# Molecular DNA identification of medicinal plants used by traditional healers in Malaysia

**N.A.A. Aziz[1], M.I. Ahmad[1] and D.M. Naim[2]**

[1]School of Industrial Technology, Universiti Sains Malaysia, Pulau Pinang, Malaysia
[2]School of Biological Sciences, Universiti Sains Malaysia, Pulau Pinang, Malaysia

Corresponding author: D.M. Naim
E-mail: darlinamdn@usm.my

Genet. Mol. Res. 14 (4): 15937-15947 (2015)
Received June 7, 2015
Accepted September 28, 2015
Published December 7, 2015
DOI http://dx.doi.org/10.4238/2015.December.7.5

**ABSTRACT.** Plants have been used throughout human history for food and medicine. However, many plants are toxic, and cannot easily be morphologically distinguished from non-toxic plants. DNA identification solves this problem and is widely used. Nonetheless, plant DNA barcode identification faces a number of challenges, and many studies have been conducted to find suitable barcodes. The present study was conducted to test the efficiency of commonly used primers, namely *ITS2*, *rpoC1*, and *trnH-psbA*, in order to find the best DNA barcode markers for the identification of medicinal plants in Malaysia. Fresh leaves from 12 medicinal plants that are commonly used by Malay traditional healers were collected from the Tropical Spice Garden, Pulau Pinang, and subjected to polymerase chain reaction amplification using *ITS2*, *rpoC1*, and *trnH-psbA* DNA markers. We found that *trnH-psbA* is the best DNA marker for the species-level identification of medicinal plants in Malaysia.

**Key words:** DNA barcode; Medicinal plant; *ITS2*; *rpoC1*; *trnH-psbA*

 

## INTRODUCTION

Medicinal plants have long been used to treat ailments, and continue to play a major role in the primary health care of about 80% of the world's inhabitants (Farnsworth et al., 1985; El Beyrouthy et al., 2013). Many countries in Asia, Africa, and Latin America use traditional medicine to meet some of their primary health care needs (Kassaye et al., 2006). In addition, large proportions of the populations of many developing countries rely on traditional practitioners and their armamentaria of medicinal plants (WHO, 2002). Malaysia is a multicultural country that has rich traditional modalities (Sooi and Keng, 2013), and the practice of traditional medicine is common amongst ethnic groups such as the Malays, Chinese, Indians, and aborigines (GLOBinMED, 2011). The use of herbal medicines is based on practical experience, observation, and rituals derived from socioreligious beliefs, and this knowledge is passed down through the generations (Sooi and Keng, 2013).

Medicinal plants contain substances that have physiological effects on the human body (Nwachukwu et al., 2010), and plant parts can be used fresh or dried for preventing and healing various forms of ailment (Joharchi and Amiri, 2012). Normally, the use of medicinal plants varies from species to species, as diseases vary from one form to another in various locations (Nwachukwu et al., 2010). However, due to the physical similarities of plant parts that cause misidentification by consumers and the lack of a standard identification system, medicinal plants and their parts are often adulterated and/or replaced by other plants in the market, resulting in a loss of efficacy and a danger of toxicity (Joharchi and Amiri, 2012). Therefore, the correct identification of medicinal plants is urgently needed to ensure the safety and quality of these natural health products.

Conventionally, botanists tend to describe medicinal plants by their appearance and morphological identification (Rydberg, 2010). However, this morphological technique can cause confusion, particularly when identifying unstructured plant parts (Mahadani and Ghosh, 2013). Therefore, the development of DNA-based markers has been important for the authentication of medicinal plants (Techen et al., 2014). DNA barcoding is a novel technique of identifying biological specimens, which uses short DNA sequences from either nuclear or organelle genomes (Hebert et al., 2003; Techen et al., 2014). This technique has been successful in animal identification at the species level, and has been used in determining species boundaries, identifying new species, and species delimitation (Hebert et al., 2003, 2004a; Techen et al., 2014). However, the standard animal DNA barcode is composed of a portion of the mitochondrial gene COI, which evolves too slowly in plants to serve as a useful DNA barcode and has led to the search for a suitable DNA barcode for plants (Fazekas et al., 2008, 2012).

Although this is a growing area of scientific interest, few studies have been conducted on species identification and the barcoding of medicinal plants in Malaysia. Therefore, this study was conducted to test the efficiency of commonly used primers, namely *ITS2, rpoC1*, and *trnH-psbA*, as potential plant identification markers for medicinal plants used by Malay traditional practitioners.

## MATERIAL AND METHODS

### Sample collection

We collected 12 plant samples from the Tropical Spice Garden (TSG), Pulau Pinang, Malaysia. Plants were randomly collected in the morning from 9.00 am until 12.00 noon. Whole plants and/or plant parts were cut using sharp scissors and placed in a sterile plastic bag. Morphological confirmation and species identification was conducted with the help of expert local

taxonomists and the Photographic Atlas of Botany and Guide to Plant Identification (Castner, 2005), and the results are summarized in Table 1. The samples were then taken to traditional herbal practitioners who confirmed the medicinal value of the plants and their effectiveness in treating ailments. Voucher samples were deposited in the herbarium of Universiti Sains Malaysia. In addition, 185 conspecific sequences of the collected plant samples were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/genbank), in which the amounts of *ITS2*, *rpoC1*, and *trnH-psbA* were 67, 35, and 83, respectively.

**Table 1.** Twelve medicinal plants collected from the Tropical Spice Garden, Pulau Pinang, Malaysia.

| Vernacular name | Family | Species name |
|---|---|---|
| Kari | Rutaceae | *Murraya koenigii* |
| Pecah beling | Acanthaceae | *Sambucus chinensis* |
| Ganda rusa | Acanthaceae | *Justicia gendarussa* |
| Pegaga | Apiaceae | *Centella asiatica* |
| Kesom | Polygonaceae | *Persicaria odorata* |
| Ulam raja | Asteraceae | *Cosmos caudatus* |
| Belalai gajah | Acanthaceae | *Clinacanthus nutans* |
| Sirih | Piperaceae | *Piper betel* |
| Hempedu bumi | Acanthaceae | *Andrographis paniculata* |
| Dukung anak | Phyllanthaceae | *Phyllanthus niruri* |
| Kaduk | Piperaceae | *Piper sarmentosum* |
| Misai kucing | Lamiaceae | *Orthosiphon stamineus* |

## DNA isolation

About 200 mg of fresh young leaves were ground to a fine paste and homogenized in a DNA extraction buffer (50 mM Tris HCL [pH 8.0], 25 mM EDTA [pH 8.0], 150 mM NaCl, 40 mL $H_2O$, and 1 g polyvinylpyrrolidone [PVP-40; molecular weight 40,000]). The mixture was made up to 100 mL with distilled water, and the pH was adjusted to 5.0 with HCl. The ground material was then transferred to a 2-mL tube. The DNA isolation procedure used for the leaf material was a modification of the cetyltrimethylammonium bromide protocol by Cota-Sánchez et al. (2006). Briefly, 500 µL of hexadecyltrimethylammonium bromide (CTAB) buffer was added to the 2-mL tube that contained the ground material. The mixture was then incubated overnight at 56°C and mixed intermittently by inversion. Subsequently, 250 µL chloroform was added to the mixture and mixed by inversion. The tube was then spun at 12,000 *g* for 5 min. Two layers of aqueous phase were formed, and the upper aqueous phase (that contained the DNA) was transferred to a clean tube. The DNA was then precipitated with 50 µL 7.5 M ammonium acetate (NaAc, pH 4.6) and two volumes of 95% EtOH, and incubated at -20°C for 1 h.

To isolate the precipitate, the tube was centrifuged at 13,000 rpm for 15 min. The supernatant was removed and the DNA pellet was washed twice by adding 500 µL ice-cold 70% EtOH. The tube was then centrifuged for 10 min, the ethanol was poured off, the tube was centrifuged for a few more seconds, and the remaining liquid was removed using a pipette. The tube was then dried in an incubator (with the cap open) for 30 min, or until the remaining liquid had evaporated. The DNA pellet was then resuspended in Milli-Q® water, and the amount of the Milli-Q® water depended on how much DNA was isolated. After resuspension, the DNA was incubated at 65°C for 20 min to destroy any DNases that may have been present, and stored at -20°C.

To obtain high-quality DNA that was free from polysaccharides and other metabolites that might interfere with polymerase chain reaction (PCR) amplification, the purified DNA concentration of each sample was estimated by comparing ethidium bromide-stained band intensities against λ DNA.

## DNA amplification, purification, and sequencing

PCR amplification of the purified DNA was performed in 200-µL reaction tubes in a total volume of 50 µL. The PCR mixture contained 10X PCR buffer, 25 mM $MgCl_2$, 1.25 mM of each dNTP, 1 U *Taq* polymerase, 10 µM of each primer, 20 ng genomic DNA, and 30 µL Milli-Q® water. The primers used were *ITS2-R*, 5'-GACGCTTCTCCAGACTACAAT-3'; *ITS2-2F*, 5'-GCGATACTTGGTGTGAAT-3'; *rpoC1-2F*, 5'-GGCAAAGAGGGAAGATTTCG-3', (Royal Botanical Gardens, 2007); *rpoC1-4R*, 5'-CCATAAGCATATCTTGAGTTGG-3', (Royal Botanical Gardens, 2007); *psbA-F*, 5'-GTTATGCATGAACGTAATGCTC-3' (Sang et al., 1997); and *trnH-R*, 5'-CGCGCATGGATTCACAAATC-3' (Sang et al., 1997). The thermal regime consisted of an initial denaturation at 95°C for 2 min; 35 cycles at 94°C for 45 s, 40°-50°C (depending on the primer used) for 45 s, 72°C for 1 min, and a final incubation at 72°C for 10 min. The PCR products were visualized on 0.8% agarose gels and the most intense bands were selected for purification, according to the PCR Purification Kit protocol of Genomics Bioscience and Technology (Taiwan). The purified PCR products were sequenced using an ABI3730XL automated sequencer (Applied Biosystems) provided by the Bioneer Corporation (Korea).

## Sequence analysis and species identification

The amplified partial sequences of *ITS2*, *rpoC1*, and *trnH-psbA* from the medicinal plants were included in a BLASTn search (http://blast.ncbi.nlm.nih.gov), which takes a sequence query and matches it against BOLD and/or GenBank. Two statistics that are used to measure species identification efficiency are the E-value and maximum identity. High identification efficiency is obtained when the hit and E-value reach 100% and 0, respectively (see Mankga et al., 2013). All of the sequences were collapsed into haplotypes using Collapse version 1.2 (Provan et al., 2005), which were then aligned using ClustalW version 2.0.12 (Larkin et al., 2007) in combination with 186 sequences for the three barcodes (*ITS2*, *rpoC1*, and *trnH-psbA*) that were downloaded from GenBank. All of the sequences were manually checked and trimmed using BioEdit version 7.2.5 (Hall, 1999); alignments were subsequently revised by eye in an effort to maximize positional homology. All of the positions containing gaps and missing data were eliminated from the datasets, and sequence GC content was calculated using a maximum composite likelihood model.

Three parameters were used to characterize the interspecific divergence of a sequence (Chen et al., 2010): 1) the average interspecific distance between all species in each genus with at least two species; 2) the average theta prime (θ'), where θ' is the mean pairwise distance within each genus with more than one species; and 3) the minimum interspecific distance, i.e., the minimum interspecific distance within each genus with at least two species. Three parameters were also used to determine intraspecific sequence variation: 1) the average intraspecific difference between all samples within each species with more than one individual; 2) theta (θ), which is the mean pairwise distance within each species with at least two individuals; and 3) the coalescent depth, which is the maximum intraspecific distance within each species with at least two individuals (see Chen et al., 2010; Yao et al., 2010).

Species discrimination was conducted using a neighbor-joining tree in MEGA version 5.0 (Tamura et al., 2011) with a K2P molecular evolutionary model; branch supports were determined using 1000 bootstrap replicates. Successful identification using this method was inferred when sequences from the same species formed a monophyletic group.

The DNA barcoding gap (defined as the spacer region between intra- and interspecific genetic variation) and identification efficiency were used to investigate breaks in the distribution of genetic pairwise distances using automatic barcode gap discovery (ABGD) (Puillandre et al., 2012). This method proposes a standard definition of the barcode gap and can be used to partition the dataset into candidate species, even when two distributions overlap (see Liu et al., 2014). Sequences of each dataset (*ITS2*, N = 20, 13 species; *rpoC1*, N = 11, 4 species; *trnH-psbA*, N = 20, 6 species) were uploaded to http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html. For this analysis, we set values for *P* (prior maximum divergence of intraspecific diversity) that ranged from 0.001 to 0.1, because if *P* is set too high the entire dataset will be considered a single species (Puillandre et al., 2012). Distance analysis was calculated based on the K2P analysis.

## RESULTS AND DISCUSSION

An ideal DNA barcode must have adequate conserved regions for universal primer design, enough variability to be used for species identification (Consortium for the Barcode of Life), and have a good ability to differentiate between closely related species, which will only be achieved when the genetic distance between species is significantly higher than within congeneric species (Hebert et al., 2004b; Mankga et al., 2013). While morphological plant identification remains controversial, several approaches are available for assigning species based on nucleotide sequences, namely phylogenetic tree- and distance-based methods. By using both methods, our study found that *trnH-psbA* is the most reliable DNA barcode for the commonly used medicinal plants (Table 1) that were collected from the Penang Botanical Garden, Malaysia.

### Success rate of PCR amplification, sequence information, and identification

A reliable barcode region must be relatively short to facilitate the screening process (i.e., genomic DNA extraction, amplification, and sequencing) for species identification (Kress et al., 2005; Gao et al., 2010). In this study, high-quality DNA was obtained from all of the samples analyzed. *ITS2, rpoC1*, and *trnH-psbA* from all of the samples were successfully amplified, with a high efficiency rate. A total of 197 sequences for the three barcodes were analyzed, 12 of which were obtained from the medicinal plants collected from the TSG, Penang, and 185 from GenBank. Each sample was successfully amplified by the 530-bp (*ITS2*), 765-bp (*rpoC1*), and 542-bp (*trnH-psbA*) fragments. *trnH-psbA* had a relatively high amplification efficiency (98%) compared to *ITS2* (95%) and *rpoC1* (90%). All of the sequences were used in the subsequent alignment analysis and were deposited in GenBank.

Our results show that of the 12 samples analyzed, *trnH-psbA* was correctly identified in 50.0 and 75.0% of the sequences at the species and genus levels, respectively, indicating that the scientific names recovered from the BLASTn analysis matched the putative scientific names expected based on the vernacular names and morphological identification (Table 2). In contrast, the correct identifications for *ITS2* and *rpoC1* were lower at the species level (33.3 and 16.7%, respectively). When BLASTn was used for samples from a wide range of taxa, *trnH-psbA* consistently exhibited a higher success rate for species- and genus-level identification (99.8%) than *ITS2* (69.5%) or *rpoC1* (52.3%). The mismatch in species identification between BLASTn and searches based on vernacular names was probably due to variations in the naming of the same plants throughout their ethnic and geographical ranges (Mankga et al., 2013). For example,

in Malaysia the name Sirih is used for *Piper betel*, *P. sarmentosum*, and *P. nigrum*. The correct identification of medicinal plant species is crucial, because adulterants, although belonging to the same genus, do not have any medicinal properties (Joharchi and Amiri, 2012). In Iran, *Bunium cylindricum* (an adulterant) was mixed with Zire-e-siah (*Bunium persicum*) and sold in the market, resulting in the degradation of the quality and efficacy of the drug (Joharchi and Amiri, 2012).

**Table 2.** BLAST analysis with the percentages of maximum identity (Max. Id.) for each sample collected.

| No. | Vernacular name | Scientific name | DNA region | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | *ITS2* | | *rpoC1* | | *trnH-psbA* | |
| | | | Scientific name | Max. Id. (%) | Scientific name | Max. Id. (%) | Scientific name | Max. Id. (%) |
| 1 | Kari | *Murraya koenigii* | *Murraya koenigii* | 99 | *Murraya exotica* | 99 | *Murraya koenigii* | 99 |
| 2 | Pecah Beling | *Strobilanthes crispa* | *Sambucus chinensis* | 91 | *Baphicacanthus cusia* | 99 | *Baphicacanthus cusia* | 99 |
| 3 | Ganda Rusa | *Justicia gendarussa* | *Justicia gendarussa* | 99 | *Sesamum indicum* | 96 | *Justicia gendarussa* | 97 |
| 4 | Pegaga | *Centella asiatica* | *Centella asiatica* | 98 | *Centella asiatica* | 99 | *Centella asiatica* | 99 |
| 5 | Kesom | *Persicaria odorata* | *Persicaria posumbu* | 99 | *Persicaria hydropiper* | 100 | *Persicaria hydropiper* | 97 |
| 6 | Ulam Raja | *Cosmos caudatus* | *Cosmos sulphureus* | 99 | *Bidens frondosa* | 99 | *Coreopsis* | 89 |
| 7 | Belalai Gajah | *Clinacanthus nutans* | *Clinacanthus magnusianus* | 94 | *Sesamum indicum* | 96 | *Clinacanthus nutans* | 100 |
| 8 | Sirih | *Piper betel* | *Piper nigrum* | 99 | *Piper galeatum* | 100 | *Piper nigrum* | 97 |
| 9 | Hempedu Bumi | *Andrographis paniculata* | *Anisomeles indica* | 94 | *Andrographis paniculata* | 99 | *Andrographis paniculata* | 99 |
| 10 | Dukung Anak | *Phyllanthus niruri* | *Phyllanthus urinaria* | 99 | *Glochidion eriocarpum* | 99 | *Phyllanthus urinaria* | 99 |
| 11 | Kaduk | *Piper sarmentosum* | *Piper sarmentosum* | 88 | *Piper galeatum* | 99 | *Piper sarmentosum* | 99 |
| 12 | Misai Kucing | *Orthosiphon stamineus* | *Orthosiphon aristatus* | 100 | *Clerodendranthus spicatus* | 99 | *Clerodendranthus spicatus* | 100 |

Although nuclear ribosomal DNA (*ITS*) has been used as a universal barcode in the discrimination of more than 6600 plant samples (Chen et al., 2010; Liu et al., 2012), we did not find this to be the case in our study. In the present study, *trnH-psbA* was relatively short (542 bp) compared to *rpoC1*, but was slightly longer than *ITS2* (Table 1). Consequently, *trnH-psbA* was relatively easy to amplify using the selected primers, and we obtained a high amplification success rate (98%). Other studies (i.e., Chen et al., 2010; Gao et al., 2010) have also shown that short barcode regions have a relatively high amplification and sequencing efficiency. In addition, *trnH-psbA* had the highest discriminatory power of the three regions, as determined by the BLASTn analysis (Table 2).

## Intra- and interspecific sequence polymorphisms at each locus

A good barcode should exhibit high interspecific divergence in order to discriminate between different species. A summary of the intra- and interspecific genetic divergence and variation of the three regions examined (*ITS2*, *rpoC1*, and *trnH-psbA*) is presented in Table 3. By comparing the interspecific divergence of the three regions, we found that *trnH-psbA* exhibited the highest interspecific divergence, with a maximum interspecific distance of 0.0595 ± 0.0025 (SD), θ' = 0.0481 ± 0.0019 (SD), and a minimum interspecific distance of 0.0083 ± 0.0096 (SD), while *rpoC1* exhibited the lowest interspecific level (Table 3). A Wilcoxon signed-rank test confirmed that *trnH-psbA* exhibited the highest interspecific divergence (Table 4). The highest and lowest intraspecific genetic divergence was exhibited by *trnH-psbA* and *rpoC1*, respectively (Table 3), which was also confirmed by a Wilcoxon signed-rank test (Table 4).

## Barcoding gap assessment

The ABGD is a tool that automatically finds the barcode gap, and can be used when two distributions overlap to partition the dataset into candidate species (Puillandre et al., 2012; Liu et

al., 2014). In this study, the ABGD method for both initial and recursive partition for the *trnH-psbA* dataset (N = 20, six species) resulted in six groups with *P* ranging from 0.001 to 0.0129, three groups with *P* at 0.0215, and two groups with *P* ranging from 0.0359 to 0.1000 (Table 5). By increasing the prior intraspecific limit and lowering the relative width of the barcoding gap (X value), more groups than the number of species defined taxonomically were obtained. In general, recursive partitions contain more groups than initial partitions (Puillandre et al., 2012); however, recursive partitions are expected to better-handle heterogeneities in the dataset, while initial partitions are typically stable on a wider range of prior values and are usually close to the number of groups described by taxonomists (Puillandre et al., 2012). Using the *ITS2* (N = 20, 13 species) and *rpoC1* (N = 11, four species) datasets, we obtained one partition using the standard setting (see Material and Methods); however, after lowering the X value, 13 and 4 groups were found for *ITS2* and *rpoC1*, respectively (Table 5). This indicates that the partition of *trnH-psbA* and *ITS2* using ABGD is consistent with the number of species defined taxonomically. However, only *trnH-psbA* exhibited a distinct DNA barcoding gap, whereas both *ITS2* and *rpoC1* exhibited a negative barcoding gap (i.e., overlapping intra- and interspecific sequence variation) (Figure 1). This suggests that *trnH-psbA* is the best marker to discriminate between medicinal plant species, and that the ABGD method is an effective technique for species identification.

**Table 3.** Inter- and intraspecific divergences of *ITS2*, *rpoC1*, and *trnH-psbA*.

| | *ITS2* | *rpoC1* | *trnH-psbA* |
|---|---|---|---|
| Size (bp) | 530 | 765 | 542 |
| Average interspecific distance | 0.0216 ± 0.0092 | 0.0096 ± 0.0048 | 0.0495 ± 0.0025 |
| Theta prime | 0.0196 ± 0.0085 | 0.0091 ± 0.0031 | 0.0481 ± 0.0019 |
| Minimum interspecific distance | 0.0074 ± 0.0065 | 0.0035 ± 0.0084 | 0.0083 ± 0.0096 |
| Average intraspecific distance | 0.0028 ± 0.0016 | 0.0012 ± 0.0026 | 0.0075 ± 0.0069 |
| Theta | 0.0032 ± 0.0018 | 0.0064 ± 0.0052 | 0.0075 ± 0.0069 |
| Coalescent depth | 0.0031 ± 0.0029 | 0.0087 ± 0.0081 | 0.0034 ± 0.0028 |

**Table 4.** Wilcoxon signed-rank test of interspecific divergences between the three DNA regions examined.

| Region | Inter-relative rank | | Sample size (N) | Significant value (P) | Result |
|---|---|---|---|---|---|
| | w+ | w- | | | |
| *ITS2/rpoC1* | 119 | 1 | 15 | <0.05 | *ITS2* > *rpoC1* |
| *trnH-psbA/ITS2* | 227 | 4 | 21 | <0.001 | *trnH-psbA* > *ITS2* |
| *trnH-psbA*/*rpoC1* | 136 | 0 | 16 | <0.01 | *trnH-psbA* > *rpoC1* |

## Phylogenetic inference

Barcoding studies normally use phylogenetic trees to assign species names, and the most commonly used tree is neighbor-joining, in which the assessment is based on phenetic distance and the evolutionary information of a species (Liu et al., 2014). Based on the neighbor-joining tree we constructed (Figure 2), each barcode marker successfully separated each genus by representing monophyletic clades, in which each clade appeared distinctly distant from other clades. For example, the clades containing the genera *Piper* (*P. galeatum* and *P. nigrum*) and *Murraya* (*M. koenigii* and *M. exotica*) as displayed in the *trnH-psbA* and *ITS2* phylogenetic trees are clearly distinct (Figure 2). *trnH-psbA* and *rpoC1* exhibited the highest possible identification success (100%), both at the genus and species levels. However, the identification success of *ITS2*

was only 98.7%, because the positions of some species were ambiguous. For example, based on *ITS2*, *Clerodendranthus spicatus* (HM595465) was separated from its origin, where this species was located within *Orthosiphon aristatus* (Figure 2).

**Table 5.** Prior intraspecific divergence and the number of groups for initial and recursive partition of the selected plant samples.

| Region | Prior intraspecific divergence (P) | Number of groups | |
|---|---|---|---|
| | | Initial partition | Recursive partition |
| *ITS2* | 0.001 | 13 | 13 |
| | 0.0017 | 13 | 13 |
| | 0.0028 | 13 | 13 |
| | 0.0046 | 13 | 13 |
| | 0.0077 | 13 | 13 |
| | 0.0129 | 1 | 1 |
| *rpoC1* | 0.001 | 4 | 4 |
| | 0.0017 | 4 | 4 |
| | 0.0028 | 3 | 3 |
| | 0.0046 | 3 | 3 |
| | 0.0077 | 1 | 1 |
| *trnH-psbA* | 0.001 | 6 | 6 |
| | 0.0017 | 6 | 6 |
| | 0.0028 | 6 | 6 |
| | 0.0046 | 6 | 6 |
| | 0.0077 | 6 | 6 |
| | 0.0129 | 6 | 6 |
| | 0.0215 | 3 | 3 |
| | 0.0359 | 2 | 2 |
| | 0.0599 | 2 | 2 |
| | 0.1 | 2 | 2 |

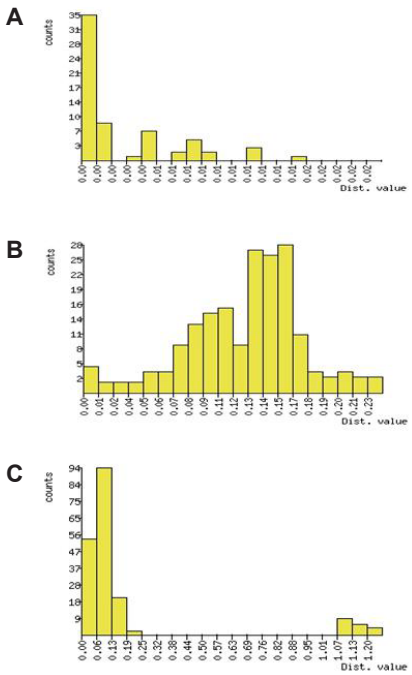


**Figure 1.** Automatic barcode gap discovery analysis of medicinal plants for (**A**) *rpoC1*, (**B**) *ITS2*, and (**C**) *trnH-psbA*.

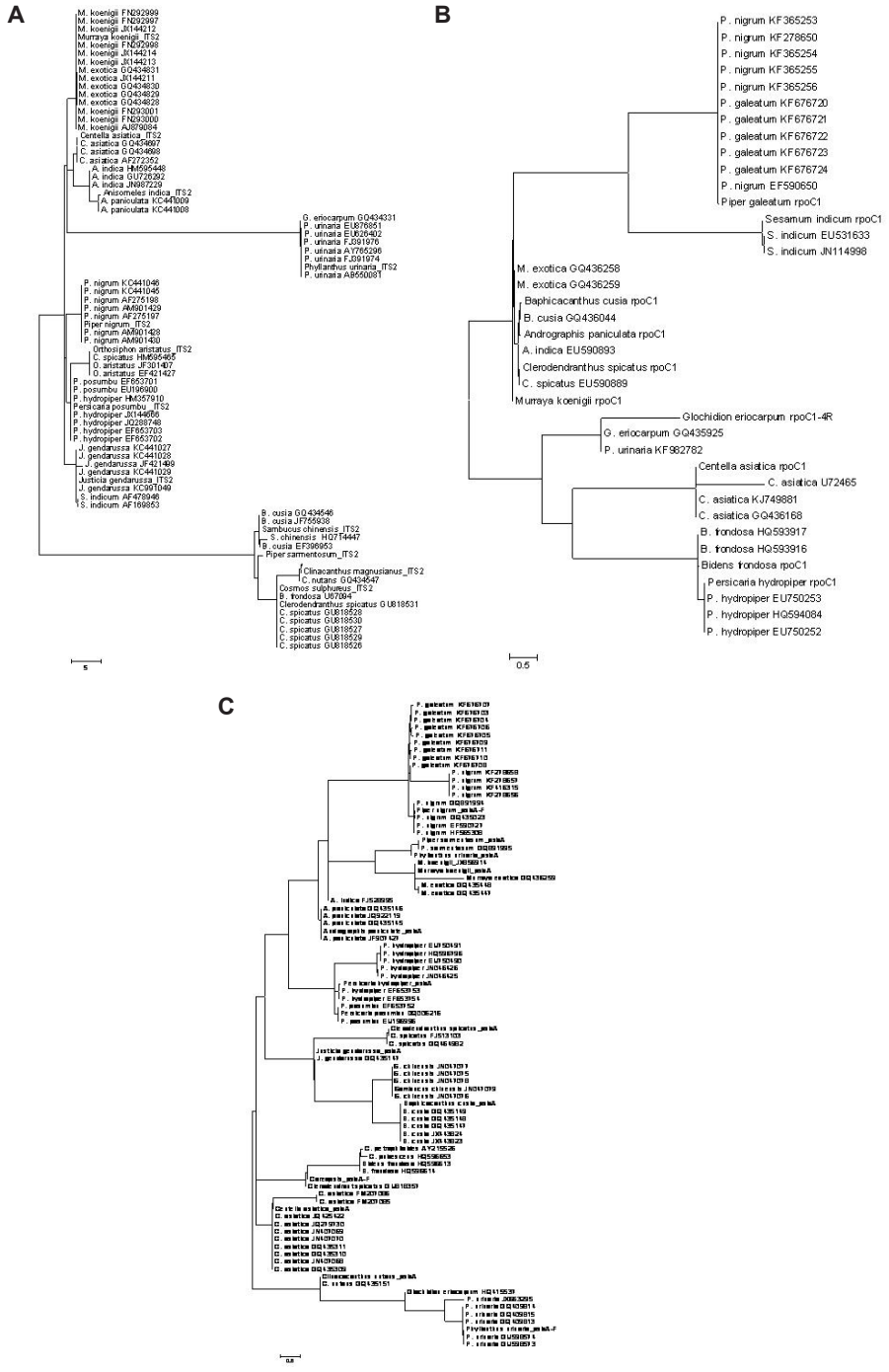


**Figure 2.** Neighbor-joining tree of medicinal plants and conspecific sequences retrieved from GenBank for (**A**) *ITS2*, (**B**) *rpoC1*, and (**C**) *trnH-psbA*.

## CONCLUSIONS

Of the three DNA regions assessed in this study, *trnH-psbA* was the best marker for discriminating between the 12 medicinal plants studied. Prior to this study, there were no documented records of these medicinal plant species in GenBank, and this is the first molecular report concerning the medicinal plants commonly used by Malay traditional healers in Malaysia, and serves as a basis for future studies regarding the conservation and management of these species.

### Conflicts of interest

The authors declare no conflict of interest.

## ACKNOWLEDGMENTS

Research supported by the Long Term Grant Scheme (Ministry of Education Malaysia), the Universiti Sains Malaysia (Local Knowledge Cluster; #203/PTS/6720007, #203/PTS/6727005, #203/PTS/6727003), and the Universiti Malaysia Kelantan (#R/LRGS/A02.00/00559A/004/2012/000089).

## REFERENCES

Castner JL (2005). Photographic atlas of botany and guide to plant identification. Feline Press, Florida.

Chen S, Yao H, Han J, Liu C, et al. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS One* 5: e8613.

Cota-Sánchez JH, Remarchuk K and Ubayasena K (2006). Ready-to-use DNA extracted with a CTAB method adapted for herbarium specimens and mucilaginous plant tissue. *Plant Mol. Biol. Rep.* 24: 161-167.

El Beyrouthy M, Kafrouny M, Arnold Na, Al-Hejin A, et al. (2013). Macroscopic, microscopic and DNA fingerprinting to fight adulteration of banafsaj (*Viola odorata* L.) sold at the Lebanese herbal shops. *Eur. J. Sci. Res.* 104: 642-651.

Farnsworth NR, Akerele O, Bingel AS, Soejarto DD, et al. (1985). Medicinal plants in therapy. *Bull. World Health Organ.* 63: 965-981.

Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, et al. (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS One* 3: e2802.

Fazekas AJ, Kuzmina ML, Newmaster SG and Hollingsworth M (2012). DNA barcoding methods for land plants. *Methods Mol. Biol.* 858: 223-252.

Gao T, Yao H, Song J, Zhu Y, et al. (2010). Evaluating the feasibility of using candidate DNA barcodes in discriminating species of the large Asteraceae family. *BMC Evol. Biol.* 10: 324.

GLOBinMED (2011). Malaysia National Medicine Policy. Available at [http://www.globinmed.com/]. Accessed April 21, 2014.

Hall TA (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41: 95-98.

Hebert PDN, Cywinska A, Ball SL and deWaard JR (2003). Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B* 270: 313-322.

Hebert PDN, Penton EH, Burns JM, Janzen DH, et al. (2004a). Ten species in one: DNA barcoding reveals cryptic species in the Neotropical skipper butterfly *Astraptes fulgerator*. *Proc. Natl. Acad. Sci. U. S. A.* 101: 14812-14817.

Hebert PDN, Stoeckle MY, Zemlak TS and Francis CM (2004b). Identification of birds through DNA barcodes. *PLoS Biol.* 2: 1657-1663.

Joharchi MR and Amiri MA (2012). Taxonomic evaluation of misidentification of crude herbal drugs marketed in Iran. *Avicenna J. Phytomed*. 2: 105-112.

Kassaye KD, Amberir A, Getachew B and Mussema Y (2006). A historical overview of traditional medicine practices and policy in Ethiopia. *Ethiop. J. Health Dev.* 20: 127-134.

Kim Sooi L and Lean Keng S (2013). Herbal medicines: Malaysia women's knowledge and practice. *Evid. Based Complement. Alt.* 2013: 438139.

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, et al. (2005). Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad.*

*Sci.* 102: 8369-8374.
Larkin MA, Blackshields G, Brown NP, Chenna R, et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.
Liu XF, Yang CH, Han HL, Ward RD, et al. (2014). Identifying species of moths (Lepidoptera) from Baihua Mountain, Beijing, China, using DNA barcodes. *Ecol. Evol*. 4: 2472-2487.
Liu Z, Zeng X, Yang D, Chu G, et al. (2012). Applying DNA barcodes for identification of plant species in the family *Araliaceae*. *Gene* 499: 76-80.
Mahadani P and Ghosh SK (2013). DNA barcoding: A tool for species identification from herbal juices. *DNA Barcodes* 1: 35-38.
Mankga LT, Yessoufou K, Moteetee AM, Daru BH, et al. (2013). Efficacy of the core DNA barcodes in identifying processed and poorly conserved plant materials commonly used in South African traditional medicine. *Zookeys* 365: 215-233.
Nwachukwu CU, Ume NC, Obasi MN, Nzewuihe GU, et al. (2010). The qualitative uses of some medicinal plants in Ikeduru L.G.A of Imo State, Nigeria. *NY Sci. J*. 3: 129-134.
Provan J, Wattier RA and Maggs CA (2005). Phylogeographic analysis of the red seaweed *Palmaria palmata* reveals a Pleistocene marine glacial refugium in the English Channel. *Mol. Ecol.* 14: 793-803.
Puillandre N, Lambert A, Brouillet S and Achaz G (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol. Ecol.* 21: 1864-1877.
Royal Botanical Gardens (2007). DNA barcoding. Available at [http://www.kew.org/barcoding/protocols.html]. Accessed April 21, 2014.
Rydberg A (2010). DNA barcoding as a tool for the identification of unknown plant material: A case study on medicinal roots traded in the medina of Marrakech. Master's thesis, Uppsala University.
Sang T, Crawford DJ and Stuessy T (1997). Chloroplast DNA phylogeny, reticulate evolution and biogeography of *Paeonia* (Paeoniaceae). *Am. J. Bot.* 84: 1120-1136.
Tamura K, Peterson D, Peterson N, Stecher G, et al. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731-2739.
Techen N, Parveen I, Pan Z and Khan IA (2014). DNA barcoding of medicinal plant material for identification. *Curr. Opin. Biotechnol.* 25: 103-110.
WHO (World Health Organization) (2002). IACR monographs on the evaluation of carcinogenic risks to humans. Vol. 82. IARC Press, France.
Yao H, Song J, Liu C, Luo K, et al. (2010). Use of ITS2 region as the universal barcode for plants and animals. *PLoS One* 5: e13102.