-Y. CHOI, R.L. NELSON, J.M. COSTA, J.E. SPECHT, ET AL. 2006. Impacts of genetic bottlenecks on soybean genome diversity. *Proceedings of the National Academy of Sciences of the United States of America* 103: 16666–71.

JANICK, J., AND R. PAULL. 2008. *Artocarpus heterophyllus / Artocarpus integer*. *In* The encyclopedia of fruit and nuts, 481–487. CABI Publishing Series.

JANSEN, P.C.M. 1991. *Artocarpus integer* (Thunb.) Merr. *In* E. W. M. Verheij, and R. E. Coronel [eds.], PROSEA (Plant Resources of South-East Asia). No. 2. Edible Fruits and Nuts., 91–94. Pudoc, Wageningen, Netherlands.

JARRETT, F.M. 1959. Studies in Artocarpus and Allied Genera I-III. *Journal of the Arnold Arboretum* XL: 1-29, 30-37, 298-368.

JARRETT, F.M. 1960. Studies in Artocarpus and Allied Genera IV-V. *Journal of the Arnold Arboretum* XLI: 73–139.

JONES, J.D.G., K. WITEK, W. VERWEIJ, F. JUPE, D. COOKE, S. DORLING, L. TOMLINSON, ET AL. 2014. Elevating crop disease resistance with cloned genes. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 369: 20130087.

KHOURY, C.K., A.D. BJORKMAN, H. DEMPEWOLF, J. RAMIREZ-VILLEGAS, L. GUARINO, A. JARVIS, L.H. RIESEBERG, AND P.C. STRUIK. 2014. Increasing homogeneity in global food supplies and the implications for food security. *Proceedings of the National Academy of Sciences of the United States of America* 111: 4001–6.

LANE, A., AND A. JARVIS. 2007. Changes in climate will modify the geography of crop suitability: agricultural biodiversity can help with adaptation. *ICRISAT* 4: 1–12.

LIM, S. BIN, C.T. CHUA, AND O.H. HASHIM. 1997. Galactose-binding lectin from the seeds of champedak (*Artocarpus integer*): sequences of its subunits and interactions with human serum O-glycosylated glycoproteins. *Journal of Immunological Methods* 209: 177–186.

MEILLEUR, B.A., AND T. HODGKIN. 2004. In situ conservation of crop wild relatives: status and trends. *Biodiversity and Conservation* 13: 663–684.

MEIRMANS, P.G., AND P.H. VAN TIENDEREN. 2004. Genotype and Genodive: Two Programs for the Analysis of Genetic Diversity of Asexual Organisms. *Molecular Ecology Notes* 4: 792–794.

MILLER, A.J., AND B.L. GROSS. 2011. From forest to field: perennial fruit crop domestication. *American Journal of Botany* 98: 1389–414.

MINISTRY OF AGRICULTURE AND AGRO-BASED INDUSTRY OF MALAYSIA. 2014. Agro-Food Statistics 2013. Putrajaya, Malaysia.

MINISTRY OF NATURAL RESOURCES AND ENVIRONMENT OF MALAYSIA. 2014. 5th National Report to Convention of Biological Diversity (Malaysia). Putrajaya, Malaysia.

MINN, Y., K. PRINZ, AND R. FINKELDEY. 2014. Genetic variation of teak (Tectona grandis Linn. f.) in Myanmar revealed by microsatellites. *Tree Genetics & Genomes* 10: 1435–1449.

MOORE, G., AND W. TYMOWSKI. 2005. Explanatory Guide to the International Treaty on Plant Genetic Resources. IUCN, Gland, Switzerland and Cambridge, UK.

MORA, C., D.P. TITTENSOR, S. ADL, A.G.B. SIMPSON, AND B. WORM. 2011. How many species are there on Earth and in the ocean? G. M. Mace [ed.],. *PLoS Biology* 9: e1001127.

MUZZALUPO, I., G.G. VENDRAMIN, AND A. CHIAPPETTA. 2014. Genetic Biodiversity of Italian Olives (Olea europaea) Germplasm Analyzed by SSR Markers. *The Scientific World Journal*1–12.

NEI, M., F. TAJIMA, AND Y. TATENO. 1983. Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* 19: 153–170.

NORDIN, M.S., Z. ARIFFIN, R. JAJULI, D.W. WAN ABDULLAH, AND M.G. DENIS. 2007. Country Report on the State of Plant Genetic Resources for Food and Agriculture in Malaysia (1997-2007). Malaysia.

OLSEN, K.M., AND B.L. GROSS. 2008. Detecting multiple origins of domesticated crops. *Proceedings of the National Academy of Sciences of the United States of America* 105: 13701–2.

VAN OOSTERHOUT, C., W.F. HUTCHINSON, D.P.M. WILLS, AND P. SHIPLEY. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538.

PEAKALL, R., AND P.E. SMOUSE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics* 28: 2537–2539.

PRESCOTT-ALLEN, C., AND R. PRESCOTT-ALLEN. 1986. The First Resource: Wild Species in the North American Economy. Yale University Press, New Haven.

PRESCOTT-ALLEN, R., AND C. PRESCOTT-ALLEN. 1990. How many plants feed the world? *Conservation Biology* 4: 365–374.

PRIMACK, R.B. 1985. Comparative studies of fruits in wild and cultivated trees of chempedak (*Artocarpus integer*) and terap (*A. odoratissimus*) in Sarawak, East Malaysia, with additional information on the reproductive biology of the Moraceae in Southeast A. *Malaysian Nature Journal* 39: 1–39.

PRITCHARD, J.K., M. STEPHENS, AND P. DONNELLY. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–59.

R DEVELOPMENT CORE TEAM. 2014. R: A Language and Environment for Statistical Computing.

RAMBAUT, A. 2014. FigTree v1. 3.1: Tree figure drawing tool. *FigTree website*.

SABAH FORESTRY DEPARTMENT. 2012. Annual Report: Chapter 18 (Forest Resource Management).

SAKAI, S., M. KATO, AND H. NAGAMASU. 2000. Artocarpus (Moraceae)-Gall midge pollination mutualism mediated by a male-flower parasitic fungus. *American Journal of Botany* 87: 440–445.

SARAWAK FORESTRY DEPARTMENT. 2012. Annual Report.

SAW, L., J. LAFRANKIE, K. KOCHUMMEN, AND S. YAP. 1991. Fruit trees in a Malaysian rain forest. *Economic Botany* 45: 120–136.

SELKOE, K.A., AND R.J. TOONEN. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9: 615–29.

SINNOTT, R.W. 1984. Virtues of the Haversine. *Sky and Telescope* 68: 159.

SMOUSE, P., AND R. PEAKALL. 1999. Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* 82: 561–573.

TAKEZAKI, N., M. NEI, AND K. TAMURA. 2010. POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. *Molecular Biology and Evolution* 27: 747–52.

TEE, E.S., M. NOOR, M. AZUDIN, AND K. IDRIS. 1988. Nutrient Composition of Malaysian Foods. ASEAN Sub-Committee on Protein, Food Habits Research and Development.

ULLSTRUP, A. 1972. The impacts of the southern corn leaf blight epidemics of 1970-1971. *Annual Review of Phytopathology* 10: 37–50.

WILSON, E.O. 1999. The Diversity of Life. 2nd ed. W. W. Norton & Company.

WITHERUP, C.E., D. RAGONE, T. WIESNER-HANKS, B. IRISH, B. SCHEFFLER, S. SIMPSON, F. ZEE, ET AL. 2013. Development of microsatellite loci in Artocarpus altilis (Moraceae) and cross-amplification in congeneric species. *Applications in Plant Sciences* 1: 1–7.

YAACOB, O., AND S. SUBHADRABANDHU. 1995. Jackfruit, Cempedak, Breadfruit, and Marang/Tarap (*Artocarpus* spp.). *In* The Production of Economic Fruits in South-East Asia, 123–128. Oxford University Press.

YAMASAKI, M., M.I. TENAILLON, I.V. BI, S.G. SCHROEDER, H. SANCHEZ-VILLEDA, J.F. DOEBLEY, B.S. GAUT, AND M.D. MCMULLEN. 2005. A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. *The Plant Cell* 17: 2859–72.\

ZEREGA, N.J.C., D. RAGONE, AND T.J. MOTLEY. 2004. Complex Origins of Breadfruit (Artocarpus altilis, Moraceae): Implications for Human Migrations in Oceania. *American Journal of Botany* 91: 760–766.

ZEREGA, N.J.C., M.N. NUR SUPARDI, AND T.J. MOTLEY. 2010. Phylogeny and Recircumscription of Artocarpeae (Moraceae) with a Focus on *Artocarpus*. *Systematic Botany* 35: 766–782.

ZEREGA, N.J.C., G. RAVIKANTH, E.M. GARDNER, K. LARICCHIA, T. MELHEM, M. WANG, C.E. WITHERUP, ET AL. 2014. New genomic resources to understand genetic diversity and promote crop improvement of jackfruit (Artocarpus heterophyllus), breadfruit (A. altilis), and other underutilized Artocarpus crops. *In* Proceedings of the International Symposium on Jackfruit (Artocarpus heterophyllus) and Breadfruit (Artocarpus altilis Fosberg) of the Tropics, India.

Appendices

Appendix 1. Interview and other miscellaneous data

Interview data
Informal interviews were conducted with tree-owners, guides or villagers from 17 sites (Table 25). Questions asked were usually free-form, with suggestions or prompts if the interviewee was unsure of the meaning of the question. Certain types of information were applicable to the entire site (e.g. care, pests), but sometimes multiple responses were recorded per site for different cultivars or different trees (e.g. cultivar names, tree ages, fruit quality).

Seasonality
The peak season of cempedak was said to be July to August, and sometimes extending to October. However, in July 2013, I was unable to purchase cempedak at markets in Selangor, but I found a few cempedak fruits in supermarkets. Whereas in the north-western states of Penang, Kedah and northern Perak, the cempedak harvest was beginning to wind down in mid-August. In Bidor, Perak, there were many cempedak fruits at the market in early August, although trees in the orchard we visited only had small, unripe fruits. In the Kluang germplasm in Johor (southernmost state), most cempedak fruits were big but not fully ripe in late August.
Most people interviewed said that cempedak bears fruit once or twice a year. In the Bt. Tangga germplasm, trees reportedly fruited three times in 2012. The fruiting season of bangkong seemed to be irregular, as I received a range of answers from interviewees about when bangkong bears fruit (March or July-August (Ulu Geroh), April-July (MARDI FR), August-October (Sungai Buloh), September-October (Sungai Rual), November-December (FRIM Kepong)). One native guide said that the bangkong fruiting comes after the durian season which is typically June-August, and others also commented that the rainy season may affect the timing of bangkong fruiting (late May-September for the Peninsular West Coast, November-March for the Peninsular East Coast).

Cultivars
Cempedak wholesalers at the Selayang Wholesale Market in Selangor described several "elite" cempedak cultivars, namely "Ya Zi (Duck) Red", "seedless" (not actually seedless, but seeds are small and flat), "Chepok" (5-6 kg fruits), CH28 or Kuang, Hulu Langat, Perak Champion, and durian-cempedak. These elite cultivars are more expensive, typically eaten fresh, have thick pulp, and some of them prized for their color (red is an auspicious color in Chinese tradition). They also differentiated between the elite cultivars and the "Kampung" or village variety, which are said to be of lower quality and commonly fried to be eaten as fritters. The elite cultivars are sold at RM4 – RM4.50/kg, while the kampung cultivars are sold at RM3-RM4/kg. It should be noted that the actual color of "red" cempedak is usually bright orange with a tint of red. I also spoke with a Federal Agricultural Marketing Authority (FAMA) officer, who described three types of cempedak, namely "Kampung", hybrid (nangka-cempedak or jackfruit-cempedak), and "cempedak madu" (honey cempedak). The "Kampung" cempedak probably represents cempedak without cultivar names, but could also include land-races, as the germplasm collection in Bt. Tangga was described as "Kampung" clones.

Major Cempedak Production Sites

Based on interviews and conversations with various people, experts and non-experts alike, I compiled a list of localities that are hubs of cempedak production and sales, as a reference for future studies (Appendix Table 2). This list is not meant to be accurate nor exhaustive.

Potential Sites for Bangkong

I identified potential sites for bangkong collection in Peninsular Malaysia based on herbarium records. I did not manage to visit all these sites due to time and technical constraints, but I list them here as a reference for future studies. In Johor, we were not able to find bangkong at the forests along Mawai-Jemaluang road, Kota Tinggi Waterfall, or Gunung Pulai 2 – although it was only a team of two (the field assistant and myself) searching in unfamiliar sites without any sort of guide or information except for herbarium labels, as the locals we talked to did not know where bangkong might be located. The Leiden Herbarium also contains many vouchers of wild cempedak and possibly bangkong, including from Borneo, Sumatra, Java, the Moluccas, New Guinea, and even one voucher as far east as Hawaii. The Leiden Herbarium database can be accessed at: http://www.virtualherbarium.org/vhportal.html.

Appendix Table 1. Information collected through interviews with tree-owners, guides or villagers from 17 sites (five bangkong sites, 11 cempedak sites, one nangka-cempedak site) in Peninsular Malaysia. Numbers in parentheses indicate how many times the response was given.

Category	**Responses**
Vernacular name/variety name	**Bangkong names**: Bangkong, temenguk, dekoh, tekik. **Cempedak varieties**: small-seeded variety, white variety, Red Ah Chai, durian-cempedak, nangka-cempedak, red variety from Balik Pulau, Liew Kim Fook var., CH29, CH30, Jenderi, Bukit Selambau, Bukit Selambau 2, Kuala Nerang, Perak Champion, Sik, Pendang, Berseri, Bantai, CH35, CH34, CH33, CH28.
Tree age	Range (6-60+)
Fruiting season (which months)	Range of months: March-Dec (bangkong), July-Oct (cempedak), all year (nangka-cempedak)
Fruiting frequency per year	**Bangkong**: Once (1), 1-2 times (1) **Cempedak**: Once (4), twice (6), 1-2 times (4)
Fruit description	Traits mentioned: Latex, taste of seed, density of achenes, size of seeds, size of perianth-balls, thickness of pulp, sweetness, texture, yield, size of fruit, whether or not it is good for frying
Color of pulp	white (5), yellow (8), yellow-orange (3), orange (3), orange-red (3), red (4), pink (1)
Source of propagation material	Bought from nursery (5), grown by elders (4), bought from DoA (1). Geographical provenance recorded for five trees.
Propagation method	Seed (3), grafted (6)
Care	Chemical fertilizers (3); no care (1)
Pests	Squirrel (4), monkey (1), wasps (1), ants (1)
Selling price (RM/kg)	Range: Wholesale: RM1-RM1.50; Market price: RM4-RM4.50

Appendix Table 2. List of localities that were reported to be hubs of cempedak production. Please not that this list is not meant to be accurate nor exhaustive.

State	**Locality (District or Town)**
Terengganu	Hulu Paka (DoA collection), Jertih, Kuala Terengganu
Kedah	Yan
Penang	Balik Pulau
Perak	Teluk Intan, Bidor, Baling
Pahang	Raub, Bentong
Selangor	Kuang, Sungai Buloh, Hulu Langat, Kuala Selangor
Negeri Sembilan	Mantin
Johor	Kluang

Appendix Table 3. Potential sampling locations for bangkong based on herbarium records. Some cempedak (Artocarpus integer) were included here if I suspect they may be bangkong based on their location in wild areas.

Taxon name on herbarium sheet	**State**	**District**	**Locality**
Artocarpus integer var. silvestris	Johor	Batu Pahat	Soga F.R.
Artocarpus integer var. silvestris	Johor	Johor Bahru	Gn. Pulai F.R., Gn. Pulai
Artocarpus integer var. silvestris	Johor	Kluang	Kluang F.R.
Artocarpus integer var. silvestris	Johor	Kota Tinggi	Mawai Rd. (Mawai-Jemaluang road)
Artocarpus integer var. silvestris	Johor	Mersing	Gn. Arong F.R.
Artocarpus integer var. silvestris	Johor	Mersing	Mersing F.R.
Artocarpus integer	Johor	Mersing	State Land Lesen E 7/65
Artocarpus integer	Johor	Pontian	P. Pisang
Artocarpus integer	Kelantan		Gn. Chamar
Artocarpus integer var. silvestris	Kelantan	Gua Musang	Sg. Lebir, Kuala Ternya
Artocarpus integer var. silvestris	Kelantan	Gua Musang	Sg. Mersing
Artocarpus integer	Kelantan	Gua Musang	Sg. Perias, Kuala May
Artocarpus integer var. silvestris	Kelantan	Jeli	Jeli
Artocarpus integer var. silvestris	Kelantan	Kuala Krai	Stong Tengah F.R.
Artocarpus integer	Kelantan		Sg. Lebir, Kuala Relai
Artocarpus integer	Kelantan		Sg. Merkill, Bukit Bakar
Artocarpus integer	Kelantan		Sg. Nenggiri, Kampung Jenera
Artocarpus integer	Negeri Sembilan	Seremban	Lenggeng F.R.
Artocarpus integer var. silvestris	Pahang	Cameron Highlands	Sg. Bertam, Kuala Mensum
Artocarpus integer	Pahang	Fraser's Hill	Fraser's Hill
Artocarpus integer	Pahang	Fraser's Hill	Jeriau Waterfall
Artocarpus integer var. silvestris	Pahang	Jerantut	Taman Negara, Sg. Sepia, Jeram Perahau

Taxon name on herbarium sheet	State	District	Locality
Artocarpus integer	Pahang	Jerantut	Taman Negara, Sg. Tahan, Lata Berkoh
Artocarpus integer var. silvestris	Pahang	Jerantut	Tekam F.R.
Artocarpus integer	Pahang	Lesong Forest Reserve	Lesong Forest Reserve
Artocarpus integer var. silvestris	Pahang	Lipis	Sg. Telom, Kuala Mesong
Artocarpus integer var. silvestris	Pahang	Lipis	Taman Negara, Merapoh
Artocarpus integer var. silvestris	Pahang	Raub	
Artocarpus integer var. silvestris	Pahang	Rompin	Lesong F.R.
Artocarpus integer	Pahang		Sungai Telom
Artocarpus integer	Pahang		Ulu Bertam, Kuala Mensum
Artocarpus integer	Perak	Hulu Perak	Fort Kemer
Artocarpus integer var. silvestris	Perak	Hulu Perak	Grik
Artocarpus integer var. silvestris	Perak	Kinta	Gopeng, Sg. Groh
Artocarpus integer	Perak	Tapah	Kampung Sekam
Artocarpus integer var. silvestris	Selangor	Gombak	Bkt. Lagong F.R.
Artocarpus integer	Selangor	Gombak	Forest Research Institute Malaysia
Artocarpus integer	Selangor	Gombak	Kepong Plantations
Artocarpus integer var. silvestris	Selangor	Gombak	Ulu Gombak (University of Malaya field station)
Artocarpus integer var. silvestris	Selangor	Hulu Langat	Sg. Lalang F.R.
Artocarpus integer	Selangor	Hulu Selangor	Semangkok F.R.
Artocarpus integer	Selangor	Kepong	Bukit Baru, IPP
Artocarpus integer	Selangor	Petaling	Bkt. Lanjan
Artocarpus integer var. silvestris	Selangor	Petaling	Sg. Buloh F.R.
Artocarpus integer var. silvestris	Terengganu	Dungun	Bkt. Bauk F.R.
Artocarpus integer var. silvestris	Terengganu	Dungun	Kpg. Padang Pulut
Artocarpus integer var. silvestris	Terengganu	Dungun	Taman Negara, Sg. Loh, Kuala Datok
Artocarpus integer	Terengganu	Hulu Terengganu	Tasik Kenyir
Artocarpus integer var. silvestris	Terengganu	Kuala Terengganu	Kpg. Sg. Rengas

Appendix 2. Data for samples used in this study.

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
30-Jul-13	MW101	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.238968	101.623934	
30-Jul-13	MW102	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.239040	101.623880	
30-Jul-13	MW103	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.239062	101.623756	
30-Jul-13	MW104	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.239011	101.623860	
30-Jul-13	MW105	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.238919	101.623924	
30-Jul-13	MW106	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.238570	101.623981	
30-Jul-13	MW107	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.238641	101.623851	
30-Jul-13	MW108	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.238456	101.623869	
30-Jul-13	MW109	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.238517	101.623712	
30-Jul-13	MW110	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.238546	101.623678	
30-Jul-13	MW111	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.238315	101.623689	
30-Jul-13	MW112	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.238415	101.623630	
30-Jul-13	MW113	cempedak		10b-FRIM Kepong	May be part of durian orchard owned by local	Selangor	3.228175	101.619914	
30-Jul-13	MW114	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.238829	101.623609	
30-Jul-13	MW115	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.238808	101.623212	
30-Jul-13	MW116	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.239029	101.623034	
30-Jul-13	MW117	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.239135	101.623186	
30-Jul-13	MW118	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.239411	101.623087	
30-Jul-13	MW119	bangkong		10b-FRIM Kepong	FRIM forest	Selangor	3.239305	101.622309	
2-Aug-13	MW201	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.985940	101.686613	
2-Aug-13	MW202	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.985765	101.686727	
2-Aug-13	MW203	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.985789	101.686685	
2-Aug-13	MW204	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.985650	101.686956	
30-Sep-13	MW205	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.985908	101.686328	
30-Sep-13	MW206	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.986661	101.686048	
30-Sep-13	MW207	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.986789	101.685897	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
30-Sep-13	MW208	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.987440	101.685502	
30-Sep-13	MW209	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.987564	101.685585	
30-Sep-13	MW210	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.987458	101.685720	
30-Sep-13	MW211	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.987486	101.685842	
30-Sep-13	MW212	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.987673	101.685521	
30-Sep-13	MW213	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.987651	101.685267	
30-Sep-13	MW214	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.988153	101.686539	
30-Sep-13	MW215	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.988306	101.686353	
30-Sep-13	MW216	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.986278	101.687416	
30-Sep-13	MW217	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.985333	101.687246	
30-Sep-13	MW218	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.986365	101.687975	
30-Sep-13	MW219	bangkong		10f-MARDI FR	Forest near MARDI HQ, Serdang	Selangor	2.999916	101.687577	
30-Sep-13	MW301	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.229911	101.586279	
30-Sep-13	MW302	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230461	101.585742	
30-Sep-13	MW303	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230601	101.585549	
30-Sep-13	MW304	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230278	101.585446	
30-Sep-13	MW305	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230367	101.585388	
30-Sep-13	MW306	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230278	101.585109	
30-Sep-13	MW307	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230236	101.585061	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
30-Sep-13	MW308	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230751	101.585087	
30-Sep-13	MW309	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.231956	101.584572	
30-Sep-13	MW310	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230449	101.585603	
30-Sep-13	MW312	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230429	101.585683	
30-Sep-13	MW313	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.231071	101.585585	
30-Sep-13	MW314	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.231153	101.585860	
30-Sep-13	MW315	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.231212	101.585754	
30-Sep-13	MW316	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.231049	101.585903	
30-Sep-13	MW317	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230961	101.586045	
30-Sep-13	MW318	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230815	101.585944	
30-Sep-13	MW319	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230820	101.585962	
30-Sep-13	MW320	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230632	101.586503	
30-Sep-13	MW321	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230632	101.586503	partial site coord. only
22-Sep-13	MW322	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230632	101.586503	partial site coord. only
22-Sep-13	MW323	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230632	101.586503	partial site coord. only
22-Sep-13	MW324	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230632	101.586503	partial site coord. only
22-Sep-13	MW325	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230632	101.586503	partial site coord. only
22-Sep-13	MW326	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230632	101.586503	partial site coord. only
22-Sep-13	MW327	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230632	101.586503	partial site coord. only

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
22-Sep-13	MW328	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230632	101.586503	partial site coord. only
22-Sep-13	MW329	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230632	101.586503	partial site coord. only
22-Sep-13	MW330	bangkong		10c-Sg Buloh	Forest near Orang Asli Village	Selangor	3.230632	101.586503	partial site coord. only
22-Sep-13	MW351	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230315	101.588317	partial site coord. only
22-Sep-13	MW352	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230315	101.588317	partial site coord. only
22-Sep-13	MW353	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230315	101.588317	partial site coord. only
22-Sep-13	MW354	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230315	101.588317	partial site coord. only
22-Sep-13	MW355	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230315	101.588317	partial site coord. only
22-Sep-13	MW356	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230315	101.588317	partial site coord. only
22-Sep-13	MW357	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230315	101.588317	partial site coord. only
22-Sep-13	MW358	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230315	101.588317	partial site coord. only
22-Sep-13	MW359	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230315	101.588317	partial site coord. only
22-Sep-13	MW360	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230315	101.588317	partial site coord. only
22-Sep-13	MW361	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230272	101.587924	
22-Sep-13	MW362	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230348	101.587930	
22-Sep-13	MW363	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230411	101.587951	
22-Sep-13	MW364	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230333	101.587913	
22-Sep-13	MW365	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.230293	101.587786	
22-Sep-13	MW366	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.229709	101.587281	
22-Sep-13	MW367	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.229664	101.587121	
22-Sep-13	MW368	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.229856	101.587027	
22-Sep-13	MW369	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.229936	101.586990	
22-Sep-13	MW370	cempedak		10c-Sg Buloh	Orang Asli Village	Selangor	3.229709	101.586449	
5-Aug-13	MW401	hybrid	nangka-cemp	5a-Pasir Mas	Private yard.	Kelantan	6.042011	102.145535	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
20-Aug-13	MW402	cempedak		5a-Pasir Mas	Private yard.	Kelantan	6.038974	102.143176	
20-Aug-13	MW403	cempedak		5a-Pasir Mas	Private yard.	Kelantan	6.038868	102.143293	
6-Aug-13	MW410	cempedak		5b-Bt Panau	Orchard.	Kelantan	5.894636	102.140710	
6-Aug-13	MW411	cempedak		5b-Bt Panau	Orchard.	Kelantan	5.894677	102.140602	
6-Aug-13	MW412	cempedak		5b-Bt Panau	Orchard.	Kelantan	5.894689	102.140588	
6-Aug-13	MW413	cempedak		5b-Bt Panau	Orchard.	Kelantan	5.894655	102.140962	
6-Aug-13	MW414	cempedak		5b-Bt Panau	Orchard.	Kelantan	5.894682	102.141039	
6-Aug-13	MW415	cempedak		5b-Bt Panau	Orchard.	Kelantan	5.894679	102.141102	
6-Aug-13	MW416	cempedak		5b-Bt Panau	Orchard.	Kelantan	5.894633	102.141214	
6-Aug-13	MW417	cempedak		5b-Bt Panau	Orchard.	Kelantan	5.894647	102.141416	
6-Aug-13	MW418	cempedak		5b-Bt Panau	Orchard.	Kelantan	5.894675	102.141446	
6-Aug-13	MW419	cempedak		5b-Bt Panau	Orchard.	Kelantan	5.894688	102.141617	
6-Aug-13	MW420	cempedak		5b-Bt Panau	Orchard.	Kelantan	5.894739	102.141594	
6-Aug-13	MW421	cempedak		5b-Bt Panau	Orchard.	Kelantan	5.894645	102.141720	
6-Aug-13	MW422	cempedak		5b-Bt Panau	Orchard.	Kelantan	5.894634	102.141883	
6-Aug-13	MW423	cempedak		5b-Bt Panau	Orchard.	Kelantan	5.894567	102.141883	
6-Aug-13	MW424	cempedak	nangka-cemp	5b-Bt Panau	Orchard.	Kelantan	5.894162	102.140715	
6-Aug-13	MW425	hybrid	nangka-cemp	5a-Chekok	Orchard.	Kelantan	5.911806	102.190317	
6-Aug-13	MW426	hybrid	nangka-cemp	5a-Chekok	Orchard.	Kelantan	5.911835	102.190191	
6-Aug-13	MW428	hybrid	nangka-cemp	5a-Chekok	Orchard.	Kelantan	5.911679	102.190196	
9-Aug-13	MW429	cempedak		5d-Kg Bt Jering	Private yard.	Kelantan	5.483191	101.903762	
9-Aug-13	MW430	cempedak		5d-Kg Bt Jering	Private yard.	Kelantan	5.482703	101.903824	
9-Aug-13	MW431	cempedak		5d-Kg Bt Jering	Private yard.	Kelantan	5.482746	101.904143	
16-Aug-13	MW441	cempedak		5a-Lubok Stol	Private yard.	Kelantan	5.948717	101.941910	
16-Aug-13	MW442	cempedak		5a-Lubok Stol	Private yard.	Kelantan	5.946306	101.940709	
16-Aug-13	MW443	cempedak		5a-Lubok Stol	Private yard.	Kelantan	5.944043	101.937340	
8-Aug-13	MW501	cempedak	Red Ah Chai	8c-Bidor	Orchard.	Perak	4.128831	101.274467	
8-Aug-13	MW502	cempedak		8c-Bidor	Orchard.	Perak	4.128784	101.274517	
8-Aug-13	MW503	cempedak		8c-Bidor	Orchard.	Perak	4.128582	101.274549	
8-Aug-13	MW504	cempedak	durian-cemp	8c-Bidor	Orchard.	Perak	4.128737	101.274644	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
8-Aug-13	MW505	cempedak	durian-cemp	8c-Bidor	Orchard.	Perak	4.128722	101.274722	
8-Aug-13	MW506	cempedak	Red Balik Pulau	8c-Bidor	Orchard.	Perak	4.128671	101.274735	
8-Aug-13	MW507	cempedak	durian-cemp	8c-Bidor	Orchard.	Perak	4.128742	101.274864	
8-Aug-13	MW508	cempedak	durian-cemp	8c-Bidor	Orchard.	Perak	4.129078	101.274742	
8-Aug-13	MW509	cempedak	durian-cemp	8c-Bidor	Orchard.	Perak	4.129616	101.272202	
8-Aug-13	MW510	cempedak	durian-cemp	8c-Bidor	Orchard.	Perak	4.129220	101.274806	
8-Aug-13	MW511	cempedak	durian-cemp	8c-Bidor	Orchard.	Perak	4.129348	101.274955	
8-Aug-13	MW512	cempedak	seedless	8c-Bidor	Orchard.	Perak	4.129241	101.274966	
8-Aug-13	MW514	cempedak	durian-cemp	8c-Bidor	Orchard.	Perak	4.111293	101.286838	
13-Aug-13	MW601	cempedak		5c-Tanah Merah	Private yard.	Kelantan	5.812382	102.134008	
13-Aug-13	MW602	cempedak		5c-Tanah Merah	In palm oil estate.	Kelantan	5.808832	102.084290	
13-Aug-13	MW603	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851649	102.006809	
13-Aug-13	MW604	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851534	102.006860	
13-Aug-13	MW605	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851543	102.006906	
13-Aug-13	MW606	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851515	102.006968	
13-Aug-13	MW607	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851510	102.007012	
13-Aug-13	MW608	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851524	102.007066	
13-Aug-13	MW609	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851481	102.007071	
13-Aug-13	MW610	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851379	102.007116	
13-Aug-13	MW611	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851328	102.007112	
13-Aug-13	MW612	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851358	102.007071	
13-Aug-13	MW613	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851329	102.007028	
13-Aug-13	MW614	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851247	102.007016	
13-Aug-13	MW615	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851309	102.007158	
13-Aug-13	MW616	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851318	102.007255	
13-Aug-13	MW617	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851386	102.007335	
13-Aug-13	MW618	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851404	102.007412	
13-Aug-13	MW619	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851425	102.007413	
13-Aug-13	MW620	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851445	102.007433	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
13-Aug-13	MW621	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851650	102.007391	
13-Aug-13	MW622	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851668	102.007285	
13-Aug-13	MW623	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851719	102.007351	
13-Aug-13	MW624	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851752	102.007303	
13-Aug-13	MW625	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851855	102.007274	
13-Aug-13	MW626	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851872	102.007239	
13-Aug-13	MW627	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851873	102.007167	
13-Aug-13	MW628	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851804	102.007153	
13-Aug-13	MW629	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851850	102.007083	
13-Aug-13	MW630	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851768	102.006943	
13-Aug-13	MW632	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851682	102.006961	
13-Aug-13	MW633	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851840	102.006818	
13-Aug-13	MW634	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851889	102.006849	
13-Aug-13	MW635	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.851965	102.006932	
13-Aug-13	MW636	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.852042	102.006975	
13-Aug-13	MW637	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.852080	102.006872	
13-Aug-13	MW638	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.852307	102.006929	
13-Aug-13	MW639	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.852376	102.006844	
13-Aug-13	MW640	cempedak		5c-FELDA Kemahang	Orchard.	Kelantan	5.852288	102.006720	
17-Aug-13	MW701	cempedak		7a-Balik Pulau	Orchard.	Penang	5.349169	100.246931	
17-Aug-13	MW702	cempedak		7a-Balik Pulau	Orchard.	Penang	5.349579	100.246944	
17-Aug-13	MW703	cempedak		7a-Balik Pulau	Orchard.	Penang	5.349672	100.247092	
17-Aug-13	MW704	cempedak		7a-Balik Pulau	Orchard.	Penang	5.349697	100.247152	
17-Aug-13	MW705	cempedak		7a-Balik Pulau	Orchard.	Penang	5.349737	100.247203	
17-Aug-13	MW706	cempedak		7a-Balik Pulau	Orchard.	Penang	5.349776	100.247288	
17-Aug-13	MW707	cempedak		7a-Balik Pulau	Orchard.	Penang	5.349810	100.247368	
17-Aug-13	MW708	cempedak		7a-Balik Pulau	Orchard.	Penang	5.349962	100.247632	
17-Aug-13	MW709	cempedak		7a-Balik Pulau	Orchard.	Penang	5.349975	100.247714	
17-Aug-13	MW710	cempedak		7a-Balik Pulau	Orchard.	Penang	5.349790	100.247991	
17-Aug-13	MW711	cempedak	Liew	7b-Liew	Orchard.	Penang	5.349678	100.248828	
17-Aug-13	MW712	cempedak		7a-Balik Pulau	Orchard.	Penang	5.349821	100.248869	
17-Aug-13	MW713	cempedak	Liew	7b-Liew	Orchard.	Penang	5.349911	100.248979	
17-Aug-13	MW714	cempedak		7a-Balik Pulau	Orchard.	Penang	5.350160	100.249162	
17-Aug-13	MW715	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350437	100.249351	
17-Aug-13	MW716	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350412	100.249382	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
17-Aug-13	MW717	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350348	100.249417	
17-Aug-13	MW718	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350277	100.249451	
17-Aug-13	MW719	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350188	100.249536	
17-Aug-13	MW720	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350188	100.249607	
17-Aug-13	MW721	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350140	100.249617	
17-Aug-13	MW722	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350142	100.249653	
17-Aug-13	MW723	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350090	100.249905	
17-Aug-13	MW724	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350167	100.249840	
17-Aug-13	MW725	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350229	100.249811	
17-Aug-13	MW726	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350314	100.249762	
17-Aug-13	MW727	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350480	100.249847	
17-Aug-13	MW728	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350444	100.249835	
17-Aug-13	MW729	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350377	100.249937	
17-Aug-13	MW730	cempedak	Liew	7b-Liew	Orchard.	Penang	5.350337	100.249992	
17-Aug-13	MW731	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326260	100.238712	
17-Aug-13	MW732	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326220	100.238701	
17-Aug-13	MW733	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326363	100.238732	
17-Aug-13	MW734	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326430	100.238561	
17-Aug-13	MW735	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326428	100.238526	
17-Aug-13	MW736	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326460	100.238499	
17-Aug-13	MW737	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326498	100.238334	
17-Aug-13	MW738	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326520	100.238223	
17-Aug-13	MW739	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326428	100.238136	
17-Aug-13	MW740	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326328	100.238272	
17-Aug-13	MW741	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326596	100.238044	
17-Aug-13	MW742	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326445	100.238049	
17-Aug-13	MW743	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326348	100.238074	
17-Aug-13	MW744	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326368	100.237981	
17-Aug-13	MW745	cempedak		7c-E&J Fruit Farm	Orchard.	Penang	5.326413	100.237951	
19-Aug-13	MW801	cempedak	CH29	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.486680	100.483289	
19-Aug-13	MW802	cempedak	CH30	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485961	100.483271	
19-Aug-13	MW803	cempedak	CH30	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485827	100.483175	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
19-Aug-13	MW804	cempedak	CH30	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485794	100.483248	
19-Aug-13	MW805	cempedak	CH30	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485790	100.483336	
19-Aug-13	MW806	cempedak	CH30	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485859	100.483482	
19-Aug-13	MW807	cempedak	Jeneri	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485513	100.483543	
19-Aug-13	MW808	cempedak	Jeneri*	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485553	100.483592	
19-Aug-13	MW809	cempedak	Bukit Selambau	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485576	100.483665	
19-Aug-13	MW810	cempedak	Bukit Selambau	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485592	100.483740	
19-Aug-13	MW811	cempedak	Bukit Selambau	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485638	100.483801	
19-Aug-13	MW812	cempedak	Bukit Selambau	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485659	100.483904	
19-Aug-13	MW813	cempedak	Kuala Nerang	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485680	100.483989	
19-Aug-13	MW814	cempedak	Kuala Nerang	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485709	100.484045	
19-Aug-13	MW815	cempedak	Kuala Nerang	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485714	100.484144	
19-Aug-13	MW816	cempedak	Kuala Nerang	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485732	100.484205	
19-Aug-13	MW817	cempedak	Bukit Selambau 2	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485768	100.484282	
19-Aug-13	MW818	cempedak	Bukit Selambau 2	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485810	100.484351	
19-Aug-13	MW819	cempedak	Bukit Selambau 2	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485873	100.484517	
19-Aug-13	MW820	cempedak	Perak Champion	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485706	100.484584	
19-Aug-13	MW821	cempedak	Perak Champion	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485653	100.484628	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
19-Aug-13	MW822	cempedak	Bukit Manik	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485657	100.484455	
19-Aug-13	MW823	cempedak	Bukit Manik	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485657	100.484379	
19-Aug-13	MW824	cempedak	Bukit Manik	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485631	100.484285	
19-Aug-13	MW825	cempedak	Bukit Manik	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485566	100.484112	
19-Aug-13	MW826	cempedak	Bayas	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485511	100.483956	
19-Aug-13	MW827	cempedak	Bayas	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485462	100.483884	
19-Aug-13	MW828	cempedak	Bayas	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485443	100.483817	
19-Aug-13	MW829	cempedak	Bayas	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485351	100.483639	
19-Aug-13	MW830	cempedak	Sik	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485362	100.483369	
19-Aug-13	MW831	cempedak	Sik	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485323	100.483326	
19-Aug-13	MW832	cempedak	Pendang	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485458	100.483601	
19-Aug-13	MW833	cempedak	Pendang	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485499	100.483715	
19-Aug-13	MW834	cempedak	Pendang	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485538	100.483752	
19-Aug-13	MW835	cempedak	Pendang	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485536	100.483835	
19-Aug-13	MW836	cempedak	Pendang	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485560	100.483909	
19-Aug-13	MW837	cempedak	Berseri	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485583	100.483998	
19-Aug-13	MW838	cempedak	Berseri	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485612	100.484082	
19-Aug-13	MW839	cempedak	Berseri*	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485650	100.484162	
19-Aug-13	MW840	cempedak	Berseri*	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485684	100.484250	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
19-Aug-13	MW841	cempedak	Bantai	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485716	100.484284	
19-Aug-13	MW842	cempedak	Bantai	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.485995	100.484370	
19-Aug-13	MW843	cempedak	CH29	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.486836	100.483664	
19-Aug-13	MW844	cempedak	CH29	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.486964	100.483535	
19-Aug-13	MW845	cempedak	CH29	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.486927	100.483385	
19-Aug-13	MW846	cempedak	CH29	6a-Bt Tangga Germplasm	MARDI germplasm collection.	Kedah	6.486937	100.483212	
19-Aug-13	MW851	cempedak		8a-Grik	Growing by restaurant near nursery.	Perak	5.416667	101.133333	
19-Aug-13	MW852	cempedak		8a-Grik	Growing by restaurant near nursery.	Perak	5.416667	101.133333	
19-Aug-13	MW853	cempedak		8a-Grik	Growing by restaurant near nursery.	Perak	5.416667	101.133333	
19-Aug-13	MW854	cempedak		8a-Grik	Growing by restaurant near nursery.	Perak	5.416667	101.133333	
19-Aug-13	MW855	cempedak		8a-Grik	Growing by restaurant near nursery.	Perak	5.416667	101.133333	
22-Aug-13	MW971	cempedak	CH35	12e-Ayer Hitam CDS	Cultivar clone collection.	Johor	1.944376	103.197678	
22-Aug-13	MW972	cempedak	CH27	12e-Ayer Hitam CDS	Cultivar clone collection.	Johor	1.944758	103.197924	
22-Aug-13	MW973	cempedak	CH34	12e-Ayer Hitam CDS	Cultivar clone collection.	Johor	1.945118	103.197543	
22-Aug-13	MW974	cempedak	CH27	12e-Ayer Hitam CDS	Cultivar clone collection.	Johor	1.945042	103.197131	
22-Aug-13	MW975	cempedak	CH35	12e-Ayer Hitam CDS	Cultivar clone collection.	Johor	1.944722	103.197115	
22-Aug-13	MW976	cempedak	CH27	12e-Ayer Hitam CDS	Cultivar clone collection.	Johor	1.944375	103.196808	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
22-Aug-13	MW901	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948758	103.366659	
22-Aug-13	MW902	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948873	103.366762	
22-Aug-13	MW903	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948959	103.366788	
22-Aug-13	MW905	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948997	103.366778	
22-Aug-13	MW906	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949055	103.366800	
22-Aug-13	MW908	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949179	103.366897	
22-Aug-13	MW909	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949208	103.366794	
22-Aug-13	MW910	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949267	103.366849	
22-Aug-13	MW911	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949328	103.366718	
22-Aug-13	MW912	cempedak	nangka-cemp	12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949370	103.366805	
23-Aug-13	MW914	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948666	103.367197	
23-Aug-13	MW915	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948885	103.367398	
23-Aug-13	MW916	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948834	103.367203	
23-Aug-13	MW917	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948936	103.367292	
23-Aug-13	MW918	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948947	103.367162	
23-Aug-13	MW921	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949085	103.367214	
23-Aug-13	MW922	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949148	103.367312	
23-Aug-13	MW924	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949229	103.367160	
23-Aug-13	MW925	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949288	103.367175	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
23-Aug-13	MW926	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949330	103.367219	
23-Aug-13	MW927	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948656	103.367792	
23-Aug-13	MW928	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948703	103.367818	
23-Aug-13	MW929.b	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948768	103.367707	
23-Aug-13	MW929	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948783	103.367905	
23-Aug-13	MW931	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948866	103.367770	
23-Aug-13	MW932	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948918	103.367772	
23-Aug-13	MW933	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948984	103.367860	
23-Aug-13	MW934	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949025	103.367884	
23-Aug-13	MW935	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949094	103.367762	
23-Aug-13	MW936	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949148	103.367845	
23-Aug-13	MW937	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949206	103.367632	
23-Aug-13	MW937.b	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949205	103.367624	
23-Aug-13	MW939	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949312	103.367651	
23-Aug-13	MW940	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948650	103.367875	
23-Aug-13	MW941	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948706	103.368109	
23-Aug-13	MW942	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948784	103.368155	
23-Aug-13	MW944	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948879	103.367788	
23-Aug-13	MW945	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948935	103.367769	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
23-Aug-13	MW946	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.948978	103.367869	
23-Aug-13	MW947	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949025	103.367878	
23-Aug-13	MW948	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949093	103.367679	
23-Aug-13	MW949	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949142	103.367835	
23-Aug-13	MW950	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949203	103.367831	
23-Aug-13	MW951	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949240	103.367863	
23-Aug-13	MW952	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949286	103.367779	
23-Aug-13	MW953	cempedak	CH26 Paya Jaras	12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949391	103.367747	
23-Aug-13	MW954	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949403	103.367738	
23-Aug-13	MW955	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949453	103.367788	
23-Aug-13	MW956	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949517	103.367783	
23-Aug-13	MW957	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949529	103.367715	
23-Aug-13	MW958	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949618	103.367773	
23-Aug-13	MW959	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949694	103.367816	
23-Aug-13	MW964	cempedak		12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949378	103.367607	
23-Aug-13	MW965	cempedak	Sri Harum 2	12d-Kluang Germplasm	MARDI germplasm collection.	Johor	1.949419	103.367563	
26-Aug-13	MW1001	cempedak		11a-Pasoh FR	In palm oil estate on the road to Pasoh 50-ha plot.	Negeri Sembilan	2.955571	102.288461	
27-Aug-13	MW1002	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree not tagged.	Negeri Sembilan	2.982000	102.313000	
27-Aug-13	MW1003	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree no. 92088	Negeri Sembilan	2.982000	102.313000	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
27-Aug-13	MW1004	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree no. 92192	Negeri Sembilan	2.982000	102.313000	
27-Aug-13	MW1005	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree no.181642.	Negeri Sembilan	2.982000	102.313000	
27-Aug-13	MW1006	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree no.201082.	Negeri Sembilan	2.982000	102.313000	
27-Aug-13	MW1007	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree no.192545.	Negeri Sembilan	2.982000	102.313000	
27-Aug-13	MW1008	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree no.192751.	Negeri Sembilan	2.982000	102.313000	
27-Aug-13	MW1009	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree no.291738.	Negeri Sembilan	2.982000	102.313000	
27-Aug-13	MW1010	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree no.311758.	Negeri Sembilan	2.982000	102.313000	
27-Aug-13	MW1011	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree no.331393.	Negeri Sembilan	2.982000	102.313000	
27-Aug-13	MW1012	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree no.413841.	Negeri Sembilan	2.982000	102.313000	
27-Aug-13	MW1013	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree no.453996.	Negeri Sembilan	2.982000	102.313000	
27-Aug-13	MW1014	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree no.355128.	Negeri Sembilan	2.982000	102.313000	
27-Aug-13	MW1015	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree no.315580.	Negeri Sembilan	2.982000	102.313000	
27-Aug-13	MW1016	bangkong		11a-Pasoh FR	Pasoh 50-ha plot. Tree no.185170.	Negeri Sembilan	2.982000	102.313000	
12-Sep-13	MW1101	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1102	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1103	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1104	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1105	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1106	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1107	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1108	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1109	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1110	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1111	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1112	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
12-Sep-13	MW1113	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1114	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1115	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1116	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1117	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1118	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1119	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1120	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1121	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1122	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1123	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1124	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1125	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1126	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1127	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1128	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1129	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1130	bangkong		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1151	cempedak		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1152	cempedak		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
12-Sep-13	MW1153	cempedak		8b-Ulu Geroh	Orang Asli Village	Perak	4.433333	101.249979	site coord. only
23-Sep-13	MW1201	bangkong		12c-Gn Lambak	On the way to campground along main trail	Johor	2.025211	103.345424	
24-Sep-13	MW1202	bangkong		12c-Gn Belumut	Gunung Belumut Educational Forest	Johor	2.065704	103.526437	
24-Sep-13	MW1203	bangkong		12c-Gn Belumut	Beside river	Johor	2.065118	103.526928	
24-Sep-13	MW1204	bangkong		12c-Gn Belumut	By entrance arch sign.	Johor	2.065542	103.526454	
24-Sep-13	MW1205	bangkong		12b-Lenggor FR	Beside dirt track outside the gate to Lenggor FR	Johor	2.183333	103.666678	site coord. only
24-Sep-13	MW1206	bangkong		12b-Lenggor FR	Park area of Lenggor Forest Reserve.	Johor	2.183333	103.666678	site coord. only
25-Sep-13	MW1207	cempedak		12f-Kg Sg Selangi	Planted, in Orang Asli Village	Johor	1.860420	103.992685	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
26-Sep-13	MW1208	bangkong		12g-Gn Pulai	Along road to telecommunications tower on forested hill.	Johor	1.590477	103.517896	
26-Sep-13	MW1209	bangkong		12g-Gn Pulai	Along road to telecommunications tower on forested hill.	Johor	1.590543	103.518544	
26-Sep-13	MW1210	bangkong		12g-Gn Pulai	Along road to telecommunications tower on forested hill.	Johor	1.590930	103.519081	
26-Sep-13	MW1211	bangkong		12g-Gn Pulai	Along road to telecommunications tower on forested hill.	Johor	1.592206	103.522915	
26-Sep-13	MW1212	bangkong		12g-Gn Pulai	Along road to telecommunications tower on forested hill.	Johor	1.592223	103.523513	
26-Sep-13	MW1213	bangkong		12g-Gn Pulai	Along road to telecommunications tower on forested hill.	Johor	1.592168	103.523482	
26-Sep-13	MW1214	bangkong		12g-Gn Pulai	Along road to telecommunications tower on forested hill.	Johor	1.592211	103.523627	
26-Sep-13	MW1215	bangkong		12g-Gn Pulai	Along road to telecommunications tower on forested hill.	Johor	1.590933	103.524436	
26-Sep-13	MW1216	bangkong		12g-Gn Pulai	Along road to telecommunications tower on forested hill.	Johor	1.591538	103.528199	
26-Sep-13	MW1217	bangkong		12g-Gn Pulai	Along road to telecommunications tower on forested hill.	Johor	1.592307	103.523436	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
26-Sep-13	MW1218	bangkong		12g-Gn Pulai	Along road to telecommunications tower on forested hill.	Johor	1.592021	103.521566	
26-Sep-13	MW1219	cempedak		12g-Setia Jaya	Next to road to Hutan Lipur Gunung Pulai 2	Johor	1.569160	103.510593	
26-Sep-13	MW1220	cempedak		12g-Setia Jaya	In palm oil estate.	Johor	1.569030	103.510768	
29-Sep-13	MW1301	bangkong		10a-Ulu Serendah	Forest near Orang Asli Village	Selangor	3.365343	101.615844	site coord. only
29-Sep-13	MW1302	bangkong		10a-Ulu Serendah	Forest near Orang Asli Village	Selangor	3.365343	101.615844	site coord. only
29-Sep-13	MW1303	bangkong		10a-Ulu Serendah	Forest near Orang Asli Village	Selangor	3.365343	101.615844	site coord. only
29-Sep-13	MW1304	bangkong		10a-Ulu Serendah	Forest near Orang Asli Village	Selangor	3.365343	101.615844	site coord. only
29-Sep-13	MW1305	bangkong		10a-Ulu Serendah	Forest near Orang Asli Village	Selangor	3.365343	101.615844	site coord. only
29-Sep-13	MW1306	bangkong		10a-Ulu Serendah	Forest near Orang Asli Village	Selangor	3.365343	101.615844	site coord. only
29-Sep-13	MW1307	bangkong		10a-Ulu Serendah	Forest near Orang Asli Village	Selangor	3.365343	101.615844	site coord. only
29-Sep-13	MW1308	bangkong		10a-Ulu Serendah	Forest near Orang Asli Village	Selangor	3.365343	101.615844	site coord. only
29-Sep-13	MW1309	bangkong		10a-Ulu Serendah	Forest near Orang Asli Village	Selangor	3.365343	101.615844	site coord. only
29-Sep-13	MW1310	bangkong		10a-Ulu Serendah	Forest near Orang Asli Village	Selangor	3.365343	101.615844	site coord. only
29-Sep-13	MW1311	bangkong		10a-Ulu Serendah	Forest near Orang Asli Village	Selangor	3.365343	101.615844	site coord. only
29-Sep-13	MW1312	bangkong		10a-Ulu Serendah	Forest near Orang Asli Village	Selangor	3.365343	101.615844	site coord. only
30-Sep-13	MW1313	bangkong		10d-Kota Damansara		Selangor	3.169178	101.580938	site coord. only
30-Sep-13	MW1314	bangkong		10d-Kota Damansara		Selangor	3.169178	101.580938	site coord. only
30-Sep-13	MW1315	bangkong		10d-Kota Damansara		Selangor	3.169178	101.580938	site coord. only
30-Sep-13	MW1316	bangkong		10d-Kota Damansara		Selangor	3.169178	101.580938	site coord. only
30-Sep-13	MW1317	bangkong		10d-Kota Damansara		Selangor	3.169178	101.580938	site coord. only
30-Sep-13	MW1318	bangkong		10d-Kota Damansara		Selangor	3.169178	101.580938	site coord. only

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
2-Oct-13	MW1401	cempedak	CH35	10f-UPM Ag Park	Teaching collection	Selangor	2.986531	101.708314	site coord. only
2-Oct-13	MW1402	cempedak	CH34	10f-UPM Ag Park	Teaching collection	Selangor	2.986531	101.708314	site coord. only
2-Oct-13	MW1403	cempedak	CH33	10f-UPM Ag Park	Teaching collection	Selangor	2.986531	101.708314	site coord. only
2-Oct-13	MW1404	cempedak	CH28	10f-UPM Ag Park	Teaching collection	Selangor	2.986531	101.708314	site coord. only
2-Oct-13	MW1405	cempedak	durian-cempedak	10f-UPM Ag Park	Teaching collection	Selangor	2.986531	101.708314	site coord. only
2-Oct-13	MW1406	cempedak	nangka-cemp	10f-UPM Ag Park	Teaching collection	Selangor	2.986531	101.708314	site coord. only
2-Oct-13	MW1407	cempedak	CH30	10f-UPM Ag Park Puchong	Teaching and commercial orchard	Selangor	2.982588	101.657621	site coord. only
10-Oct-13	MW1501	bangkong		5d-Sg Rual	Orang Asli Village	Kelantan	5.646200	101.783337	
10-Oct-13	MW1502	bangkong		5d-Sg Rual	Orang Asli Village	Kelantan	5.646280	101.783675	
10-Oct-13	MW1503	bangkong		5d-Sg Rual	Orang Asli Village	Kelantan	5.646245	101.783715	
10-Oct-13	MW1504	bangkong		5d-Sg Rual	Orang Asli Village	Kelantan	5.658742	101.798165	
10-Oct-13	MW1505	bangkong		5d-Sg Rual	Orang Asli Village	Kelantan	5.655603	101.796971	
10-Oct-13	MW1506	bangkong		5d-Sg Rual	Orang Asli Village	Kelantan	5.655009	101.797223	
7-Oct-13	MW1601	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1602	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1603	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1604	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1605	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1606	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1607	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1608	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1609	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1610	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1611	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
7-Oct-13	MW1612	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1613	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1614	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1615	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1616	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1617	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1618	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1619	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1620	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1621	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1622	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1623	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1624	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1625	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1626	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1627	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1628	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1629	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1630	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
7-Oct-13	MW1631	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1632	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
7-Oct-13	MW1633	bangkong		10e-Puchong (Ayer Hitam) FR		Selangor	3.011222	101.645658	site coord. only
13-Jun-13	NZ616	cempedak		14b-Sepilok Rainforest Discovery Center	evergreen lowland Dipterocarp Rainforest	Sabah	5.892000	117.939083	
14-Jun-13	NZ632	hybrid	nangka-cemp	14g-Ulu Dusun Ag Stn	Orchard. Grid 4.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ633	hybrid	nangka-cemp	14g-Ulu Dusun Ag Stn	Orchard. Grid 4.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ635	hybrid	nangka-cemp	14g-Ulu Dusun Ag Stn	Orchard. Grid 4.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ636	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 5.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ637	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 5.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ638	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 5.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ639	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 6.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ640	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 9.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ641	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 11.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ642	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 16.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ643	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 15.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ644	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 15.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ645	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 14.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ646	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 18.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ647	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 18.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ648	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 3.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ649	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 3.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ650	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 23.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ651	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 24.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ652	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 1.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ653	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 1.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ654	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 1.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ655	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 2.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ656	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 10.	Sabah	5.790544	117.775433	site coord. only
14-Jun-13	NZ657	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 10.	Sabah	5.790544	117.775433	site coord. only

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
14-Jun-13	NZ658	cempedak		14g-Ulu Dusun Ag Stn	Orchard. Grid 10.	Sabah	5.790544	117.775433	site coord. only
15-Jun-13	NZ689	cempedak		14f-Kebun Cina FR	Evergreen lowland Dipterocarp Rainforest. A portion of th area used to be a rubber planation up until the 1940s.	Sabah	5.840306	118.051222	site coord. only
15-Jun-13	NZ690	cempedak		14f-Kebun Cina FR	Evergreen lowland Dipterocarp Rainforest. A portion of th area used to be a rubber planation up until the 1940s.	Sabah	5.840306	118.051222	site coord. only
15-Jun-13	NZ691	cempedak		14f-Kebun Cina FR	Evergreen lowland Dipterocarp Rainforest. A portion of th area used to be a rubber planation up until the 1940s.	Sabah	5.840306	118.051222	site coord. only
17-Jun-13	NZ703	cempedak		14c-Sepilok Bed and Breakfast	Cultivated area.	Sabah	5.877806	117.946056	
17-Jun-13	NZ710	cempedak		14a-Kg Sg Batang	Cultivated area in a home garden	Sabah	5.935528	118.011528	site coord. only
17-Jun-13	NZ711	cempedak		14a-Kg Sg Batang	Cultivated area in a home garden	Sabah	5.935528	118.011528	site coord. only
17-Jun-13	NZ712	hybrid		14e-Sepilok Jungle Resort	Cultivated area.	Sabah	5.867244	117.950639	
18-Jun-13	NZ725	cempedak		14k-Sabah Forestry Inst	Cultivated area near lake	Sabah	5.629867	117.122647	
18-Jun-13	NZ732	cempedak		14i-Sg Wasai	Along roadside near forested area near the Wasai River in very red soil.	Sabah	5.604033	117.096833	
18-Jun-13	NZ738	cempedak		14j-Sg Telupid	Along roadside near forested area near the Wasai River in very red soil.	Sabah	5.676083	116.936800	
21-Jun-13	NZ785	cempedak		14h-Kg Kopozon	Village with cultivated and wild growing trees	Sabah	5.708333	116.016650	site coord. only

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
21-Jun-13	NZ786	cempedak		14h-Kg Kopozon	Village with cultivated and wild growing trees	Sabah	5.708333	116.016650	site coord. only
21-Jun-13	NZ787	cempedak		14h-Kg Kopozon	Village with cultivated and wild growing trees	Sabah	5.708333	116.016650	site coord. only
21-Jun-13	NZ792	cempedak		14h-Kg Kopozon	Village with cultivated and wild growing trees	Sabah	5.708333	116.016650	site coord. only
21-Jun-13	NZ793	cempedak		14h-Kg Kopozon	Village with cultivated and wild growing trees	Sabah	5.708333	116.016650	site coord. only
21-Jun-13	NZ812	hybrid		14k-Kipandi Butterfly Park	Along a managed evergreen lowland forest trail	Sabah	5.871167	116.250400	
23-Jun-13	NZ836	cempedak		14l-Beaufort Hill	Disturbed evergreen lowland forest	Sabah	5.346867	115.749150	
23-Jun-13	NZ843	cempedak		14l-Beaufort Hill	Disturbed evergreen lowland forest	Sabah	5.346650	115.750383	
23-Jun-13	NZ845	cempedak		14l-Beaufort Hill	Disturbed evergreen lowland forest	Sabah	c	115.749467	
23-Jun-13	NZ846	cempedak		14l-Beaufort Hill	Disturbed evergreen lowland forest	Sabah	5.346104	115.749521	
24-Jun-14	NZ850	cempedak		14l-Klias-Sianggau Road	In an overgrown area of a yard, on road between Klias FR and Sianggau FR	Sabah	5.248928	115.648897	
24-Jun-14	NZ854	cempedak		14n-Sianggau FR	Disturbed evergreen lowland forest	Sabah	5.179000	115.607350	site coord. only
24-Jun-14	NZ858	cempedak		14n-Sianggau FR	Disturbed evergreen lowland forest	Sabah	5.179000	115.607350	site coord. only
24-Jun-14	NZ859	cempedak		14n-Sianggau FR	Disturbed evergreen lowland forest	Sabah	5.179000	115.607350	site coord. only
24-Jun-14	NZ860	hybrid		14n-Sianggau FR	Disturbed evergreen lowland forest	Sabah	5.179000	115.607350	site coord. only
24-Jun-14	NZ861	cempedak		14n-Sianggau FR	Disturbed evergreen lowland forest	Sabah	5.179000	115.607350	site coord. only
24-Jun-14	NZ863	cempedak		14n-Sianggau FR	Disturbed evergreen lowland forest	Sabah	5.178283	115.607117	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
24-Jun-14	NZ865	cempedak		14n-Sianggau FR	Disturbed evergreen lowland forest	Sabah	5.175545	115.609596	
24-Jun-14	NZ869	cempedak		14n-Sianggau FR	Disturbed evergreen lowland forest	Sabah	5.171400	115.610933	
27-Jun-14	NZ918	cempedak		14m-Tenom Ag Park	Cultivated botanical garden	Sabah	5.173472	115.983700	
27-Jun-14	NZ941	cempedak		14m-Tenom Ag Park	Edge of hill forest near Artocarpus orchard	Sabah	5.186500	116.000444	site coord. only
27-Jun-14	NZ942	cempedak		14m-Tenom Ag Park	Edge of hill forest near Artocarpus orchard	Sabah	5.186500	116.000444	site coord. only
27-Jun-14	NZ943	cempedak		14m-Tenom Ag Park	Edge of hill forest near Artocarpus orchard	Sabah	5.186500	116.000444	site coord. only
1-Feb-02	NZ220	cempedak		13-Singapore Botanical Garden		Singapore	1.313883	103.815914	site coord. only
1-Feb-02	NZ222	cempedak		13-Singapore Botanical Garden		Singapore	1.313883	103.815914	site coord. only
year 2000	SBG3	cempedak		13-Singapore Botanical Garden		Singapore	1.313883	103.815914	site coord. only
1-Dec-12	YQL36	cempedak		13-Singapore Botanical Garden	Found in SBG Rainforest (M2-275)	Singapore	1.313883	103.815914	site coord. only
1-Mar-02	NZ227	bangkong		11a-Pasoh FR	Pasoh 50-ha plot.	Negeri Sembilan	2.982000	102.313000	
1-Mar-02	NZ236	bangkong		11b-Triang FR		Negeri Sembilan	2.978000	102.158639	
1-Mar-02	NZ260	cempedak		14d-Sepilok Arboretum		Sabah	5.873956	117.949599	
12-Mar-12	NZ466	cempedak		4-Kao Luk Lam Ru NP, Prachuap Khiri Khan	In a patch of forest between the road and beach	Thailand	8.623117	98.236350	site coord. only
12-Mar-12	NZ467	cempedak		4-Kao Luk Lam Ru NP, Prachuap Khiri Khan	In a patch of forest between the road and beach	Thailand	8.623117	98.236350	site coord. only
13-Mar-12	NZ476	cempedak		4-Khao Pra Taew Wildlife Center, Phuket	Growing near restaurant in park HQ.	Thailand	8.027940	98.362150	
13-Mar-12	NZ485 b	cempedak		4-Pattani	Peat swamp forest	Thailand	6.867800	101.250000	coord. of Pattani province

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
15-Mar-12	NZ509	cempedak		4-Surat Thani	Growing in a private yard.	Thailand	8.785210	98.699360	site coord. only
15-Mar-12	NZ510	cempedak		4-Surat Thani	Growing in a private yard.	Thailand	8.785210	98.699360	site coord. only
15-Mar-12	NZ511	cempedak		4-Surat Thani	Growing in a private yard.	Thailand	8.785210	98.699360	site coord. only
3-Jul-13	EG21	hybrid		1-The Kampong, Hissar estate, Miami	May be hybrid *A. integer x A. heterophyllus* based on poorly-defined annulus, elongate leaves, and elongate syncarps.	USA	25.716944	-80.251389	
16-Jul-13	EG55	cempedak		1-Fairchild Tropical Botanical Garden, Miami	Rare fruits pavillion.	USA	25.676080	-80.275836	
2-Oct-13	FRI78710	bangkong		9-Kuala Tahan	Taman Negara	Pahang	4.383333	102.400000	site coord. only
3-Oct-13	FRI78711	bangkong		9-Kuala Tahan	Taman Negara	Pahang	4.383333	102.400000	site coord. only
3-Oct-13	FRI78712	bangkong		9-Kuala Tahan	Taman Negara	Pahang	4.383333	102.400000	site coord. only
3-Oct-13	FRI78713	bangkong		9-Kuala Tahan	Taman Negara	Pahang	4.383333	102.400000	site coord. only
24-Jul-13	FRI79724	cempedak		12a-Gn Ledang		Johor	2.402000	102.590500	
20-Jan-13	CAS9	cempedak		2-Xishuangbanna Tropical Botanical Garden	No provenance information.	China	21.919603	101.277334	coord. of herbarium location
29-Jun-89	Gentry & La Frankie 66897 (GL66897)	bangkong		11a-Pasoh FR	Pasoh 50-ha plot.	Negeri Sembilan	3.000000	102.333300	site coord. only
7-Sep-87	Burley & Lee 246 (F.mus.int.1)	cempedak		16-Similajau FR	Lowland mixed Dipterocarp forest on yellow clay soil.	Sarawak	3.333333	113.416667	
9-Oct-97	Berau 865 (Leiden865)	cempedak		17-Berau, East Kalimantan	Logged-over forest. Indonesia, East Kalimantan, Berau Inhutani area, km 16, along trayek E	Indonesia	1.590000	117.150000	
12-Nov-90	J. Dransfield 6877 (Leiden6877)	cempedak		15-Ladan Hills FR	Disturbed lowland dipterocarp forest.	Brunei	4.583589	114.749829	

Date collected	Coll. No.	Taxa	Cultivar	Site	Site Notes	State/Country	lat (N)	long (N)	latlongNotes
25-Nov-82	Mogea 3665 (Kew3665)	cempedak		18-Tumbang Riang, Central Kalimantan	Primary rainforest.	Indonesia	-0.833333	112.833333	
1-Oct-33	Rahmat Si Boeea 5753 (NYBG1)	cempedak		20-Pargambiran, East Sumatra		Indonesia	2.723814	98.492788	coord. of Pargambiran
15-Sep-06	P. K. Loc 642 (NYBG2)	cempedak		3-Lai Thieu, Binh Duong	Orchard	Vietnam	10.936000	106.686000	
21-May-93	Jong 383 (NYBG3)	cempedak		19-Ngira, Noyan, West Kalimantan	Trailside rubber garden, Ngira.	Indonesia	0.687359	110.592711	coord. of Noyan district

Appendix 3. Measures of genetic diversity for all cempedak and bangkong populations.

Data for populations with N ≤ 4, which were not presented in Table 11, are in **bold**.
Cultivation status codes: cult = cultivated cempedak; silv = bangkong; uncult = wild cempedak; unk = unknown cempedak; hyb = hybrid cempedak-jackfruit.
Numbers in population code indicate state, and letters indicate locality within states. Please see Table 4 for full population names.
Allelic richness could not be calculated for populations N ≤ 4 due to missing data.
N = sample size; Na = no. of different alleles; Ne = no. of effective alleles; Ar = allelic richness; Pa = no. of private alleles; I = Shannon's Information Index; Ho = Observed heterozygosity; He = Expected heterozygosity; F = Fixation Index, [1 – (Ho/He)].

Cultivation status	Population Code	N	Na	Ne	Ar	Pa	I	Ho	He	F
cult	cult5b-BtPanau	14	3.1	1.9	2.1	0	0.737	0.401	0.424	0.141
cult	cult5c-FELDAKemahang	37	5.8	2.4	2.5	6	0.989	0.399	0.471	0.117
cult	cult6a-BtTanggaGermplasm	46	3.8	1.9	2.1	0	0.723	0.404	0.378	0.007
cult	cult7a-BalikPulau	12	3.4	2.0	2.3	0	0.788	0.355	0.423	0.194
cult	cult7b-Liew	18	2.3	1.5	1.6	1	0.449	0.380	0.277	-0.159
cult	cult7c-E&JFruitFarm	15	2.7	1.7	1.9	0	0.600	0.361	0.342	-0.015
cult	cult8a-Grik	5	2.5	2.2	2.2	1	0.680	0.517	0.398	-0.324
cult	cult8c-Bidor	13	2.7	1.6	1.8	0	0.517	0.305	0.289	0.058
cult	cult10c-SgBuloh	20	3.9	2.0	2.2	0	0.782	0.334	0.409	0.197
cult	cult10f-UPMAgPark	5	2.3	1.8	2.1	0	0.606	0.400	0.363	-0.047
cult	cult12d-KluangGermplasm	53	6.2	2.3	2.5	2	1.021	0.412	0.497	0.213
cult	cult12e-AyerHitamCDS	6	2.7	2.1	2.2	0	0.694	0.347	0.395	0.139
cult	cult14g-UluDusunAgStn	23	3.8	1.8	2.1	2	0.742	0.385	0.393	0.100
cult	cult14m-TenomAgPark	4	1.9	1.5	1.8	0	0.432	0.271	0.266	0.013
cult	**cult1-Miami**	**1**	**1.3**	**1.3**	**-**	**0**	**0.173**	**0.250**	**0.125**	**-1.000**
cult	**cult2-China**	**1**	**1.3**	**1.3**	**-**	**0**	**0.231**	**0.333**	**0.167**	**-1.000**
cult	**cult3-Vietnam**	**1**	**1.5**	**1.5**	**-**	**0**	**0.347**	**0.500**	**0.250**	**-1.000**
cult	**cult4-KhaoPraTaew**	**1**	**1.4**	**1.4**	**-**	**0**	**0.289**	**0.417**	**0.208**	**-1.000**
cult	**cult4-SuratThani**	**3**	**2.8**	**2.3**	**-**	**0**	**0.831**	**0.389**	**0.491**	**0.224**
cult	**cult5a-LubokStol**	**3**	**2.1**	**1.9**	**-**	**0**	**0.600**	**0.361**	**0.384**	**0.003**
cult	**cult5a-PasirMas**	**2**	**1.5**	**1.4**	**-**	**0**	**0.307**	**0.292**	**0.208**	**-0.387**

Cultivation status	Population Code	N	Na	Ne	Ar	Pa	I	Ho	He	F
cult	**cult5c-TanahMerah**	**2**	**1.9**	**1.7**	**-**	**0**	**0.498**	**0.500**	**0.313**	**-0.581**
cult	**cult5d-Jering**	**3**	**1.9**	**1.7**	**-**	**0**	**0.478**	**0.389**	**0.301**	**-0.304**
cult	**cult8b-UluGeroh**	**3**	**2.1**	**1.9**	**-**	**0**	**0.594**	**0.500**	**0.380**	**-0.302**
cult	**cult10b-FRIM-Kepong**	**1**	**1.4**	**1.4**	**-**	**1**	**0.289**	**0.417**	**0.208**	**-1.000**
cult	**cult10f-UPMAgParkPuchong**	**1**	**1.3**	**1.3**	**-**	**0**	**0.173**	**0.250**	**0.125**	**-1.000**
cult	**cult12f-KgSgSelangi**	**1**	**1.1**	**1.1**	**-**	**0**	**0.231**	**0.333**	**0.167**	**-1.000**
cult	**cult12g-KgSetiaJaya**	**2**	**1.7**	**1.5**	**-**	**0**	**0.383**	**0.292**	**0.250**	**-0.178**
cult	**cult13-SBG**	**3**	**2.2**	**1.9**	**-**	**0**	**0.573**	**0.417**	**0.343**	**-0.240**
cult	**cult14a-KgSgBatang**	**2**	**1.6**	**1.4**	**-**	**0**	**0.321**	**0.250**	**0.208**	**-0.244**
cult	**cult14c-SepilokBnB**	**1**	**1.3**	**1.3**	**-**	**0**	**0.231**	**0.333**	**0.167**	**-1.000**
cult	**cult14k-SabahForestryInst**	**1**	**1.4**	**1.4**	**-**	**0**	**0.289**	**0.417**	**0.208**	**-1.000**
cult	**cult14l-KliasRoad**	**1**	**1.2**	**1.2**	**-**	**0**	**0.116**	**0.167**	**0.083**	**-1.000**
silv	silv10a-UluSerendah	12	4.1	2.7	2.3	2	0.966	0.422	0.483	0.171
silv	silv10b-FRIM-Kepong	18	3.7	2.2	3.1	0	0.860	0.433	0.465	0.049
silv	silv10c-SgBuloh	29	4.8	2.4	2.0	1	0.894	0.365	0.440	0.167
silv	silv10d-KotaDamansara	6	2.0	1.8	2.4	0	0.463	0.346	0.272	-0.280
silv	silv10e-PuchongFR	33	4.4	2.3	2.4	2	0.846	0.352	0.425	0.127
silv	silv10f-MARDIFR	19	5.2	2.6	2.2	5	1.000	0.426	0.477	0.090
silv	silv11a-Pasoh	17	4.1	2.4	2.5	0	0.828	0.373	0.408	0.128
silv	silv12g-GnPulai	11	3.0	2.0	2.3	0	0.723	0.342	0.408	0.303
silv	silv5d-SgRual	6	3.6	2.6	2.3	3	0.981	0.465	0.534	0.124
silv	silv8b-UluGeroh	30	4.7	2.6	2.4	3	1.029	0.517	0.533	0.001
silv	silv9-KualaTahan	4	2.9	2.3	2.7	0	0.826	0.458	0.471	0.003
silv	**silv11b-Triang**	**1**	**1.4**	**1.4**	**-**	**1**	**0.289**	**0.417**	**0.208**	**-1.000**
silv	**silv12a-GnLedang**	**1**	**1.3**	**1.3**	**-**	**0**	**0.231**	**0.333**	**0.167**	**-1.000**
uncult	uncult14l-BeaufortHill	4	1.7	1.4	2.8	0	0.328	0.208	0.211	0.002
uncult	**uncult14b-SepilokRDC**	**1**	**1.5**	**1.5**	**-**	**0**	**0.347**	**0.500**	**0.250**	**-1.000**
uncult	**uncult14f-KebunCinaFR**	**3**	**1.4**	**1.3**	**-**	**0**	**0.239**	**0.250**	**0.162**	**-0.480**
uncult	**uncult16-Sarawak**	**1**	**1.1**	**1.1**	**-**	**0**	**0.058**	**0.083**	**0.042**	**-1.000**

Cultivation status	Population Code	N	Na	Ne	Ar	Pa	I	Ho	He	F
uncult	**uncult17-KalimantanEast**	**1**	**1.4**	**1.4**	**-**	**0**	**0.289**	**0.417**	**0.208**	**-1.000**
uncult	**uncult18-KalimantanCentral**	**1**	**1.3**	**1.3**	**-**	**0**	**0.231**	**0.333**	**0.167**	**-1.000**
unk	unk14h-KgKopozon	5	1.9	1.4	2.3	0	0.371	0.200	0.218	0.067
unk	unk14n-SianggauFR	7	1.8	1.4	2.6	0	0.357	0.214	0.222	0.021
unk	**unk11a-Pasoh**	**1**	**1.3**	**1.3**	**-**	**0**	**0.231**	**0.333**	**0.167**	**-1.000**
unk	**unk12b-LenggorFR**	**2**	**1.9**	**1.8**	**-**	**0**	**0.527**	**0.375**	**0.344**	**-0.083**
unk	**unk12c-GnBelumut**	**3**	**2.5**	**2.3**	**-**	**0**	**0.735**	**0.403**	**0.432**	**0.063**
unk	**unk12c-GnLambak**	**1**	**1.3**	**1.3**	**-**	**0**	**0.173**	**0.250**	**0.125**	**-1.000**
unk	**unk13-SBGforest**	**1**	**1.2**	**1.2**	**-**	**0**	**0.116**	**0.167**	**0.083**	**-1.000**
unk	**unk14d-SepilokArboretum**	**1**	**1.3**	**1.3**	**-**	**0**	**0.173**	**0.250**	**0.125**	**-1.000**
unk	**unk14i-SgWasai**	**1**	**1.2**	**1.2**	**-**	**0**	**0.116**	**0.167**	**0.083**	**-1.000**
unk	**unk14j-SgTelupid**	**1**	**1.5**	**1.5**	**-**	**0**	**0.347**	**0.500**	**0.250**	**-1.000**
unk	**unk15-Brunei**	**1**	**1.1**	**1.1**	**-**	**0**	**0.058**	**0.083**	**0.042**	**-1.000**
unk	**unk19-KalimantanWest**	**1**	**1.3**	**1.3**	**-**	**0**	**0.231**	**0.333**	**0.167**	**-1.000**
unk	**unk20-Sumatra**	**1**	**1.3**	**1.3**	**-**	**0**	**0.231**	**0.333**	**0.167**	**-1.000**
unk	**unk4-KaoLukLamRu**	**2**	**1.7**	**1.5**	**-**	**0**	**0.383**	**0.375**	**0.250**	**-0.511**
unk	**unk4-Pattani**	**1**	**1.3**	**1.3**	**-**	**1**	**0.231**	**0.333**	**0.167**	**-1.000**
hybcult	**hybcult10f-UPMAgPark**	**1**	**1.8**	**1.8**	**-**	**0**	**0.520**	**0.750**	**0.375**	**-1.000**
hybcult	**hybcult12d-KluangGermplasm**	**1**	**1.9**	**1.9**	**-**	**0**	**0.635**	**0.917**	**0.458**	**-1.000**
hybcult	**hybcult14e-SepilokJungleResort**	**1**	**1.4**	**1.4**	**-**	**1**	**0.347**	**0.500**	**0.250**	**-1.000**
hybcult	**hybcult14g-UluDusunAgStn**	**3**	**1.8**	**1.6**	**-**	**0**	**0.408**	**0.389**	**0.259**	**-0.514**
hybcult	**hybcult1-Miami**	**1**	**1.8**	**1.8**	**-**	**0**	**0.578**	**0.833**	**0.417**	**-1.000**
hybcult	**hybcult5a-Chekok**	**3**	**2.1**	**2.0**	**-**	**1**	**0.626**	**0.750**	**0.412**	**-0.851**
hybcult	**hybcult5a-PasirMas**	**1**	**1.6**	**1.6**	**-**	**0**	**0.404**	**0.583**	**0.292**	**-1.000**
hybcult	**hybcult5b-BtPanau**	**1**	**1.7**	**1.7**	**-**	**0**	**0.462**	**0.667**	**0.333**	**-1.000**
hybunk	**hybunk14k-KipandiButterflyPark**	**1**	**1.6**	**1.6**	**-**	**0**	**0.404**	**0.583**	**0.292**	**-1.000**
hybunk	**hybunk14n-SianggauFR**	**1**	**1.6**	**1.6**	**-**	**0**	**0.520**	**0.750**	**0.375**	**-1.000**

Appendix 4. Results of clonal assignment at different thresholds for samples sharing the same genotype with one or more other individuals.

Color codes indicate changes in clonal assignment results according to the threshold used (1-pink, 2-yellow, 4-tan, 6-blue, 8-purple, 10-gray). Numbers under the thresholds refer to the clonal genotype no. associated with the individual, while numbers after decimal points indicate the order in which individuals were assigned to that particular genotype. Proportion of membership to the Structure-inferred cempedak ("KC") and bangkong ("KB") genetic groups are indicated by "prop[KC]" and "prop[KB]" respectively. A threshold of zero indicates that two members of the same clone have no allelic differences.

			Genotype No. at Clonal Assignment Threshold (t)							Structure	
Coll no.	**Taxa/Cultivar**	**Site**	**t=0**	**t=1**	**t=2**	**t=4**	**t=6**	**t=8**	**t=10**	**prop[KC]**	**prop[KB]**
MW505	duri-cemp	8c-Bidor	1.1	1.1	1.1	**1.1**	1.01	1.01	1.01	0.97	0.03
MW507	duri-cemp	8c-Bidor	1.2	1.2	1.2	**1.2**	1.02	1.02	1.02	0.97	0.03
MW508	duri-cemp	8c-Bidor	1.3	1.3	1.3	**1.3**	1.03	1.03	1.03	0.97	0.03
MW509	duri-cemp	8c-Bidor	1.4	1.4	1.4	**1.4**	1.04	1.04	1.04	0.97	0.03
MW510	duri-cemp	8c-Bidor	1.5	1.5	1.5	**1.5**	1.05	1.05	1.05	0.97	0.03
MW351	cempedak	10c-Sg Buloh	1.6	1.6	1.6	**1.6**	1.06	1.06	1.06	0.97	0.03
NZ918	cempedak	14m-Tenom Ag Park	1.7	1.7	1.7	**1.7**	1.07	1.07	1.07	0.97	0.03
MW1405	duri-cemp a	10f-UPM Ag Park	18.1	18.1	1.8	**1.8**	1.08	1.08	1.08	0.96	0.04
MW514	duri-cemp b	8c-Bidor	18.2	18.2	1.9	**1.9**	1.09	1.09	1.09	0.96	0.04
MW638	cempedak	5c-FELDA Kemahang	25.1	25.1	25.1	**25.1**	1.10	1.10	1.10	0.98	0.02
MW708	cempedak	7a-Balik Pulau	25.2	25.2	25.2	**25.2**	1.11	1.11	1.11	0.98	0.03
MW367	cempedak	10c-Sg Buloh			32.1	**32.1**	1.12	1.12	1.12	0.98	0.02
MW1201	bangkong	12c-Gn Lambak			32.2	**32.2**	1.13	1.13	1.13	0.98	0.02
FMusint	cempedak	16-Similajau FR					1.14	1.14	1.14	0.96	0.04
MW639	cempedak	5c-FELDA Kemahang					1.15	1.15	1.15	0.98	0.02
MW732	E&J Orange	7c-E&J Fruit Farm					1.16	1.16	1.16	0.97	0.04
MW940	cempedak	12d-Kluang Germplasm					1.17	1.17	1.17	0.88	0.12
EG55	cempedak	1-Fairchild Tropical Botanical Garden, Miami						1.18	1.18	0.90	0.10
MW734	E&J Orange	7c-E&J Fruit Farm						1.19	1.19	0.97	0.03
NZ509	cempedak	4-Surat Thani						1.20	1.20	0.94	0.06
MW972	CH27	12e-Ayer Hitam CDS	2.1	2.1	2.1	**2.1**	2.1	2.1	2.1	0.65	0.35

			Genotype No. at Clonal Assignment Threshold (t)							Structure	
Coll no.	**Taxa/Cultivar**	**Site**	**t=0**	**t=1**	**t=2**	**t=4**	**t=6**	**t=8**	**t=10**	**prop[KC]**	**prop[KB]**
MW353	cempedak	10c-Sg Buloh	2.2	2.2	2.2	**2.2**	2.2	2.2	2.2	0.65	0.35
MW354	cempedak	10c-Sg Buloh	2.3	2.3	2.3	**2.3**	2.3	2.3	2.3	0.65	0.35
MW403	cempedak	5a-Pasir Mas	2.4	2.4	2.4	**2.4**	2.4	2.4	2.4	0.65	0.35
NZ647	cempedak	14g-Ulu Dusun Ag Stn	2.5	2.5	2.5	**2.5**	2.5	2.5	2.5	0.66	0.34
MW937	cempedak	12d-Kluang Germplasm			2.6	**2.6**	2.6	2.6	2.6	0.74	0.26
MW1205	bangkong	12b-Lenggor FR				**2.7**	2.7	2.7	2.7	0.51	0.49
MW352	cempedak	10c-Sg Buloh						2.8	2.8	0.47	0.53
NZ646	cempedak	14g-Ulu Dusun Ag Stn						2.9	2.9	0.57	0.43
MW814	Kuala Nerang	6a-Bt Tangga Germplasm	3.1	3.1	3.1	**3.1**	3.1	3.1	3.1	0.92	0.08
MW815	Kuala Nerang	6a-Bt Tangga Germplasm	3.2	3.2	3.2	**3.2**	3.2	3.2	3.2	0.92	0.08
MW816	Kuala Nerang	6a-Bt Tangga Germplasm	3.3	3.3	3.3	**3.3**	3.3	3.3	3.3	0.92	0.08
MW817	Bukit Selambau 2	6a-Bt Tangga Germplasm	3.4	3.4	3.4	**3.4**	3.4	3.4	3.4	0.92	0.08
MW903	cempedak	12d-Kluang Germplasm	3.5	3.5	3.5	**3.5**	3.5	3.5	3.5	0.93	0.07
MW813	Kuala Nerang	6a-Bt Tangga Germplasm			3.6	**3.6**	3.6	3.6	3.6	0.95	0.05
MW801	CH29	6a-Bt Tangga Germplasm	4.1	4.1	4.1	**4.1**	4.01	4.01	4.01	0.98	0.02
MW846	CH29	6a-Bt Tangga Germplasm	4.2	4.2	4.2	**4.2**	4.02	4.02	4.02	0.98	0.02
NZ653	cempedak	14g-Ulu Dusun Ag Stn	4.3	4.3	4.3	**4.3**	4.03	4.03	4.03	0.98	0.02
NZ654	cempedak	14g-Ulu Dusun Ag Stn	4.4	4.4	4.4	**4.4**	4.04	4.04	4.04	0.98	0.02
NZ655	cempedak	14g-Ulu Dusun Ag Stn	4.5	4.5	4.5	**4.5**	4.05	4.05	4.05	0.98	0.02
MW843	CH29	6a-Bt Tangga Germplasm	12.1	12.1	12.1	**12.1**	4.06	4.06	4.06	0.95	0.05
MW844	CH29	6a-Bt Tangga Germplasm	12.2	12.2	12.2	**12.2**	4.07	4.07	4.07	0.96	0.04
MW845	CH29	6a-Bt Tangga Germplasm	12.3	12.3	12.3	**12.3**	4.08	4.08	4.08	0.95	0.05
MW410	cempedak	5b-Bt Panau					4.09	4.09	4.09	0.95	0.05
NZ222	cempedak	13-Singapore Botanical Garden					4.10	4.10	4.10	0.91	0.09
MW1402	CH34 a	10f-UPM Ag Park	5.1	5.1	5.1	**5.1**	5.1	5.01	5.01	0.73	0.27
MW1401	CH35 a	10f-UPM Ag Park	5.2	5.2	5.2	**5.2**	5.2	5.02	5.02	0.74	0.26
MW358	cempedak	10c-Sg Buloh	5.3	5.3	5.3	**5.3**	5.3	5.03	5.03	0.74	0.26
MW601	cempedak	5c-Tanah Merah	5.4	5.4	5.4	**5.4**	5.4	5.04	5.04	0.74	0.26
NZ941	cempedak	14m-Tenom Ag Park	5.5	5.5	5.5	**5.5**	5.5	5.05	5.05	0.74	0.26

			Genotype No. at Clonal Assignment Threshold (t)							Structure	
Coll no.	**Taxa/Cultivar**	**Site**	**t=0**	**t=1**	**t=2**	**t=4**	**t=6**	**t=8**	**t=10**	**prop[KC]**	**prop[KB]**
MW742	E&J Orange	7c-E&J Fruit Farm			5.6	**5.6**	5.6	5.06	5.06	0.67	0.33
MW743	E&J Orange	7c-E&J Fruit Farm			5.7	**5.7**	5.7	5.07	5.07	0.70	0.30
MW921	cempedak	12d-Kluang Germplasm				**5.8**	5.8	5.08	5.08	0.74	0.26
MW744	E&J Orange	7c-E&J Fruit Farm						5.09	5.09	0.91	0.09
MW745	E&J Orange	7c-E&J Fruit Farm						5.10	5.10	0.89	0.11
MW973	CH34	12e-Ayer Hitam CDS						5.11	5.11	0.84	0.16
MW716	Liew	7b-Liew	6.1	6.1	6.1	**6.1**	6.01	6.01	6.01	0.98	0.02
MW718	Liew	7b-Liew	6.2	6.2	6.2	**6.2**	6.02	6.02	6.02	0.98	0.02
MW721	Liew	7b-Liew	6.3	6.3	6.3	**6.3**	6.03	6.03	6.03	0.98	0.02
MW726	Liew	7b-Liew	6.4	6.4	6.4	**6.4**	6.04	6.04	6.04	0.98	0.02
MW715	Liew	7b-Liew			6.5	**6.5**	6.05	6.05	6.05	0.97	0.03
MW727	Liew	7b-Liew	11.1	11.1	11.1	**11.1**	6.06	6.06	6.06	0.97	0.03
MW728	Liew	7b-Liew	11.2	11.2	11.2	**11.2**	6.07	6.07	6.07	0.97	0.03
MW729	Liew	7b-Liew	11.3	11.3	11.3	**11.3**	6.08	6.08	6.08	0.97	0.03
MW719	Liew	7b-Liew				**34.1**	6.09	6.09	6.09	0.97	0.03
MW720	Liew	7b-Liew				**34.2**	6.10	6.10	6.10	0.94	0.07
MW802	CH30	6a-Bt Tangga Germplasm	7.1	7.1	7.1	**7.1**	7.1	7.1	7.1	0.98	0.02
MW805	CH30	6a-Bt Tangga Germplasm	7.2	7.2	7.2	**7.2**	7.2	7.2	7.2	0.98	0.02
MW806	CH30	6a-Bt Tangga Germplasm	7.3	7.3	7.3	**7.3**	7.3	7.3	7.3	0.98	0.02
MW911	cempedak	12d-Kluang Germplasm	7.4	7.4	7.4	**7.4**	7.4	7.4	7.4	0.98	0.02
MW803	CH30	6a-Bt Tangga Germplasm			7.5	**7.5**	7.5	7.5	7.5	0.97	0.03
MW905	cempedak	12d-Kluang Germplasm			7.6	**7.6**	7.6	7.6	7.6	0.98	0.02
MW804	CH30	6a-Bt Tangga Germplasm				**7.7**	7.7	7.7	7.7	0.93	0.07
MW939	cempedak	12d-Kluang Germplasm				**7.8**	7.8	7.8	7.8	0.98	0.02
MW971	CH35	12e-Ayer Hitam CDS	8.1	8.1	8.1	**8.1**	8.1	8.01	8.01	0.96	0.04
MW975	CH35	12e-Ayer Hitam CDS	8.2	8.2	8.2	**8.2**	8.2	8.02	8.02	0.96	0.04
MW917	cempedak	12d-Kluang Germplasm	8.3	8.3	8.3	**8.3**	8.3	8.03	8.03	0.96	0.04
MW918	cempedak	12d-Kluang Germplasm	8.4	8.4	8.4	**8.4**	8.4	8.04	8.04	0.96	0.04
MW1407	CH30 a	10f-UPM Ag Park Puchong			8.5	**8.5**	8.5	8.05	8.05	0.96	0.04

			Genotype No. at Clonal Assignment Threshold (t)							Structure	
Coll no.	**Taxa/Cultivar**	**Site**	**t=0**	**t=1**	**t=2**	**t=4**	**t=6**	**t=8**	**t=10**	**prop[KC]**	**prop[KB]**
MW810	Bukit Selambau	6a-Bt Tangga Germplasm	16.1	16.1	16.1	**16.1**	16.1	8.06	8.06	0.98	0.02
MW812	Bukit Selambau	6a-Bt Tangga Germplasm	16.2	16.2	16.2	**16.2**	16.2	8.07	8.07	0.98	0.02
MW811	Bukit Selambau	6a-Bt Tangga Germplasm			16.3	**16.3**	16.3	8.08	8.08	0.98	0.02
MW909	cempedak	12d-Kluang Germplasm			16.4	**16.4**	16.4	8.09	8.09	0.98	0.02
MW954	cempedak	12d-Kluang Germplasm					16.5	8.10	8.10	0.93	0.07
MW821	Perak Champion	6a-Bt Tangga Germplasm						8.11	8.11	0.96	0.04
NZ648	cempedak	14g-Ulu Dusun Ag Stn	9.1	9.1	9.1	**9.1**	9.1	9.1	9.1	0.98	0.02
NZ649	cempedak	14g-Ulu Dusun Ag Stn	9.2	9.2	9.2	**9.2**	9.2	9.2	9.2	0.98	0.02
NZ656	cempedak	14g-Ulu Dusun Ag Stn	9.3	9.3	9.3	**9.3**	9.3	9.3	9.3	0.98	0.02
NZ657	cempedak	14g-Ulu Dusun Ag Stn	9.4	9.4	9.4	**9.4**	9.4	9.4	9.4	0.98	0.02
NZ658	cempedak	14g-Ulu Dusun Ag Stn					9.5	9.5	9.5	0.98	0.02
MW368	cempedak	10c-Sg Buloh						9.6	9.6	0.93	0.07
MW837	Berseri	6a-Bt Tangga Germplasm	10.1	10.1	10.1	**10.1**	10.1	10.1	10.1	0.66	0.35
MW838	Berseri	6a-Bt Tangga Germplasm	10.2	10.2	10.2	**10.2**	10.2	10.2	10.2	0.65	0.35
MW835	Pendang	6a-Bt Tangga Germplasm	10.3	10.3	10.3	**10.3**	10.3	10.3	10.3	0.65	0.35
MW836	Pendang	6a-Bt Tangga Germplasm	10.4	10.4	10.4	**10.4**	10.4	10.4	10.4	0.66	0.35
MW834	Pendang	6a-Bt Tangga Germplasm					10.5	10.5	10.5	0.70	0.30
MW826	Bayas	6a-Bt Tangga Germplasm	13.1	13.1	13.1	**13.1**	13.1	13.1	13.1	0.96	0.05
MW828	Bayas	6a-Bt Tangga Germplasm	13.2	13.2	13.2	**13.2**	13.2	13.2	13.2	0.96	0.04
MW839	Berseri*	6a-Bt Tangga Germplasm	13.3	13.3	13.3	**13.3**	13.3	13.3	13.3	0.95	0.05
MW820	Perak Champion	6a-Bt Tangga Germplasm	14.1	14.1	14.1	**14.1**	14.1	14.1	14.1	0.83	0.17
MW822	Bukit Manik	6a-Bt Tangga Germplasm	14.2	14.2	14.2	**14.2**	14.2	14.2	14.2	0.83	0.17
MW945	cempedak	12d-Kluang Germplasm	14.3	14.3	14.3	**14.3**	14.3	14.3	14.3	0.83	0.17
MW711	Liew	7b-Liew	15.1	15.1	15.1	**15.1**	15.1	15.1	15.1	0.98	0.02
MW713	Liew	7b-Liew	15.2	15.2	15.2	**15.2**	15.2	15.2	15.2	0.98	0.02
MW722	Liew	7b-Liew					15.3	15.3	15.3	0.98	0.02
MW624	cempedak	5c-FELDA Kemahang				**35.1**	35.1	35.1	**15.4**	0.98	0.02
MW630	cempedak	5c-FELDA Kemahang				**35.2**	35.2	35.2	**15.5**	0.97	0.03
MW611	cempedak	5c-FELDA Kemahang					35.3	35.3	**15.6**	0.95	0.06

			Genotype No. at Clonal Assignment Threshold (t)							Structure	
Coll no.	**Taxa/Cultivar**	**Site**	**t=0**	**t=1**	**t=2**	**t=4**	**t=6**	**t=8**	**t=10**	**prop[KC]**	**prop[KB]**
MW626	cempedak	5c-FELDA Kemahang					35.4	35.4	**15.7**	0.97	0.03
MW823	Bukit Manik	6a-Bt Tangga Germplasm	17.1	17.1	17.1	**17.1**	17.1	17.1	17.1	0.96	0.04
MW824	Bukit Manik	6a-Bt Tangga Germplasm	17.2	17.2	17.2	**17.2**	17.2	17.2	17.2	0.96	0.04
MW825	Bukit Manik	6a-Bt Tangga Germplasm				**17.3**	17.3	17.3	17.3	0.91	0.09
MW501	Red Ah Chai	8c-Bidor	19.1	19.1	19.1	**19.1**	19.1	19.1	19.1	0.90	0.10
MW421	cempedak	5b-Bt Panau	19.2	19.2	19.2	**19.2**	19.2	19.2	19.2	0.91	0.09
NZ643	cempedak	14g-Ulu Dusun Ag Stn		20.1	20.1	**20.1**	20.1	20.1	20.1	0.91	0.09
NZ644	cempedak	14g-Ulu Dusun Ag Stn		20.2	20.2	**20.2**	20.2	20.2	20.2	0.96	0.04
MW833	Pendang	6a-Bt Tangga Germplasm	21.1	21.1	21.1	**21.1**	21.1	21.1	21.1	0.92	0.08
MW831	Sik	6a-Bt Tangga Germplasm	21.2	21.2	21.2	**21.2**	21.2	21.2	21.2	0.92	0.08
MW830	Sik	6a-Bt Tangga Germplasm				**21.3**	21.3	21.3	21.3	0.89	0.11
MW832	Pendang	6a-Bt Tangga Germplasm					21.4	21.4	21.4	0.93	0.07
NZ467	cempedak	4-Kao Luk Lam Ru NP, Prachuap Khiri Khan						21.5	21.5	0.97	0.03
MW906	cempedak	12d-Kluang Germplasm	22.1	22.1	22.1	**22.1**	22.1	22.1	22.1	0.85	0.15
MW934	cempedak	12d-Kluang Germplasm	22.2	22.2	22.2	**22.2**	22.2	22.2	22.2	0.84	0.16
MW935	cempedak	12d-Kluang Germplasm			22.3	**22.3**	22.3	22.3	22.3	0.82	0.18
MW914	cempedak	12d-Kluang Germplasm	23.1	23.1	23.1	**23.1**	23.1	23.1	23.1	0.97	0.03
MW924	cempedak	12d-Kluang Germplasm	23.2	23.2	23.2	**23.2**	23.2	23.2	23.2	0.97	0.03
MW622	cempedak	5c-FELDA Kemahang	24.1	24.1	24.1	**24.1**	24.1	24.1	24.1	0.75	0.25
MW629	cempedak	5c-FELDA Kemahang	24.2	24.2	24.2	**24.2**	24.2	24.2	24.2	0.74	0.26
MW620	cempedak	5c-FELDA Kemahang				**24.3**	24.3	24.3	24.3	0.74	0.26
MW1601	bangkong	10e-Puchong (Ayer Hitam) FR	26.1	26.1	26.1	**26.1**	26.1	26.1	26.1	0.32	0.68
MW1605	bangkong	10e-Puchong (Ayer Hitam) FR	26.2	26.2	26.2	**26.2**	26.2	26.2	26.2	0.31	0.69
MW1606	bangkong	10e-Puchong (Ayer Hitam) FR						26.3	26.3	0.18	0.82
MW1625	bangkong	10e-Puchong (Ayer Hitam) FR			27.1	**27.1**	27.1	27.1	27.1	0.04	0.97
MW1626	bangkong	10e-Puchong (Ayer Hitam) FR			27.2	**27.2**	27.2	27.2	27.2	0.04	0.96
NZ942	cempedak	14m-Tenom Ag Park			28.1	**28.1**	28.1	28.1	28.1	0.98	0.02
NZ943	cempedak	14m-Tenom Ag Park			28.2	**28.2**	28.2	28.2	28.2	0.98	0.02
MW731	E&J Orange	7c-E&J Fruit Farm						28.3	28.3	0.98	0.02

			Genotype No. at Clonal Assignment Threshold (t)							Structure	
Coll no.	**Taxa/Cultivar**	**Site**	**t=0**	**t=1**	**t=2**	**t=4**	**t=6**	**t=8**	**t=10**	**prop[KC]**	**prop[KB]**
NZ636	cempedak	14g-Ulu Dusun Ag Stn			29.1	**29.1**	29.1	29.1	29.1	0.88	0.12
NZ637	cempedak	14g-Ulu Dusun Ag Stn			29.2	**29.2**	29.2	29.2	29.2	0.96	0.04
NZ689	cempedak	14f-Kebun Cina FR			30.1	**30.1**	30.1	30.1	30.1	0.97	0.03
NZ690	cempedak	14f-Kebun Cina FR			30.2	**30.2**	30.2	30.2	30.2	0.97	0.03
MW922	cempedak	12d-Kluang Germplasm			31.1	**31.1**	31.1	31.1	31.1	0.70	0.30
MW953	CH26 Paya Jaras	12d-Kluang Germplasm			31.2	**31.2**	31.2	31.2	31.2	0.75	0.25
MW608	cempedak	5c-FELDA Kemahang			33.1	**33.1**	33.1	33.1	33.1	0.80	0.20
MW714	cempedak	7a-Balik Pulau			33.2	**33.2**	33.2	33.2	33.2	0.84	0.16
MW420	cempedak	5b-Bt Panau				**36.1**	36.1	36.1	36.1	0.98	0.02
NZ863	cempedak	14n-Sianggau FR				**36.2**	36.2	36.2	36.2	0.98	0.03
MW703	cempedak	7a-Balik Pulau						36.3	36.3	0.97	0.03
NZ859	cempedak	14n-Sianggau FR						36.4	36.4	0.97	0.03
MW809	Bukit Selambau	6a-Bt Tangga Germplasm					37.1	37.1	37.1	0.93	0.07
MW901	cempedak	12d-Kluang Germplasm					37.2	37.2	37.2	0.93	0.07
MW807	Jeneri	6a-Bt Tangga Germplasm						37.3	37.3	0.96	0.04
MW808	Jeneri*	6a-Bt Tangga Germplasm						37.4	37.4	0.93	0.07
MW1220	cempedak	12g-Setia Jaya					38.1	38.1	38.1	0.96	0.04
NZ220	cempedak	13-Singapore Botanical Garden					38.2	38.2	38.2	0.93	0.07
NZ738	cempedak	14j-Sg Telupid					39.1	39.1	39.1	0.98	0.03
NZ792	cempedak	14h-Kg Kopozon					39.2	39.2	39.2	0.97	0.03
MW606	cempedak	5c-FELDA Kemahang						39.3	39.3	0.83	0.17
MW614	cempedak	5c-FELDA Kemahang						39.4	39.4	0.82	0.18
MW818	Bukit Selambau 2	6a-Bt Tangga Germplasm						40.1	40.1	0.90	0.11
MW819	Bukit Selambau 2	6a-Bt Tangga Germplasm						40.2	40.2	0.80	0.20
MW615	cempedak	5c-FELDA Kemahang						41.1	41.1	0.82	0.18
MW637	cempedak	5c-FELDA Kemahang						41.2	41.2	0.69	0.31
MW733	E&J Orange	7c-E&J Fruit Farm						42.1	42.1	0.97	0.03
MW739	E&J Orange	7c-E&J Fruit Farm						42.2	42.2	0.94	0.06
MW1406	nangka-cemp	10f-UPM Ag Park						43.1	43.1	0.24	0.76

			Genotype No. at Clonal Assignment Threshold (t)							Structure	
Coll no.	**Taxa/Cultivar**	**Site**	**t=0**	**t=1**	**t=2**	**t=4**	**t=6**	**t=8**	**t=10**	**prop[KC]**	**prop[KB]**
MW424	nangka-cemp	5b-Bt Panau						43.2	43.2	0.26	0.74
MW412	cempedak	5b-Bt Panau						44.1	44.1	0.93	0.07
NZ854	cempedak	14n-Sianggau FR						44.2	44.2	0.98	0.02
MW1612	bangkong	10e-Puchong (Ayer Hitam) FR						45.1	45.1	0.32	0.68
MW1633	bangkong	10e-Puchong (Ayer Hitam) FR						45.2	45.2	0.54	0.46

Appendix 5. Summary of leaf color data.

Colors were recorded as codes from mini RHS color chart (Flower Council of Holland, UK). Colors that were recorded only rarely are not shown on this table. "bang" = bangkong; "cemp" = cempedak.

Underside leaf color			Midrib color			Withering leaf color (upperside)			Withering leaf color (bottom-side)		
RHS code	***bang***	***cemp***	**RHS code**	***bang***	***cemp***	**RHS code**	***bang***	***cemp***	**RHS code**	***bang***	***cemp***
146c	55	220	**146c**	2	13	**n199c**	58	66	**n167a**	60	125
137c	72	44	**152d**	33	87	**165a**	29	55	**n199c**	36	21
144a	1	0	**152b**	19	34	**n170a**	13	32	**n170a**	4	33
145c	1	0	**145a**	3	27	**n167a**	11	31	**165a**	7	25
194a	3	0	**153c**	1	18	**mixed colors**	5	26	**163a**	6	22
145a	0	2	**145b**	8	5	**166a**	3	19	**163b**	1	9
152b	0	2	**145c**	6	7	**200c**	10	5	**n167a+163a**	0	7
136c	0	1	**151d**	0	10	**163a**	1	6	**167a**	1	3
Total	**135**	**269**	**Total**	**132**	**265**	**Total**	**141**	**261**	**Total**	**127**	**258**

Appendix 6. Summary of hairiness data for several leaf and stem parts.

"bang" = bangkong; "cemp" = cempedak. Asterisks indicate when the type of glabrousness or hairiness look different from other samples (for example, some samples had short, stiff hairs instead of long wiry hairs typical of cempedak/bangkong).

Plant part	leaf edge		leaf underside		petiole		veins		upperside midrib		node		stipule		twig	
Hairiness	***bang***	***cemp***	***bang***	***cemp***	***bang***	***cemp***	***bang***	***cemp***	***bang***	***cemp***	***bang***	***cemp***	***bang***	***cemp***	***bang***	***cemp***
glabrous	31	23	98	113	10	10	84	21	29	25	8	11	0	1	48	52
medium	14	42	2	18	17	4	10	20	0	0	14	3	0	0	23	9
med-glabrous	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1
hairy	104	188	22	128	121	237	53	223	77	62	128	251	94	29	72	189
very hairy	2	0	2	3	3	15	4	0	0	2	1	2	0	3	8	16
glabrous*	0	3	0	3	0	0	0	0	0	0	0	0	0	0	0	0
hairy*	1	1	2	2	1	0	1	3	3	3	0	0	0	0	0	0
Total	**152**	**257**	**126**	**267**	**152**	**267**	**152**	**267**	**109**	**93**	**151**	**267**	**94**	**33**	**151**	**267**

Appendix 7. Images of cempedak and bangkong.

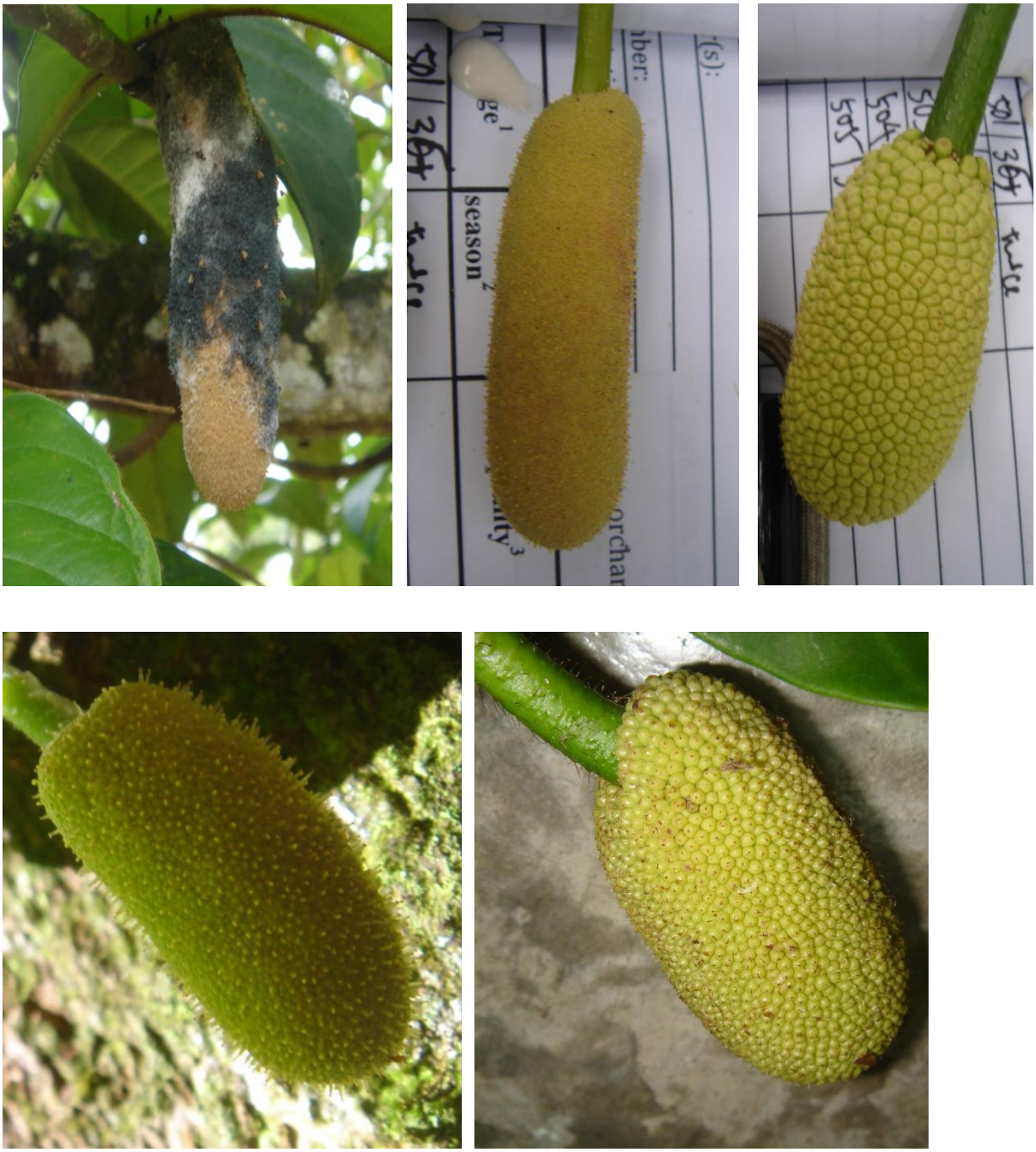
All photos are by Maria Wang, except for cempedak seeds which were from http://cempakabiru-nieda.blogspot.com/2012/01/biji-cempedak-rebus.html. All photos are not to scale.

1. Cempedak (left) and bangkong (right) tree with fruits.

2. Male and female inflorescences in cempedak and bangkong. Photos are not to scale.

Top panel: Cempedak inflorescences. From left to right: male inflorescence with fungus; male inflorescence at anthesis; female inflorescence at anthesis.

Bottom panel: Bangkong inflorescences. From left to right: male inflorescence at anthesis; female inforescence at anthesis.



3. Variety of fruit shapes and appearances in cempedak cutivars. Photos are not to scale.

A. Perak Champion; B. Durian-cempedak; C. CH34; D. Jeneri; E. Nangka-cempedak.

4. Variation of pulp color in cempedak.

A. “white” (fruit courtesy of Mr. Gan); B. pale yellow (Bukit Selambau cultivar);
C. bright yellow or yellow-orange (CH30); D. pale yellow-orange (unknown variety from FELDA Kemahang, grown from seeds from Thailand); E. bright orange (from E&J Fruit Farm);
F. orange-pink (Liew variety).

5. Variety in bangkong fruit shapes (all from Sungai Rual, Kelantan). All were unripe.

6. Longitudinal sections of ripe cempedak (top left), unripe cempedak (top right), and bangkong fruit (bottom; both unripe).

7. Cempedak and bangkong seeds. After boiling, the seeds taste like potato or chestnuts.

Top: Boiled cempedak seeds. Picture from http://cempakabiru-nieda.blogspot.com/2012/01/biji-cempedak-rebus.html

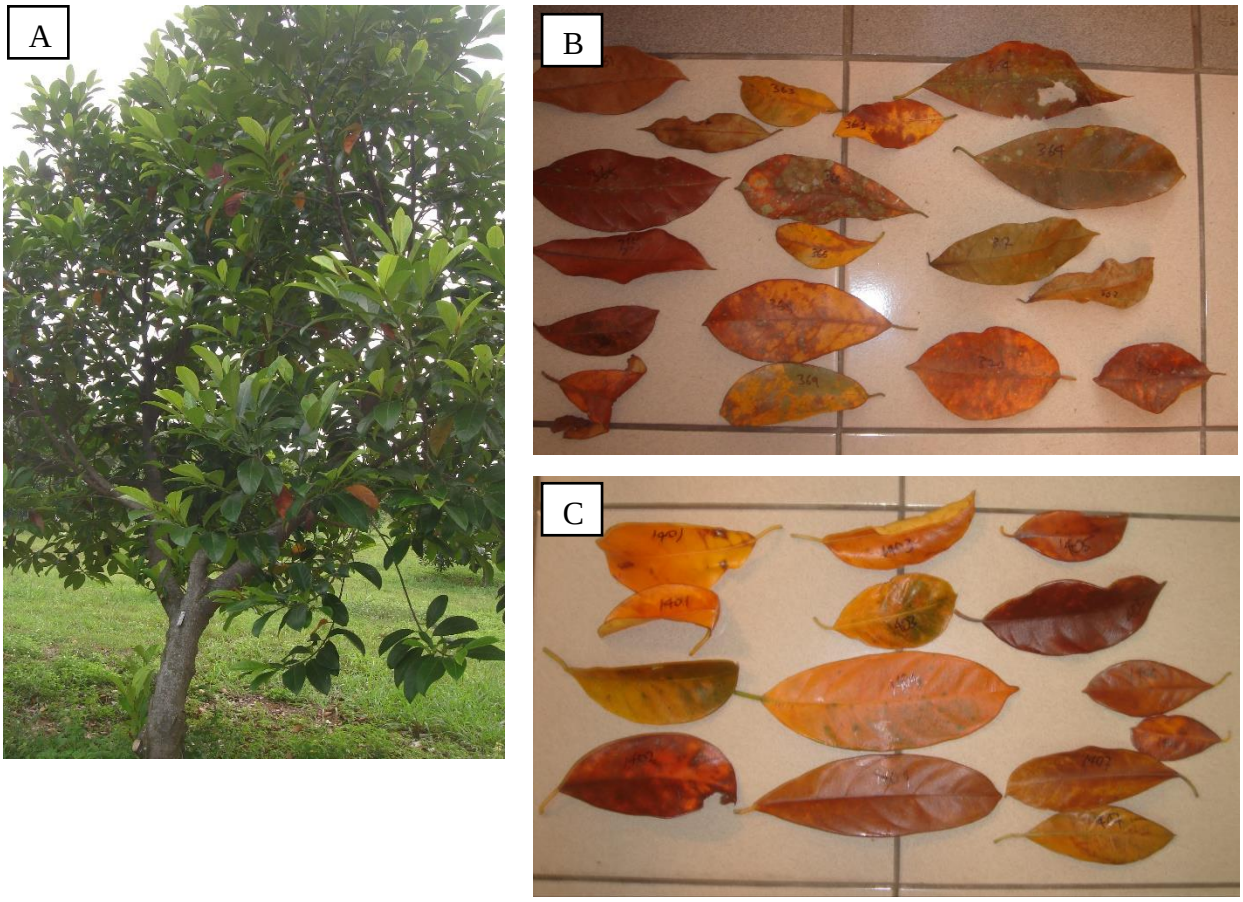
Bottom: Raw bangkong seeds from unripe fruit.

8. “Autumn tints” in withering leaves of cempedak and bangkong trees. Photos are not to scale.

One of the differences between cempedak and bangkong that Corner (1938) observed was the lack of “autumn tints”, or rich orange-ochre to brown-ochre colors in withering leaves of bangkong trees. From my observations, “autumn tints” do occur in bangkong leaves, but less frequently than cempedak. This may be partly because it was more difficult to detect withering leaves on the bangkong tree when they are growing in a leafy, dense forest.

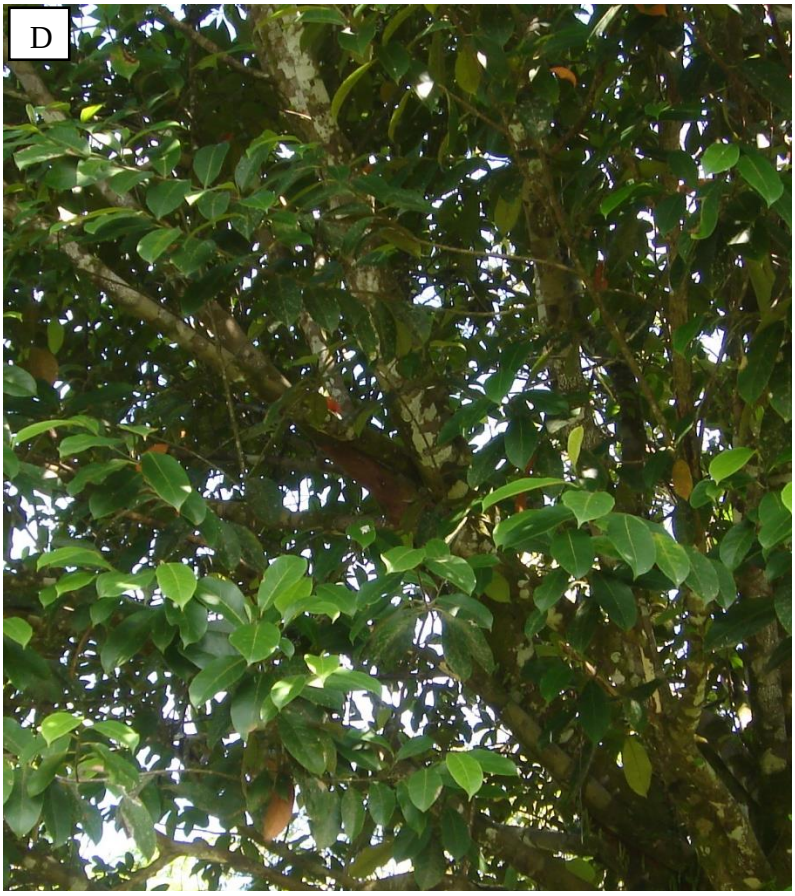
A. Withering leaves with autumn tints on cempedak tree (left).

B-C. Variation in color of withering leaves of cempedak trees sampled from Sungai Buloh (top) and UPM Agro-Park (bottom).



D. Withering leaves with autumn tints on bangkong tree, which may be difficult to detect.

E-G. Variation in color of withering leaves of bangkong trees from FRIM Kepong (top and middle) and Kota Damansara FR (bottom).



9. Morphology of nangka-cempedak (jackfruit-cempedak) hybrid cultivar. Photos are not to scale.

Top left: Nangka-cempedak tree.
Top right: Fruits on the tree.
Bottom: Close-up of leaves of nangka-cempedak (left) and cempedak (right). Compared to cempedak which typically has long wiry hairs on the leaves/twigs/stipules, nangka-cempedak has stiffer and shinier leaves, and glabrous or with short rough hairs on the leaves/twigs/stipules.

Nangka-cempedak

Cempedak