Pertanika J. Trop. Agric. Sc. 42 (3): 921 - 937 (2019)

PERTANIKA JOURNALS

TROPICAL AGRICULTURAL SCIENCE

Journal homepage: http://www.pertanika.upm.edu.my/

Assessment of IRAP Markers to Evaluate the Genetic Diversity of *Eurycoma longifolia*

Wook Nor Fadilah[1,3]*, Osman Mohamad[2], Abdullah Mohd Zaki[1] and Abdullah Shamsiah[3]

[1]*Plant Improvement Programme of Forest Biotechnology Division, Forest Research Institute Malaysia (FRIM), 52109 Kepong, Selangor, Malaysia*
[2]*Malaysian Industry-Government Group for High Technology (MiGHT), Jalan Impact, 63000 Cyberjaya, Selangor, Malaysia*
[3]*Faculty of Plantation and Agrotechnology, University Technology MARA (UiTM), 40450 Shah Alam, Selangor, Malaysia*

ABSTRACT

Eurycoma longifolia, among the more precious medicinal herbs, has many promising benefits. However, the limited information on genetic diversity of the species has hindered many further related studies such as breeding programme. This study aims to evaluate the transferability of retrotransposon-based markers, Inter-Retrotransposon Amplified Polymorphism (IRAP) to evaluate the genetic diversity within and between two selected provenances. Thirteen IRAP markers were initially screened for their transferability properties. Findings showed that retrotransposon elements were present in the genome of *E. longifolia*. Four single primers and two combinations of primers were effectively utilized to analyse the genetic diversity of *E. longifolia*. Genetic diversity assessment indicated that the total diversity (H_T) was 0.2396 and the diversity within the population (H_S) was 0.2233. The coefficient of gene differentiation (G_{ST}) was 0.0680, signifying that there were 6.8% total genetic variations between the provenances and 93.2% variations among individual accessions within the provenances. Thus, the genetic variations between selected provenances were lower than the genetic variations within the provenances.

Keywords: Genetic variations, molecular marker, retrotransposon, transferability

ARTICLE INFO

Article history:
Received: 21 February 2019
Accepted: 13 May 2019
Published: 19 August 2019

E-mail addresses:
norfadilah@frim.gov.my (Wook Nor Fadilah)
mbopar2004@yahoo.com (Osman Mohamad)
zaky@frim.gov.my (Abdullah Mohd Zaki)
shamsiah3938@uitm.edu.my (Abdullah Shamsiah)
* Corresponding author

INTRODUCTION

Eurycoma longifolia or locally known as "Tongkat Ali" is the most prominent herb species in Malaysia, currently receiving great attention from the government and

ISSN: 1511-3701
e-ISSN: 2231-8542

people due to its medicinal importance. *E. longifolia* has also been listed as the top five herbs of focus in the National Key Economic Areas (NKEA) in Malaysia (Performance Management and Delivery Unit [PEMANDU], 2013). Many studies have been conducted on the chemical profiling and pharmacological importance such as aphrodisiac (Ang et al., 2000; Low et al., 2013; Mohd Azmi et al., 2004; Subramaniam, 2013), antimalarial (Kuo et al., 2004), and anticancer (Kuo et al., 2004; Wong et al., 2012). Therefore, the market demand for *E. longifolia* in the Malaysian herbs industries is very high.

Nowadays, most of the raw materials of *E. longifolia* needed for the industries especially the root parts are obtained from the natural forest reserve areas. This practice, in the long run, could endanger the plant species diversity and be economically impractical. Thus, in order to cater the problem of providing raw materials of *E. longifolia*, the Forest Research Institute Malaysia (FRIM) (n. d.) has taken an initiative to conduct a research on "Provenance Trial of *E. longifolia*" via a breeding strategy research to provide important breeding information and conservation strategies. In forestry, the term provenance refers to the original place where the plant samples were collected.

The provenance trial is conceptually defined as the selection on a population level with a special type of plantation to understand the well-adapted organisms to various environmental conditions through genetic adaptation and phenotypic plasticity. Provenance variation can be found in any plant species that occupies a broad range of geographic conditions. The variations or differences in environmental conditions result in the adaptation of the plant species to the environment, consequently changing the population's genetic pool. Findings in the provenance trial would provide the best match between specific environmental conditions and the plant species of various origins (Hettasch, 2002).

Provenance trial research has included a wide range of sampling activities of *E. longifolia* throughout Peninsular Malaysia, establishment of the plantation trial to monitor the growth performances and the establishment of provenance resource stands for conservation strategies (Nor Fadilah et al., 2016), screening for significant phytochemical content (Mohd Zaki et al., 2015a, 2015b) and plant pathology study (Wan-Muhammad-Azrul et al., 2018). Breeding strategies of any plant species are improved by complimenting them with the knowledge of the genetic variation at the molecular level.

Molecular markers are used to evaluate the genetic variation within and between provenances. They have proven to provide a precise assessment and identification of plant cultivars (Tnah et al., 2011) compared to biochemical and phenotypic markers which are negatively affected by low polymorphisms (Kalendar et al., 1999). The DNA profiling and genetic information of plant cultivars are very important for the plant improvement programs such as germplasm and conservation management. The best molecular markers have abundant

genomic sequence information, high polymorphisms, co-dominant markers to differentiate between homozygotes and heterozygotes, and easily reproducible technique (Tnah et al., 2011).

Therefore, due to the limited information on the genetic diversity of *E. longifolia* in the study of provenance trial, this study aims to evaluate the transferability of retrotransposon-based markers, Inter-Retrotransposon Amplified Polymorphism (IRAP) to analyze the genetic diversity within and between two selected provenances. IRAP relies on PCR amplifications between two retrotransposon insertion sites (Kalendar et al., 1999). The high copy number of retrotransposon elements makes them abundant and prevalent in plant genomes, and gives a strong basis as a marker system. Furthermore, its structure and replication strategy are advantageous as markers. Additionally, this simpler technique uses fewer numbers of markers but has high polymorphism detection, high reproducibility, and high reliability (Guo et al., 2006; Yuying et al., 2011).

MATERIALS AND METHODS

Plant Materials

A total of 56 accessions of *E. longifolia* were collected as test materials; 30 accessions were from Forest Reserve of Kapas Island (FRPK) located in the state of Terengganu, and 26 accessions were from Forest Reserve of Maokil (FRMJ) located in the state of Johor. The provenances for this study were selected based on the promising growth performances in the provenance trial study (data not shown). The GPS coordinates of every mother plants found during the sampling activities were recorded (data not shown). The characteristics of the selected provenances are shown in Table 1.

Table 1
Origin of the provenances (based on sampling locations) characterized by latitude, longitude, altitude and forest type

Provenances	Latitude (N°)	Longitude (E°)	Altitude (m)	Forest type
Forest Reserve of Maokil, Johor	2°01' - 2°12'	102°49' - 103°21'	94 – 228	Hill forest
Forest Reserve of Kapas Island, Terengganu	4°35' - 5°12'	103°15' - 103°16'	30 - 169	Hill forest (Island)

DNA Extraction, Purification and Quantitation

Genomic DNA (gDNA) was extracted from leaves of each accession using a modified CTAB method (Murray & Thompson, 1980). The gDNA stock was stored at -80˚C. The purification of total gDNA was conducted following the Roche Diagnostic (2008) standard procedure: High Pure PCR Template Preparation Kit. Then, the gDNA and purified gDNA were quantified by NanoDrop 2000 system, UV-VIS Spectrophotometer (Thermo Fisher Scientific). The gDNA and purified DNA

were measured at 260 nm, with optical density at 600 nm. The purity of the genomic DNA and purified DNA was calculated at ratios 260/280 nm and 260/230 nm.

Additionally, the quality of extracted gDNA and purified DNA was evaluated by running a gel electrophoresis on 0.8% w/v agarose at 100 volt and 300 mA for 20 minutes. The DNA calf thymus (at the concentration of 5, 10, 15, 25, and 50 ng/µl) was used as a standard marker for the evaluation. The preparation and dilution of the template DNA were conducted based on both results of quantification by NanoDrop 2000 system and bands intensity observed on the agarose gel.

Analysis of IRAP Markers

PCR amplifications were performed in 25 µl total volume consisted of approximately 100 ng genomic DNA, 5X Green GoTaq® Flexi PCR Buffer (Promega Corporation), 200 µM dNTP (Promega Corporation), 3.0 mM $MgCl_2$ (Promega Corporation), 0.8 µM of each primer and 0.05 U of Taq DNA polymerase (Promega Corporation). The PCR (Applied Biosystems® 2720 Thermal Cycler) cycling conditions were as followed: 95°C for 5 min; followed by 35 cycles of 95°C for 1 min, 42°C for 1 min (T_a depended on primers used) and 72°C for 1 min; and the final extension step at 72°C for 10 minutes. PCR products were separated in 1.0% agarose gel in 1X TAE buffer for 90 minutes at 60V and 400mA (Agisimanto et al., 2008). The characteristics of retrotransposon-based primers are shown in Table 2.

Data Analysis

For each IRAP fragment, the presence or absence was scored on gel images and binary matrices were assembled in Microsoft Excel spreadsheets. Descriptive statistics including the Number of Scored Band (NSB), Number of Polymorphic Band (NPB), Percentage of Polymorphic Band (PPB), Polymorphism Information Content (PIC), Effective Multiplex Ratio (EMR), Marker Index (MI), and Heterozygosity (*He*) were calculated to determine the informativeness of the marker system.

Assessment of the genetic diversity parameters such as percentage (%) of polymorphism, Nei's genetic diversity (*He*), Shannon's information index (*I*), the observed number of alleles (*Na*), and the effective number of alleles (*Ne*) was done using POPGENE software version 1.32 (Yeh et al., 1999). Determination of genetic structure was also carried out using POPGENE software based on the Nei's method. Total gene diversity (H_T), genetic diversity within the population (H_{TS}), the coefficient of gene differentiation (G_{ST}) and gene flow (N_m) were analyzed.

NTSYS-pc version 2.2 (Rohlf, 2009) was used to perform the cluster analysis. SIMQUAL option was chosen to calculate the genetic similarities based on Jaccard's coefficient. Unweighted Pair Group Method with Arithmetic Average (UPGMA) was employed to construct the dendogram using SAHN module. Graphical 3D image of Principal Component Analysis (PCA) representing genetic diversity was generated using the EIGEN and MOD-3D programs of this software.

Table 2
Characteristics of IRAP primers used for amplifications

No	Primers' name	Retrotransposon name and orientation	Sequence (5'-3')	Tm (°C)	Reference/Source
1.	Copia-F	Ty1 Copia - Forward	5' ACNGCNTTYYTNCAYGG 3'	41.9	Flavell et al. (as cited in Agisimanto et al., 2008, p. 247)
2.	Copia-R	Ty1 copia - Reverse	5' ARCATRTCRTCNACRTA 3'	35.9	Flavell et al. (as cited in Agisimanto et al., 2008, p. 247)
3.	Sabrina-C0945	Sabrina - Forward	5' GCAAGCTTCCGTTTCCGC 3'	47.5	Leigh et al., 2003
4.	Sukkula-9900	Sukkula - Forward	5' GATAGGGTCGCATCTTGGGCGTGAC 3'	57.5	Leigh et al., 2003
5.	Sukkula-91673	Sukkula - Forward	5' TGTGACAGCCCGATGCCGACGTTCC 3'	59.1	Leigh et al., 2003
6.	Sukkula-E0228	Sukkula - Reverse	5' GGAACGTCGGCATCGGGCTG 3'	54.9	Leigh et al., 2003
7.	Nikita-57	Nikita - Forward	5' CGCATTTGTTCAAGCCTAAACC 3'	47.9	Leigh et al., 2003
8.	Nikita-E2647	Nikita - Forward	5' ACCCCTCTAGGCGACATCC 3'	50.3	Leigh et al., 2003
9.	LTR 6150	BARE-1 - Reverse	5' CTGGTTCGGCCCATGTCTATATGTATCCACACATGTA 3'	60.5	Kalendar et al., 1999
10.	LTR 6149	BARE-1 - Forward	5' CTCGCTCGCCCACTACATCAACCGCGCGTTTATT 3'	60.6	Kalendar et al., 1999
11.	3 LTR	BARE-1 - Forward	5' TGTTTCCCATGCGACGTTCCCCAACA 3'	56.0	Teo et al., 2005
12.	5' LTR 1	BARE-1 - Reverse	5' TTGCCTCTAGGGCATATTTCCAACA 3'	50.9	Teo et al., 2005
13.	5' LTR 2	BARE-1 - Reverse	5' ATCATTCCCTCTAGGGCATAAATTC 3'	48.9	Teo et al., 2005

Note: Tm = melting temperature

RESULTS AND DISCUSSION

Informativeness of the IRAP Markers

Initially, 13 IRAP primers were screened for their transferability for the genetic diversity study of *E. longifolia*. However, only four primers; Copia-F (Figure 1), Nikita-57 (Figure 2), Sabrina-C0945 (Figure 3) and Sukkula-9900 (Figure 4) out of the 13 tested primers yielded polymorphic banding patterns with high resolutions and reproducible bands. Additionally, the two combinations of primers [(Sukkula-9900 + Copia-F) and (Sukkula-9900 + Sabrina-C0945)] successfully yielded unique polymorphic banding patterns out of six combinations primers tested.

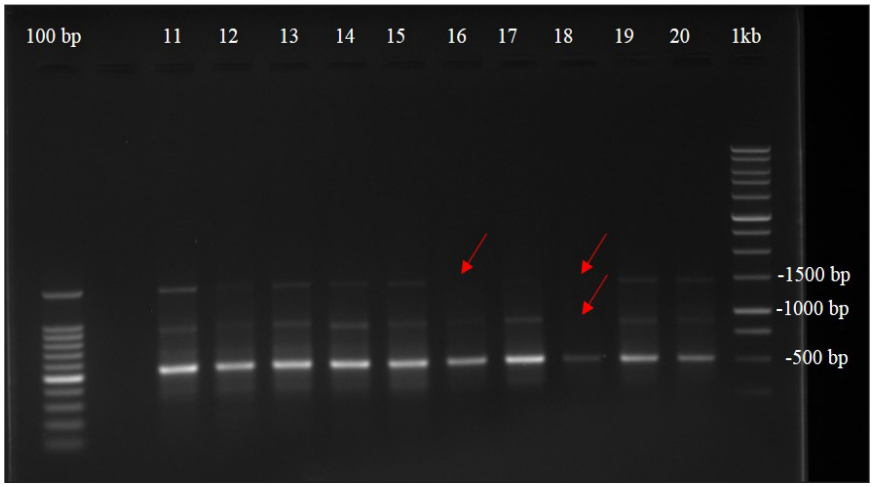

Figure 1. IRAP amplification profile of primer Copia-(F). The accessions shown were provenance of FRPK, accession number 11 to 20. DNA ladder of 100 bp and 1 kb (Promega Corporation) were located on the left and right well. The arrows show variations observed in the bands

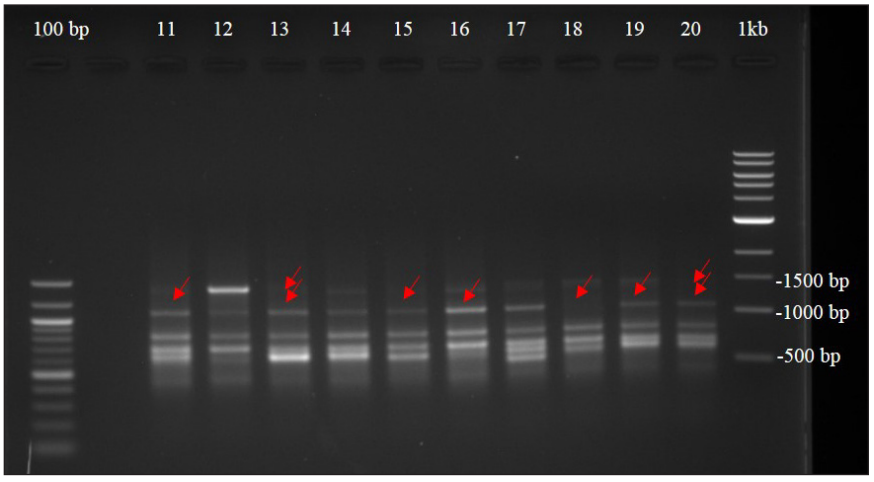

Figure 2. IRAP amplification profile of primer Nikita-57. The accessions shown were provenance of FRMJ, accession number 11 to 20. DNA ladder of 100 bp and 1 kb (Promega Corporation) were located on the left and right well. The arrows show variations observed in the bands

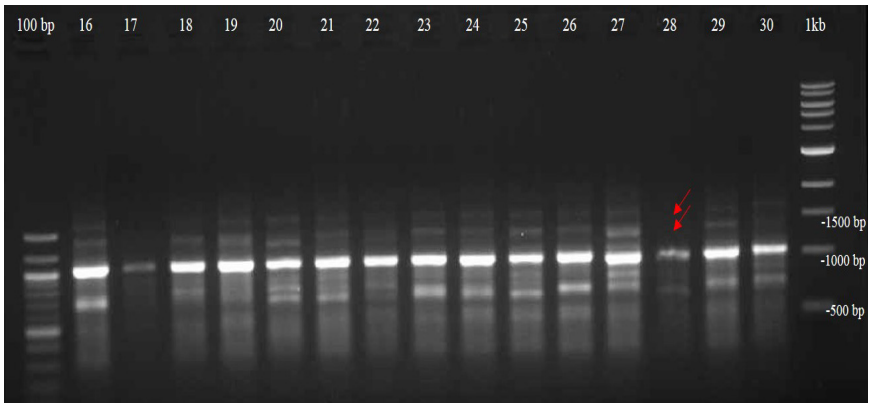

Figure 3. IRAP amplification profile of primer Sabrina-C0945. The accessions shown were provenance of FRMJ, accession number 16 to 30. DNA ladder of 100 bp and 1 kb (Promega Corporation) were located on the left and right well. The arrows show variations observed in the bands

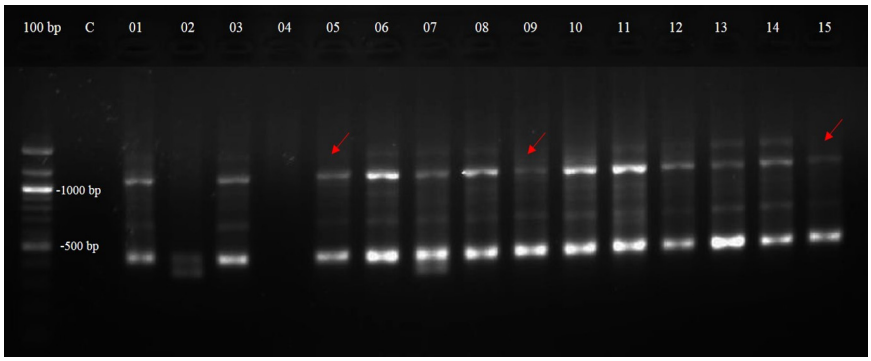

Figure 4. IRAP amplification profile of primer Sukkula-9900. The accessions shown were provenance of FRMJ, accession number 1 to 15 and the C lane shows Control. DNA ladder of 100 bp (Promega Corporation) is located on the left well. The arrows show variations observed in the bands

As a result, a total of 1247 polymorphic bands were scored from the range of 300 – 1500 base pairs (Table 3). The Percentage of Polymorphic Bands (PPB) ranged from 50.0% to 100.0% and the Polymorphism Information Content (PIC) values of the markers ranged from 0.469 to 0.833 with the average of 0.6816 (Table 4). The highest recorded PIC value was by primer Nikita-57 while the lowest PIC value was by the combination of primers Sukkula-9900 and Copia-F. The PIC values for all primers were generally more than 0.50 except for the combination of primers Sukkula-9900 and Copia-F. The Heterozygosity (*He*) values were positively related to PIC values, ranging from 0.565 to 0.852 (Table 4). Then, the Marker Index (MI) was calculated by multiplying PIC value and Effective Multiplex Ratio (EMR) value. The EMR values were positively correlated with the PPB values; thus, the higher the PPB values, the higher the EMR values. The MI values for all the primers tested in this study ranged from 0.68 to 3.72 while the EMR values ranged from 1.00 to 5.00 (Table 4). These

criteria, the PPB, PIC, MI, EMR and *He* are usually used to assess the informativeness of the marker system or to evaluate the discriminatory power of a marker.

The retrotransposon families of Copia, Nikita, Sukkula and Sabrina were initially identified in the plant genome of barley, mostly present as solo long terminal repeats (LTRs) (Leigh et al., 2003). Leigh et al. (2003) also successfully identified the efficiency of 15 retrotransposon-based primers in barley for their applications in single-primer IRAP, two-primer IRAP, and REMAP technique. Nikita-57 primer was observed to produce very clear profiles for all the marker techniques (single-primer

Table 3
Comparison of the observed band size (bp) to the expected band size (bp) and optimized annealing temperature [Ta (°C)]

No.	Primers' name	Expected band size range (bp) (Leigh et al., 2003)	Observed band size range (bp)	Ta (°C)
1.	Copia-F	~270	500 - 1500	37.0
2.	Sabrina-C0945	1000 - 4000	800 - 1500	42.5
3.	Sukkula-9900	500 - 4000	300 - 1500	50.5
4.	Nikita-57	1000 - 4000	400 - 1500	42.9
5.	Combination: Copia-F + Sukkula-9900	Non recorded	400 - 1500	43.8
6.	Combination: Sabrina-C0945 + Sukkula-9900	Non recorded	300 - 700	46.5

Table 4
Informativeness assessment of the IRAP markers

No.	Primers' name	NSB	NPB	NMB	PPB	PIC	EMR	MI	*He*
1.	Copia-F	3	2	1	66.7%	0.549	1.33	0.73	0.627
2.	Sabrina-C0945	7	5	2	71.4%	0.819	3.57	2.92	0.839
3.	Sukkula-9900	5	5	0	100%	0.743	5.00	3.72	0.779
4.	Nikita-57	7	5	2	71.4%	0.833	3.57	2.97	0.852
5.	Sukkula-9900 + Copia-F	5	3	2	60.0%	0.469	1.80	0.844	0.565
6.	Sukkula-9900 + Sabrina-C0945	4	2	2	50.0%	0.677	1.00	0.677	0.728
	Total	31	22	9	-	-	-	-	-
	Minimum	3	2	0	50.0%	0.469	1.00	0.73	0.565
	Maximum	7	5	2	100%	0.833	3.57	3.72	0.852
	Mean	5.2	3.7	1.5	69.9%	0.682	2.18	1.98	0.732

Note: NSB – Number of Scored Band, NPB – Number of Polymorphic Band, NMB – Number of Monomorphic Band, PPB – Percentage of Polymorphic Band, PIC – Polymorphism Information Content, EMR – Effective Multiplex Ratio, MI - Markers Index, *He* - Heterozygosity

IRAP, two-primer IRAP and REMAP). In the primer comparison study, the clarity profile was differentiated as poor < clear bands < very clear. The same indication of the clarity profile was also observed from Sukkula-9900 primer. However, the rate of polymorphism of Sukkula-9900 primer ranged from 40 to 80% for all the marker techniques; it was higher compared to Nikita-57 primer which only ranged from 10 to 60%. For Sabrina-C0945 primer, the capacity of clarity profile was indicated as clear, and the rate of polymorphisms observed from all the marker techniques ranged from 10 to 30%.

Findings in this study demonstrated that the retrotransposon elements were present in the genomic DNA of *E. longifolia*. Many studies have also demonstrated that IRAP markers are universal and successful for many tested species to date (Kalendar & Schulman, 2007). Among the advantages of IRAP markers is only a little or no sequence information of the species is needed, meaning that IRAP markers are applicable to be used for any plant species without the prior knowledge of gene sequences. In this study, primer sequences were obtained from the study of barley retrotransposon families (Leigh et al., 2003) because most of the primers used in this study had been utilized for other species in several different studies without the information of the species DNA sequences. Additionally, according to Kalendar et al. (2011) retrotransposon elements found in barley families are abundant, which is more than 50%, in comparison to other plant species (Ragupathy et al., 2013).

For an example, IRAP markers had been utilized for the following plant species; *Triticum* species (Farouji et al., 2015), saffron (Alsayied et al., 2015), Indian potato (Sharma & Nandieneni, 2014), plum (Senkova et al., 2013), persimmon (Yuan et al., 2012), Japanese apricot (Yuying et al., 2011), some exotic plant species from Perm region (Boronnikova & Kalendar, 2010), *Citrus suhuiensis* (Agisimanto et al., 2008), Japanese persimmon (Guo et al., 2006) and even fungus species (Santana et al., 2012). Furthermore, according to Ragupathy et al. (2013), based on the list of retroelements found in the plant genomes that have whole genome sequences (consisting of 32 different species), retrotransposon elements can be found as low as 7.02% in *Arabidopsis thaliana* (weed) and as high as 75.6% in *Zea mays* (maize). However, according to the list, none of the retroelements could be found in *Phoenix dactylifera* (date-palm). Retrotransposon elements were also abundant in eukaryotic genome, therefore making the elements a great choice as molecular markers (Kalendar et al., 2017).

On the other hand, among the advantages of IRAP is the replication strategy also known as "copy and paste" mechanism, in which the transposition occurs via RNA intermediate resulting in new insertions that increases polymorphisms in the genome (Kalendar et al., 1999). The structure

of retrotransposon elements is highly conserved. Additionally, the dispersion of the elements is abundant and prevalent, and the markers are also highly reproducible. Grzebelus (2006) and Guo et al. (2006) recorded that both IRAP markers were applicable to evaluate the intraspecific relationship since their markers were extremely polymorphic. Along with the explained characteristics, Alsayied et al. (2015) mentioned that retrotransposon-based markers could be used as indicators in biodiversity assessment.

Genetic Diversity of *Eurycoma longifolia* Provenances

The genetic diversity parameters such as the observed number of alleles (*Na*), effective number of alleles (*Ne*), Nei's gene diversity (*He*), Shannon's information index (*Na*), and percentage of polymorphic loci (%) were generated using POPGENE software version 1.32 (Yeh et al., 1999) (Table 5).

The analysis indicated that the observed number of alleles (*Na*) of FRPK (1.65 ±0.486) was higher than and FRMJ (1.55 ±0.506) while the effective number of alleles (*Ne*) of FRPK was 1.43 ±0.422 and FRMJ was 1.37 ±0.398. Nei's gene diversity (*He*) recorded from the provenance of FRPK was 0.238 ±0.221 and FRMJ was 0.209 ±0.215. The Shannon's information index (*I*) reflected the same order as *He* did: FRPK (0.348 ±0.308) and FRMJ (0.306 ±0.307). For the percentage of polymorphic loci (%), both provenances scored more than 50% with FRPK having 64.5% and FRMJ having 54.8% (Table 5).

In the SNPs study, the values of observed heterozygosity (H_0) was lower than the expected heterozygosity (*He*) (Osman et al., 2003). The *He* values generated by IRAP markers in this study were considered comparable to SNPs which ranged from 0.177 to 0.246. However, the limitation for dominant markers such as IRAP is that the data obtained could not calculate the H_0 values. Nonetheless, a study conducted on the genetic diversity of *E. longifolia* in five populations within the province of Riau using RAPD as markers revealed a very similar range of *He* value which ranged from 0.13 to 0.27 (Zulfahmi, 2013).

In a comparison of the *He* values of *E. longifolia* sampled in this study with that of *Shorea leprosula* Miq., which was listed in the Malaysia Plant Red List (Chua et al., 2010), the *He* values recorded in *E.*

Table 5
Genetic diversity of E. longifolia *provenances inferred by IRAP markers*

No.	Prove-nances	Observed number of alleles (*Na*) ±StDev	Effective number of alleles (*Ne*) ±StDev	Nei's gene diversity (*He*) ±StDev	Shannon's information index (*I*) ±StDev	Percentage of polymorphic loci (%)	The number of polymorphic loci
1.	FRPK	1.65 ±0.486	1.43 ±0.422	0.238 ±0.221	0.348 ±0.308	64.5	20
2.	FRMJ	1.55 ±0.506	1.37 ±0.398	0.209 ±0.215	0.306 ± 0.307	54.8	17

longifolia in this study was lower. The study by Lee et al. (2000), utilizing the allozyme markers to assess the genetic diversity of *S. leprosula* recorded that the *He* value for *S. leprosula* sampled throughout Peninsular Malaysia was surprisingly high (0.369 ± 0.025). Nevertheless, it should be noted that allozyme markers are codominant markers whereas IRAP markers in this study are dominant markers. Weising et al. (2005) mentioned that when associating estimation value of *He*, dominant markers data could generally produce the maximum value of 0.5 since only two alleles could be differentiated at each locus while codominant markers could produce up to 1.0 value of *He*. On the other hand, Lee et al. (2000) also mentioned that the value recorded for *S. leprosula* was among the highest *He* recorded for tropical tree species.

Nevertheless, the percentage of polymorphic loci (%) observed in SNP markers ranged from 49.0% to 75.0% as assessed from 47 individuals originated from six populations; this value was also considered comparable to the percentage of polymorphic loci (%) observed in IRAP markers which were 54.8% and 64.5% (Table 5) as assessed from 56 accessions from two provenances. The findings were also similar to the percentage of polymorphic loci (%) observed in Riau province, ranging from 40.9% to 70.5%.

The genetic structure of the two provenances of *E. longifolia* inferred by IRAP markers is revealed in Table 6. The total genetic diversity (H_T) was found to be 0.239 and the diversity within the population (H_S) was 0.223 (Table 6). The value for H_T and H_S as demonstrated by IRAP markers were also parallel to the study in SNPs (0.288 and 0.219, respectively) (Osman et al., 2003) and RAPD in Riau province (0.29 and 0.2, respectively) (Zulfahmi, 2013). Furthermore, the coefficient of gene differentiation (G_{ST}) in this study was found to be 0.068 (Table 6) which was considered as a moderate value following the classification by Nei (1978). According to Nei (1978), the G_{ST} value can be determined by three levels: low when the G_{ST} value is less than 0.05, moderate when the G_{ST} value is in the range of 0.05 to 0.15 and high when the G_{ST} value is more than 0.15. In this study, G_{ST} value signified that there were 6.8% of the total genetic variations between the provenances tested and there were 93.2% variations among individual accessions within the provenances. Both studies using SNPs and RAPD markers showed the high values of G_{ST}, 0.24 (Osman et al., 2003) and 0.31 (Zulfahmi, 2013) respectively, meaning that the genetic variations between the selected populations in the study were lower than genetic variations within populations.

Table 6
Genetic structure of two provenances (FRMJ and FRPK) of E. longifolia *identified by IRAP markers*

H_T = Total diversity	0.239
H_{TS} = Diversity within population	0.223
G_{ST} = Coefficient of gene differentiation	0.068
N_m = Gene flow based on Gst	6.85
Sample size	56

Contrary to the findings by SNPs (Osman et al., 2003) and RAPD (Zulfahmi, 2013), the G_{ST} value by IRAP in this study was lower, indicating higher chances of gene flow probably because of a higher distance of seeds dispersal. However, Osman et al. (2003) mentioned that higher G_{ST} might had resulted from smaller number of individuals sampled per population in the study. Similarly to the conditions of RAPD (Zulfahmi, 2013), the study of genetic diversity in Riau Provinces (Indonesia) also sampled about five individuals per provenances whereas the study by SNPs (Osman et al., 2003) in Malaysia sampled about six to nine individuals per populations. For this study by IRAP markers in two provenances, the sample sizes were 26 and 30 individuals per provenances.

The gene flow (N_m) based on N_m was found to be 6.85 (Table 6). The N_m value observed in this study was considered as high. In comparison to the N_m value of *E. longifolia* in Riau province (N_m = 1.11), the findings by IRAP in two provenances was higher. Govindaraju (1989) classified gene flow values into three levels: high when the N_m value is more than 1, moderate when the N_m value is in the range of 0.25 to 0.99, and low when the N_m value is less than 0.25.

According to the Oxford Dictionary of Biology, conceptually, gene flow is defined as the exchange of genetic materials by interbreeding between populations of the same species or between individuals within a population. Gene flow increases the variation in the genetic composition of a population but decreasing the genetic variations among the populations. Thus, higher gene flow value indicates higher genetic variations ("Gene flow", 2007). In the case of *E. longifolia*, its seeds dispersal is believed to be aided by wild birds. The pulpy layer of the fruits is the main attractive characteristic that appeals the wild birds (FRIM, n. d.). Therefore, since wild birds fly farther, the tendency of seeds dispersal distance increases.

Cluster Analysis of *E. longifolia* Provenances

A UPGMA dendogram is generated based on the Jaccard similarity coefficient and for the dendogram (Figure 5) varied from 0.7 to 0.96. This means that the similarity of the two provenances is about 70 to 96%. Cluster analysis revealed three broad groups where group 1 could be further divided into 3 subgroups, group 2 could be further divided into 5 subgroups, while group 3 only consisted of 2 accessions from FRMJ. However, the results indicated that *E. longifolia* accessions from two provenances were not differentiated by origin. From the dendogram (Figure 5), it could be observed that group 1 and 2 consisted of almost proportionate number of accessions from the provenances of FRPK and FRMJ.

Findings in this study were contradicted with findings by RAPD and SNPs markers on the study of genetic diversity of *E. longifolia* in Malaysia and Indonesia, respectively. In this study, cluster analysis showed that there was no clear indication that the clustering was based on the

provenances or the geographic origins. Mohd Razi et al. (2013) found that *E. longifolia* cultivars in Malaysia could be clustered into three clusters in which i) East Malaysia (Sabah), ii) middle part of west Malaysia (Terengganu and Pahang) and iii) west part of East Malaysia (Kedah and Kelantan). Furthermore, Osman et al. (2003) demonstrated that the *E. longifolia* sampled from Johor, Langkawi, Terengganu, Pahang, Melaka and tissue culture collection could be grouped into two major groups with geographical origin as the clustering basis. Additionally, Zulfahmi (2013) on the study of five provinces of *E. longifolia* in Indonesia also further subdivided the five provinces into two major groups.

On the other hand, the study of the genetic diversity of *Shorea leprosula* using allozyme marker by Lee et al. (2000) showed that the cluster analysis of eight populations in Malaysia did not reflect the geographical origin of the populations sampled in the study. According to Lee et al. (2000), this might indicate that the populations of *S. leprosula* sampled were part of a continuous population and fragmentation that might have occurred in the past.

Even though the accessions from the two provenances have high similarities, the dendogram (Figure 5) showed clear differentiation between the two provenances. The distribution by Principle Component Analysis (PCA) also revealed the differentiation of two broad groups with moderate genetic similarities (Figure 6).

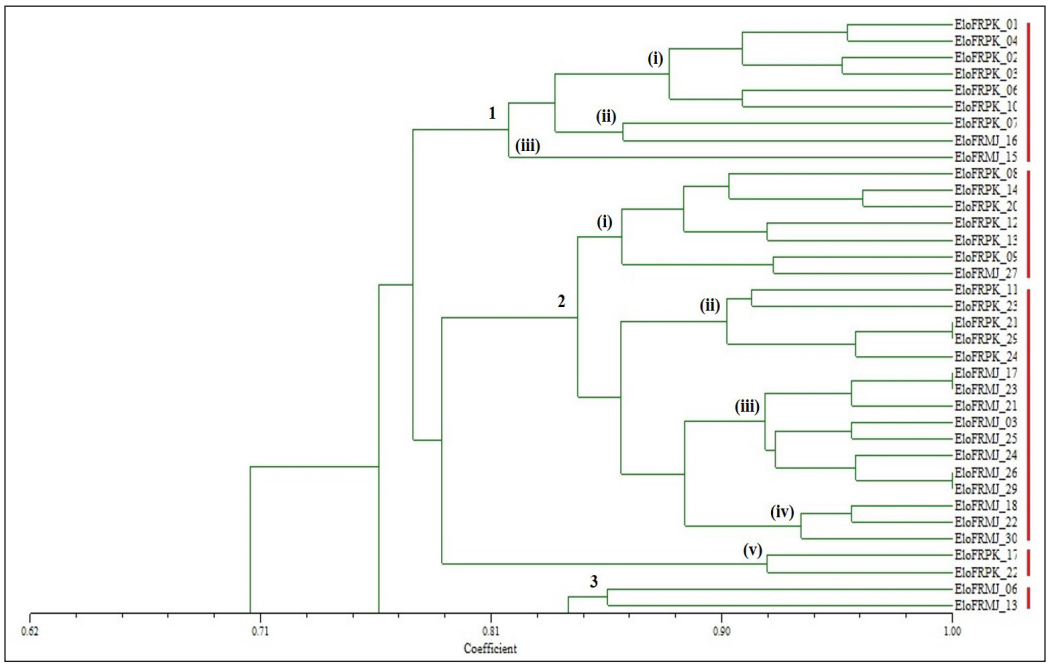

Figure 5. Dendogram constructed from IRAP markers of 56 accessions (FRPK and FRMJ) based on genetic similarity (Jaccard coefficient)

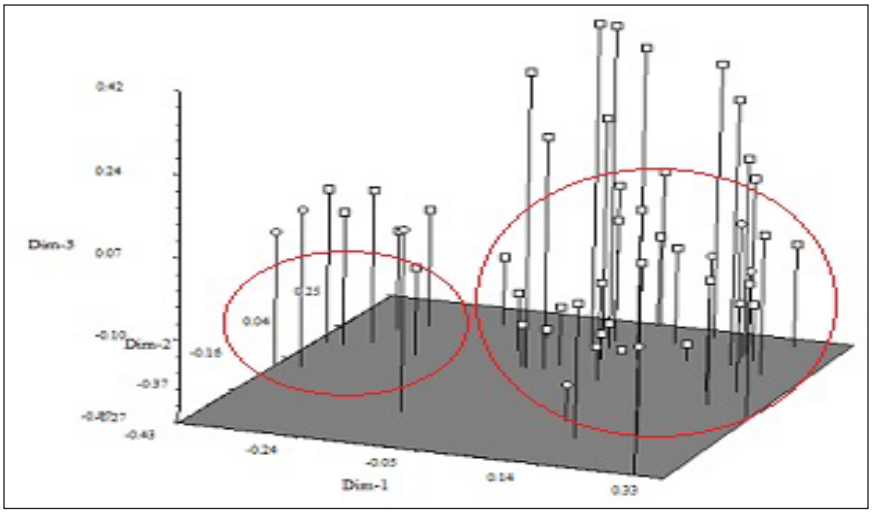

Figure 6. 3D distribution of 56 *E. longifolia* accessions (FRPK and FRMJ) revealed by IRAP markers

CONCLUSION

This study showed that IRAP markers can be effectively used to analyse the genetic diversity of *E. longifolia* based on the informative assessment of the marker system (PPB, PIC, MI, EMR and *He*). The markers are also proven to have the transferability properties, making it interchangeably useable across different species without prior knowledge of gene sequences. The total genetic diversity of two selected *E. longifolia* provenances' were found to be low (0.239). Findings in this study also suggested that there is high gene flow between the provenances (contrary to the previous findings). This implied that the plant breeders may take into consideration the available individuals within the provenances in their breeding strategies and perhaps can overlook the variations between the provenances.

ACKNOWLEDGEMENT

Our sincere appreciation goes to fellow colleagues in FRIM, especially staff from Plant Improvement Program and Genetic Laboratory, along with the others for helping us in various occasions during the period of completing the study.

REFERENCES

Agisimanto, D., Normah, M. N., Ibrahim, R., & Mohamad, A. (2008). *Retros markers for genetic variations in Limau Madu (*Citrus suhuiensis *hort. ex Tanaka)*. Retrieved October 27, 2014, from https://www.pgm-my.org/book/pgm-congress08-proceedings/

Alsayied, N., Fernandez, J. A., Schwarzacher, T., & Heslop-Harrison, J. S. (2015). Diversity and relationship of *Crocus sativus* and its relatives analysed by inter-retroelement amplified polymorphisms (IRAP). *Annals of Botany*, *116*(3), 359–68. DOI: 10.1093/aob/mcv103

Ang, H. H., Cheang, H. S., & Md. Yusof, A. P. (2000). Effects of *Eurycoma longifolia* (Tongkat Ali) on the initiation of sexual performance of inexperienced castrated male rats. *Experimental Animals*, *49*(1), 35–38.

Boronnikova, S. V., & Kalendar, R. N. (2010). Using IRAP markers for analysis of genetic variability in populations of resource and rare species of plants. *Russian Journal of Genetics*, *46*(1), 36–42. doi: 10.1134/S1022795410010060

Chua, L. S. L., Suhaida, M., Hamidah, M., & Saw, L. G. (2010). *Malaysia plant red list (Peninsular Malaysian Dipterocarpaceae)*. Kepong, Malaysia: Forest Research Institute Malaysia.

Farouji, A. E., Khodayari, H., Saeidi, H., & Rahiminejad, M. R. (2015). Genetic diversity of diploid *Triticum* species in Iran assessed using inter-retroelement amplified polymorphisms (IRAP) markers. *Biologia*, *70*(1), 52–60. doi: 10.1515/biolog-2015-0002

Forest Research Institute Malaysia. (n.d.). *Colours of FRIM: The rare sight of fruiting Tongkat Ali*. Retrieved April 30, 2014, from https://www.frim.gov.my/en/the-rare-sight-of-fruiting-tongkat-ali/

"Gene flow". (2007). In *Oxford dictionary of biology* (6th Ed.). Oxford, United Kingdom: Oxford University Press.

Govindaraju, R. D. (1989). Variation in gene flow levels among predominantly self-pollinated plants. *Journal of Evolution Biology*, *2*(3), 173–181. doi: 10.1046/j.1420-9101.1989.2030173.x

Grzebelus, D. (2006). Transposon insertion polymorphism as a new source of molecular markers. *Journal of Fruit and Ornamental Plant Research*, *14*, 21–29.

Guo, D., Zhang, H., & Luo, Z. (2006). Genetic relationships of *Diospyros kaki* Thunb. and related species revealed by IRAP and REMAP analysis. *Plant Science*, *170*(3), 528–533.

Hettasch, M. H., Lunt, K. A., Pierce, B. T., Snedden, C. L., Steyn, D. J., Venter, H. M., & Verryn, S. D. (2002). *Tree breeding course manual*. Pretoria, South Africa: CSIR Environmentek.

Kalendar, R. N, Grob, T., Regina, M., Suoniemi, A., & Schulman, A. H. (1999). IRAP and REMAP: Two new retrotransposon-based DNA fingerprinting techniques. *Theoretical and Applied Genetics*, *98*(5), 704–711.

Kalendar, R. N., Aizharkyn, K. S., Khapilina, O. N., Amenov, A. A., & Tagimanova, D. S. (2017). Plant diversity and transcriptional variability assessed by retrotransposon-based molecular markers. *Journal of Genetics and Breeding*, *21*(1), 128–134.

Kalendar, R. N., Flavell, A. J., Ellis, T. H. N., Sjakste, T., Moisy, C., & Schulman, A. H. (2011). Analysis of plant diversity with retrotransposon-based molecular markers. *Heredity*, *106*(4), 520–530.

Kalendar, R., & Schulman, A. H. (2007). IRAP and REMAP for retrotransposon-based genotyping and fingerprinting. *Nature Protocols*, *1*(5), 2478–2484. doi: 10.1038/nprot.2006.377

Kuo, P. C., Damu, A. G., Lee, K. H., & Wu, T. S. (2004). Cytotoxic and antimalarial constituents from the roots of *Eurycoma longifolia. Bioorganic and Medicinal Chemistry*, *12*(3), 537–544.

Lee, S. L., Wickneswari R., Mahani, M. C., & Zakri A. H. (2000). Genetic diversity of a tropical tree species, *Shorea leprosula* Miq. (Dipterocarpaceae), in Malaysia: Implications for conservation of genetic resources and tree improvement. *Biotropica*, *32*(2), 213–224. doi: 10.1111/j.1744-7429.2000.tb00464.x

Leigh, F., Kalendar, R. N, Lea, V., Lee, D., Donini, P., & Schulman, A. H. (2003). Comparison of the utility of barley retrotransposon families for genetic analysis by molecular marker techniques. *Molecular Genetics and Genomics*, *269*(4), 464–474.

Low, B. S., Choi, S. B., Abdul Wahab, H., Das, P. K., & Chan, K. L. (2013). Eurycomanone, the major quassinoid in *Eurycoma longifolia* root extract increases spermatogenesis by inhibiting the activity of phosphodiesterase and aromatase in steroidogenesis. *Journal of Ethnopharmacology*, *149*(1), 201–207.

Mohd Azmi, M. I., Ahmad Fauzi, P., & Norini, H. (2004). Economic analysis of *Eurycoma longifolia* (Tongkat Ali) harvesting in Peninsular Malaysia. In C. Y. Shyun, M. Mohtar, & N. M. Yunos (Eds.), *Tongkat Ali, Kacip Fatimah and Pegaga new dimensions in complementary healthcare: Proceedings of the seminar on medicinal plants, 20-21 August 2002* (pp. 3-6). Kepong, Malaysia: Forest Research Institute Malaysia.

Mohd Razi, A. R., Abdul Aziz, A., Syed Alwi, S. S. B., & Aziz Ramlan, A. (2013). Relationships between Malaysians cultivars of Tongkat Ali (*Eurycoma longifolia* Jack.) obtained through RAPD analysis. *International Journal of Biotechnology for Wellness Industries*, *2*(1), 45-50.

Mohd Zaki, A., Nor Fadilah, W., Abdul Rashid, L., Nurnadiah, R., Mohd Radzi, A., Mohamad Lokmal, N., … Mohamad, O. (2015a). Preliminary phytochemical screening of polysaccharides content for selection of high quality planting materials: *Eurycoma longifolia*. *Jurnal Teknologi*, *77*(24), 101–105.

Mohd Zaki, A., Nor Fadilah, W., Mohd Radzi, A., Sui Kiong, L., Ab. Rasip, A. G., Mohamad Lokmal, N., … Regina Mariah, J. (2015b). Preliminary phytochemical screening of eurycomanone for selection of high quality planting materials: *Eurycoma longifolia*. *Malaysian Applied Biology*, *44*(1), 23–25.

Murray, M. G., & Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research*, *8*(19), 4321–4325.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, *89*(3), 583–590.

Nor Fadilah, W., Mohd Zaki, A., & Mohamad Lokmal, N. (2016). Breeding strategies of *Eurycoma longifolia*: Present and future. In B. J. Chee, M. Mastura, K. Getha, M. G. H. Khoo, J. Mailina, M. A. Adiana, & M. S. Roshan Jahn (Eds.), *Proceedings of the 14th Seminar on Medicinal and Aromatic Plants* (pp. 18–23). Kepong, Malaysia: Forest Research Institute Malaysia.

Osman, A., Jordan, B., Lessard, P. A., Norwati, M., Mohd Rosli, H., Norifiza, M. R., … Housman, D. E. (2003). Genetic diversity of *Eurycoma longifolia* inferred from single nucleotide polymorphisms. *Plant Physiology*, *131*(1), 1294–1301.

Performance Management and Delivery Unit. (2013). *Economic Transformation Programme (ETP) annual report 2012*. Retrieved August 25, 2013, from http://etp.pemandu.gov.my/annualreport/upload/Eng_ETP2012_Full.pdf

Ragupathy, R., You, F-M., & Cloutier, S. (2013). Arguments for standardizing transposable element annotation in plant genomes. *Trends in Plant Science*, *18*(7), 367–376.

Roche Diagnostic. (2008). *High pure PCR template preparation kit: Rapidly purify genomic DNA for diverse applications*. Retrieved August 25, 2013, from https://lifescience.roche.com/documents/High-Pure-PCR-Template-Preparation-Kit.pdf

Rohlf, F. J. (2009). *NTSYSpc: Numerical taxonomy system (Version. 2.21c)*. New York, NY: Exeter Software.

Santana, M. F., De Araujo, E. F., De Souza, J. T., Mizubuti, E. S. G., & De Queiroz, M. V. (2012). Development of molecular markers based on retrotransposons for the analysis of genetic variability in *Moniliophthora perniciosa*.

European Journal of Plant Pathology, *134*(3), 497–507. doi: 10.1007/s10658-012-0031-4

Senkova, S., Ziavroska, J., Bezo, M., Stefunova, V., & Razna, K. (2013). Utilization of IRAP technique for plums genotypes differentiation. *Bioscience Research*, *10*(1), 1–7.

Sharma, V., & Nandineni, M. R. (2014). Assessment of genetic diversity among Indian potato (*Solanum tuberosum* L.) collection using microsatellite and retrotransposon-based marker systems. *Molecular Phylogenetics and Evolution*, *73*, 10–17. doi: 10.1016/j.ympev.2014.01.003

Subramaniam, V. (Ed.) (2013). *Malaysian herbal heritage*. Kepong, Malaysia: Forest Research Institute Malaysia.

Teo, C. H., Tan, S. H., Ho, C. L., & Faridah, Q. Z. (2005). Genome constitution and classification using retrotransposon-based markers in the orphan crop banana. *Journal of Plant Biology*, *48*(1), 96–105.

Tnah, L. H., Lee, C. T., Lee, S. L., Ng, K. S., Ng, C. H., & Hwang, S. S. (2011). Microsatellite markers of an important medicinal plant, *Eurycoma longifolia* (Simaroubaceae), for DNA profiling. *American Journal of Botany*, *98*(5), e130–e132. doi: 10.3732/ajb.1000469

Wan-Muhammad-Azrul, W. A., Mohd-Farid, A., Lee. S. Y., Sajap, A. S., Omar, D., & Mohamed, R. (2018). A survey on the occurrence of pests and diseases in Tongkat Ali (*Eurycoma longifolia*) plantations in Peninsular Malaysia. *Journal of Tropical Forest Science*, *30*(3), 362-375. doi: 10.26525/jtfs2018.30.3.362375

Weising, K., Nybom, H., Wolff, K., & Kahl, G. (2005). *DNA fingerprinting in plants: Principles, methods, and applications* (2nd ed.). Boca Raton, USA: Taylor and Francis Group.

Wong, P. F., Cheong, W. F., Shu, M. H., Teh, C. H., Chan, K. L., & Abu Bakar, S. (2012). Eurycomanone suppresses expression of lung cancer cell tumor markers, prohibitin, annexin 1 and endoplasmic reticulum protein 28. *Phytomedicine*, *19*(2), 138–144.

Yeh, F. C., Yang, R-C., & Boyle, T. (1999). *POPGENE version 1.31: Microsoft Window-based freeware for population genetic analysis (Quick user guide)*. Retrieved June 24, 2016, from https://sites.ualberta.ca/~fyeh/popgene.pdf

Yuan, L., Zhang, Q., Guo, D., & Luo, Z. (2012). Genetic differences among 'Luotian-tianshi' (*Diospyros kaki* Thunb.) genotypes native to China revealed by ISSR and IRAP markers. *Scientia Hortuculturae*, *137*, 75–80.

Yuying, S., Xiajun, D., Fei, W., Binhua, C., Zhihong, G., & Zhen, Z. (2011). Analysis of genetic diversity in Japanese apricot (*Prunus mume* Sieb. et Zucc.) based on REMAP and IRAP molecular markers. *Scientia Horticulturae*, *132*, 50–58.

Zulfahmi, R. (2013). Genetic diversity of *Eurycoma longifolia* Jack. based on random amplified polymorphic DNA marker. *Journal of Tropical Forest Management*, *19*(2), 138–144.