Kasetsart J. (Nat. Sci.) 49 : 335 - 349 (2015)

# Cryptic Biodiversity in Two Closely Related *Curcuma* (Zingiberaceae) Species in Thailand Revealed by Molecular and Morphometric Analyses

**Sutthira Khumkratok[1], Kriangsuk Boontiang[1], Prasit Chutichudet[1] and Pairot Pramual[2,*]**


## ABSTRACT

The genus *Curcuma* L. has several significant uses including in herbal medicine, as food and for decoration. However, the taxonomy of this genus has posed a challenge to traditional taxonomy due to the great morphological variation related to high levels of polyploidization and extensive hybridization. This study investigated the morphological and genetic differentiation between two closely related *Curcuma* species—*Curcuma gracillima* and *C. parviflora*—in Thailand. Principal components analysis of the 15 morphological characters revealed three morphological groups. Specimens of *C. parviflora* formed one group while those of *C. gracillima* separated into two groups. Genetic analysis based on *mat*K and internal transcribed spacer (ITS) sequences revealed greater genetic diversity in *C. gracillima* corresponding with its high morphological variation. Phylogenetic study revealed that *C. parviflora* is monophyletic; while, *C. gracillima* is polyphyletic due to samples of this species forming two distinct clades consistent with two morphologically divergent groups. Comparisons of the efficiency for discriminating species indicated that ITS is more effective than the *mat*K sequences with nearly perfect species identification (95%) compared to 58% for *mat*K sequences. Therefore, the ITS sequences should be used as the barcoding sequences, at least for *Curcuma* species.


**Keywords**: *Curcuma gracillima*, *Curcuma parviflora*, DNA barcode, Zingiberaceae, Thailand

## INTRODUCTION

The genus *Curcuma* is native to South and Southeast Asia, although some species are also found in China, Australia and the South Pacific; the number of species ranges between 50 and 100 (Smith, 1981; Sirirugsa, 1996; Larsen *et al*., 1998). In Thailand, 38 species of the genus *Curcuma* have been reported, with most found in the north and the northeast regions of the country (Maknoi, 2006). Many *Curcuma* species play a significant role for humans as they are used as food, ornamental plants and herbal medicine (Purseglove, 1974; Heywood, 1985; Majeed *et al*., 1995; Apavatjrut *et al*., 1999; Mahadtanapuk *et al*., 2006). Several studies have found medicinal properties in *Curcuma* species, such as anti-inflammatory, hepatoprotective, anti-tumor, stomachic, carminative properties and as a regenerator of liver tissue (Kumar *et al*., 2006; Farombi *et al*., 2008; Hatcher *et al*., 2008;


1 Department of Agricultural Technology, Faculty of Technology, Mahasarakham University, Maha Sarakham 44150, Thailand.

2 Department of Biology, Faculty of Science, Mahasarakham University, Maha Sarakham 44150, Thailand.

* Corresponding author, e-mail: pairot.p@msu.ac.th


Received date : 06/05/14 Accepted date : 29/12/14

Venkateshwara *et al.*, 2008; Gantait *et al*., 2011; Wilken *et al*., 2011). *Curcuma* spp. have also been used in the textile and pharmaceutical industries (Srimal and Dhawan, 1973).

Accurate species identification is critical for direct usage and further applications such as breed improvement and conservation management programs. However, these species are difficult to identify due to great morphological variation related to a high rate of interspecific hybridization and polyploidization (Záveská *et al*., 2012). Due to the difficulty of morphological identification, molecular genetic markers have been introduced for species diagnosis (Kress *et al*., 2005; Kress and Erickson, 2007). DNA barcoding is among the most commonly used for molecular based identification. This technique uses a short DNA sequence to assign an unknown specimen to a known species based on the level of genetic divergence (Hebert *et al*., 2003). Although this technique is well established in the animal kingdom by using the cytochrome *c* oxidase 1 (COI) sequences as the barcoding marker, development of DNA barcoding in plants is not as advanced. The major obstacle to plant DNA barcoding is the lack of a good-performing DNA barcoding region. The Consortium for the Barcode of Life (CBOL) proposed *rbc*L and *mat*K genes as standard barcoding sequences in plants (CBOL Plant Working Group, 2009). However, later studies found that these DNA sequences showed a low percentage of species identification success (Chen *et al*., 2010; Zhi *et al*., 2011). Thus, another DNA region (that is, internal transcribed spacers of nuclear ribosomal DNA, nrDNA ITS) was recommended as a complementary or as an alternative for plant barcoding (Chen *et al*., 2010; Hollingsworth, 2011; Zhi *et al*., 2011).

*Curcuma gracillima* Gagnep. and *C. parviflora* Wall. are important ornamental species in Thailand. Genetic variation is necessary for breed selection and improvement; thus quantifying the level of genetic variation is crucial for further study. However, there has been no reported morphological and genetic variation assessment of these species in natural populations. These species are also morphologically very similar with only a few morphological characters used for species diagnosis (Maknoi, 2006). Thus, the development of molecular markers for species identification will be useful. Therefore, the objectives of this study were to examine the morphological and genetic variations of *C. gracillima* and *C. parviflora* in Thailand and to assess the efficiency of the *mat*K and internal transcribed spacer (ITS) sequences in discriminating closely related *Curcuma* species.

## MATERIALS AND METHODS

### Collection, identification and morphological variation analysis

Specimens were collected from four areas in the northeast and lower north of Thailand (Table 1, Figures 1 and 2). Young leaves from each sample were collected for molecular study and kept in plastic bags with silica gel. Species were identified using the keys and descriptions of *Curcuma* from Wu and Larsen (2000) and Maknoi (2006). Fifteen morphological characters (Table 2) were measured from 70 specimens of *C. gracillima* and 20 specimens of *C. parviflora*. The voucher specimens were deposited in the Walai Rukhavej Botanical Research Institute, Mahasarakham University, Maha Sarakham, Thailand.

### DNA extraction, polymerase chain reaction and sequencing

Genomic DNA was extracted using the Genomic DNA extraction kit (RBC Bioscience; Xindian City, Taiwan). A fragment of approximately 700 bp of 18S rRNA+ITS1 was amplified using the primers ITS1 (5′-GTCCACTGAACCTTATCATTTAG-3′) and ITS4 (5′-TCCTCCGCTTATTGATATTATGC-3′) (White *et al*., 1990; Fischer *et al*., 2004). The *mat*K gene was amplified using the primers *mat*KF (5′-CCTATCCATCTGGAAATCTTAG-3′) and *mat*KR (5′-GTTCTAGCACAA GAAAGTCG-3′) (Yu *et al*., 2011). Polymerase chain reaction (PCR)

**Table 1** Details of the sampling locations and Genbank accession numbers of the internal transcribed spacer (ITS) and *mat*K gene for the 36 specimens of *Curcuma gracillima* and *C. parviflora* from Thailand used in this study.

| Location (code) | Species (code) | Latitude | Longitude | Altitude | Accession number | |
|---|---|---|---|---|---|---|
| | | | | | ITS | *mat*K |
| Chaiyaphum province | | | | | | |
| Pa Hin Ngam National Park (PH) | *C. gracillima* (PH1) | 15° 38' 39.00" | 101° 23' 10.00" | 792 | KF709126 | KF709147 |
| | *C. gracillima* (PH2) | 15° 39' 04.00" | 101° 23' 20.00" | 801 | KF709127 | KF709148 |
| Sai Thong National Park (ST) | *C. gracillima* (ST11) | 15° 53' 14.92" | 101° 27' 21.84" | 865 | KF709125 | KF709149 |
| | *C. gracillima* (ST12) | 15° 53' 16.06" | 101° 27' 16.51" | 865 | KF709124 | KF709150 |
| Phetchabun province | | | | | | |
| Nam Nao National Park (NN) | *C. parviflora* (NN1) | 16° 44' 22.00" | 101° 34' 25.00" | 844 | KF709136 | KF709142 |
| | *C. parviflora* (NN2) | 16° 40' 28.12" | 101° 41' 42.98" | 390 | KF709135 | KF709143 |
| | *C. parviflora* (NN3) | 16° 40' 26.88" | 101° 41' 42.16" | 391 | KF709134 | KF709156 |
| | *C. parviflora* (NN4) | 16° 40' 27.28" | 101° 41' 45.98" | 390 | KF709133 | KF709157 |
| | *C. parviflora* (NN5) | 16° 40' 26.36" | 101° 41' 44.88" | 391 | KF709132 | KF709144 |
| | *C. parviflora* (NN6) | 16° 40' 27.43" | 101° 41' 48.08" | 390 | - | KF709145 |
| | *C. parviflora* (NN7) | 16° 40' 25.34" | 101° 41' 48.90" | 390 | KF709131 | KF709146 |
| | *C. parviflora* (NN8) | 16° 40' 25.05" | 101° 41' 51.97" | 384 | - | KF709140 |
| | *C. parviflora* (NN254) | 16° 40' 24.85" | 101° 41' 47.91" | 392 | KF709137 | KF709141 |
| Ubon Ratchathani province | | | | | | |
| Kongchiam district (UB) | *C. gracillima* (UB2) | 15° 16' 17.40" | 105° 26' 53.05" | 107 | KF709123 | KF709154 |
| | *C. gracillima* (UB4) | 15° 17' 08.99" | 105° 27' 07.43" | 124 | KF709122 | - |
| Greenhouse, Mahasarakham | *C. gracillima* (GM1) | - | - | - | KF709130 | KF709155 |
| University (GM) | *C. gracillima* (GM12) | - | - | - | KF709129 | KF709151 |
| | *C. gracillima* (GM13) | - | - | - | - | KF709152 |
| | *C. gracillima* (GM14) | - | - | - | KF709128 | KF709153 |

was performed in a total reaction volume of 50 µL containing 2 ng genomic DNA, 3 mM $MgCl_2$, 0.32 mM dNTPs, 2 µM of each primer and 2 units *Taq* DNA polymerase (Vivantis; Malaysia). The temperature profile for ITS1 was: 94 °C for 4 min followed by 36 cycles of 94 °C for 30 s, 50 °C 45 s, 72 °C 1 min with a final extension at 72 °C for 7 min. The temperature profile for the *mat*K gene was: 94°C for 4 min followed by 40 cycles of 94 °C, for 30 s, 50 °C for 1 min, 72 °C for 1 min with a final extension at 72 °C for 7 min. PCR products were checked using 1% agarose gel electrophoresis and cleaned using the HiYield Gel/PCR DNA Extraction Kit (RBC Bioscience;, Xindian City, Taiwan). Cleaned PCR products were sequenced by the Macrogen (Seoul, Korea) sequencing service.

**Data analysis**

Morphological differentiations between species were tested using one-way analysis of variance. Morphological characters were also subjected to principal components analysis (PCA). The principle components (PCs) with eigenvalues greater than 1.0 were retained as variables and the PC scores of each sample were plotted against the PC axes.

DNA sequences were aligned using CLUSTAL X (Thompson *et al*., 1997), with a final visual inspection. Intraspecific and interspecific sequence divergences were calculated based on the Kimura 2-parameter separately for each gene and for the combined dataset using MEGA 5 (Tamura *et al*., 2011). To test the accuracy rate for species identification of ITS and *mat*K sequences in the dataset, the program TaxonDNA version 1.6.2 (Meier *et al*., 2006) was used for each marker sequence and the combinations of ITS and *mat*K sequences. The best match criteria (Meier *et al*., 2006) with a threshold value calculated from each dataset were employed.

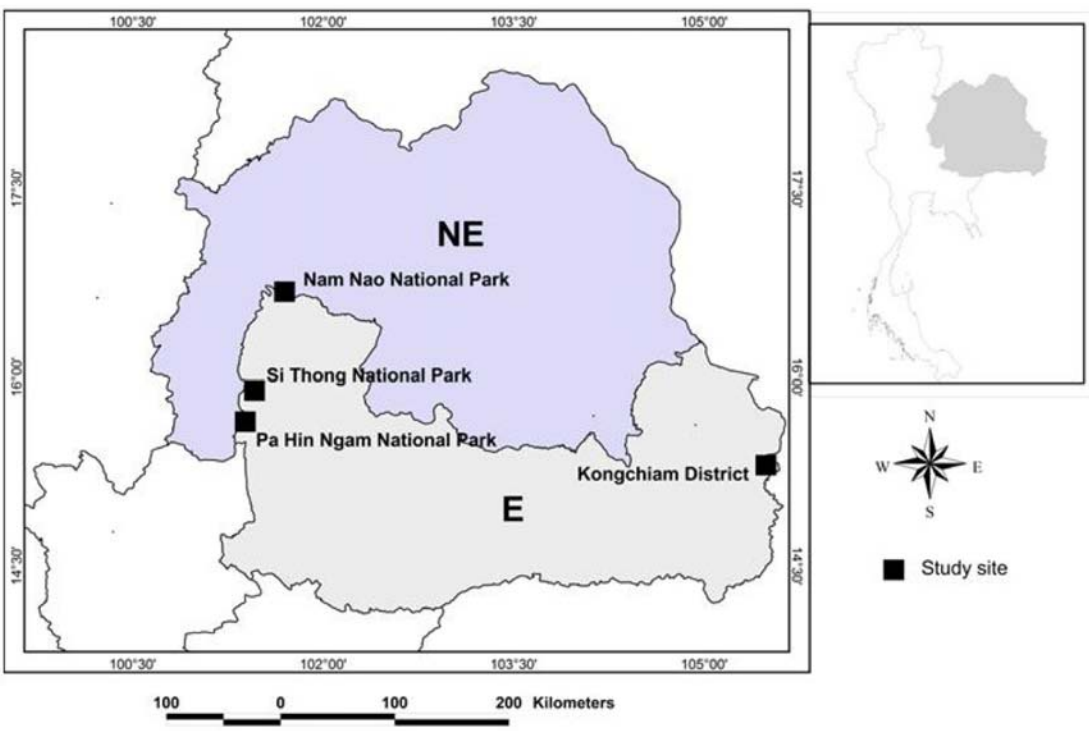


**Figure 1** Collection sites for *Curcuma gracillima* and *C. parviflora* from Thailand. (E = Eastern region and NE = Northeastern region.) Details of sampling locations are given in Table 1.

Phylogenetic analyses were conducted based on ITS and the combined data set from ITS and *mat*K sequences. The *mat*K sequences were not used for individual gene phylogenetic analysis because this gene contained very low phylogenetic signals (six parsimony informative sites). A neighbor-joining (NJ) tree was calculated in PAUP 4.10b (Swofford, 2002). Non-parametric bootstrap support was estimated using 1000 pseudoreplicates. Maximum parsimony (MP)

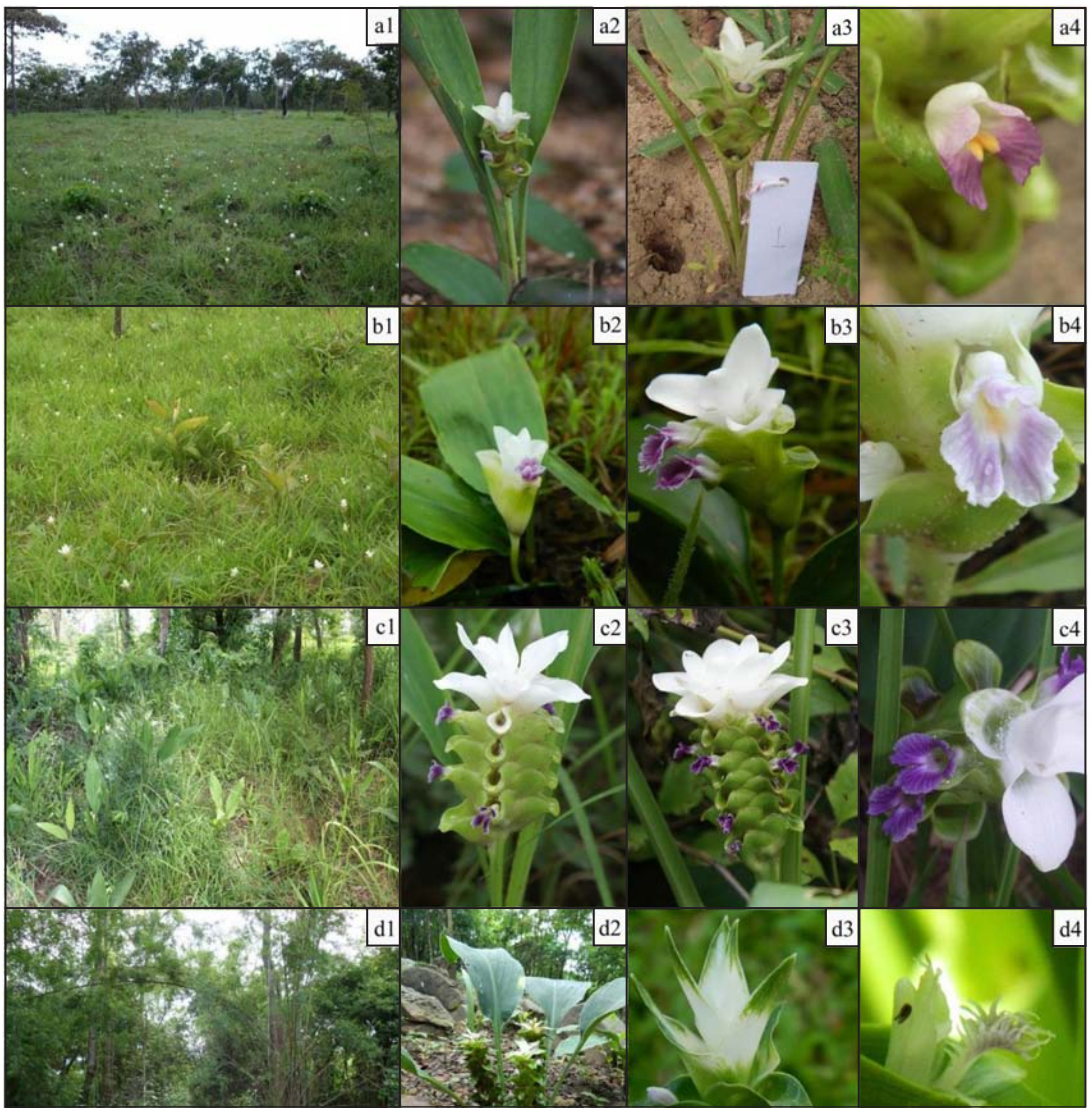


**Figure 2** Habitat, habit, bract inflorescence and floral characters of *Curcuma*: (a1–a4) *Curcuma gracillima* from Sai Thong National Park, Chaiyaphum province; (b1–b4) *C. gracillima* from Pha Hin Ngam National Park, Chaiyaphum province; (c1–c4) *C. gracillima* from Kongchiam district, Ubon Ratchathani province; (d1–d4) *C. parviflora* from Nam Nao National Park, Phetchabun Province.

**Table 2** Morphological characters of *Curcuma gracillima* and *C. parviflora* from Thailand and the results of the one-way analysis of variance among the morphological groups.

| Character | *C. gacillima* I (n = 50) | | *C. gacillima* II (n = 20) | | *C. parviflora* ( n = 20) | | *F*-test |
|---|---|---|---|---|---|---|---|
| | Min–Max | Mean (±SE) | Min–Max | Mean (±SE) | Min–Max | Mean (±SE) | |
| LW (cm) | 4.00–10.00 | 6.74$^{a}$ (0.17) | 2.00– 9.00 | 5.68$^{c}$ (0.42) | 7.00–14.50 | 11.82$^{a}$ (0.46) | 89.561** |
| LL (cm) | 8.00–29.00 | 19.46$^{a}$ (0.52) | 24.00–57.00 | 39.65$^{c}$ (1.83) | 12.00–26.00 | 19.88$^{a}$ (0.81) | 122.954** |
| LN | 1.00– 3.00 | 2.08$^{a}$ (0.49) | 2.00– 5.00 | 3.10$^{b}$ (0.12) | 2.00– 4.00 | 2.71$^{c}$ (0.14) | 36.464** |
| CBN | 5.00–13.00 | 8.44$^{a}$ (0.27) | 7.00–20.00 | 13.90$^{b}$ (0.72) | 0.00– 4.00 | 1.71$^{a}$ (0.30) | 161.996** |
| BTN | 7.00–17.00 | 11.79$^{a}$ (0.43) | 13.00–36.00 | 22.95$^{c}$ (1.42) | 7.00–13.00 | 9.23$^{a}$ (0.45) | 76.914** |
| CBW (cm) | 1.00– 2.20 | 1.72$^{a}$ (0.39) | 1.50– 3.00 | 2.31$^{b}$ (0.82) | 0.00– 0.50 | 0.27$^{c}$ (0.25) | 313.416** |
| CBL (cm) | 2.00– 3.70 | 2.82$^{a}$ (0.64) | 2.50– 4.00 | 3.26$^{b}$ (0.11) | 0.00– 1.20 | 0.82$^{c}$ (0.62) | 204.138** |
| BTW (cm) | 2.00– 3.00 | 2.61$^{a}$ (0.49) | 2.20– 4.00 | 2.86$^{a}$ (0.21) | 0.20– 1.50 | 0.88$^{c}$ (0.90) | 156.758** |
| BTL (cm) | 2.00– 3.00 | 2.59$^{a}$ (0.51) | 2.00– 4.00 | 3.11$^{b}$ (0.15) | 1.00– 4.00 | 2.42$^{c}$ (0.79) | 94.851** |
| BW (cm) | 3.00– 5.00 | 3.91$^{a}$ (0.70) | 5.00– 8.00 | 5.92$^{b}$ (0.17) | 2.70– 6.00 | 3.45$^{c}$ (0.15) | 98.173** |
| BL (cm) | 5.50–12.00 | 8.59$^{a}$ (0.23) | 10.00–15.00 | 12.27$^{b}$ (0.35) | 5.00–15.00 | 8.89$^{c}$ (0.26) | 81.572** |
| BDBH (cm) | 2.50– 5.00 | 3.96$^{a}$ (0.86) | 4.00– 8.50 | 5.76$^{b}$ (0.23) | 0.80– 4.00 | 2.38$^{c}$ (0.17) | 97.787** |
| SPH (cm) | 7.00–25.00 | 15.23$^{a}$ (0.63) | 10.00–31.00 | 20.75$^{b}$ (1.21) | 8.00–35.00 | 17.95$^{ab}$ (1.90) | 6.461** |
| LSH (cm) | 7.00–39.00 | 24.97$^{a}$ (1.19) | 40.00–91.00 | 66.00$^{b}$ (3.33) | 7.00–91.00 | 32.20$^{c}$ (2.18) | 169.157** |
| LSDBH (cm) | 0.50– 2.30 | 1.42$^{a}$ (0.69) | 1.60–4.00 | 2.57$^{b}$ (0.13) | 0.12– 1.00 | 0.49$^{c}$ (0.44) | 103.042** |

Min = Minimum; Max = Maximum.

Mean values followed by different letters are statistically different at $P < 0.05$.

** = $P < 0.001$

LW = Leaf width; LL = Leaf length; LN = Leaf number; CBN = Coma bract number; BTN = Bracteole number; CBW = Coma bract width; CBL = Coma bract length; BTW = Bracteole width; BTL = Bracteole length; BW = Bract width; BL = Bract length; BDBH = Bract diameter; SPH = Spike height; LSH = Leafy shoot height; LSDBH = Leafy shoot diameter.

analyses were performed in PAUP* using a heuristic search with 1,000 random addition sequence replicates, tree bisection-reconnection (TBR) branch swapping and the MulTrees effect. Bootstrap support was estimated for 1,000 replicates. Phylogenetic relationships were also analyzed by Bayesian methods using MrBayes v.3.20 (Huelsenbeck and Ronquist, 2001). The general time-reversible model of substitution with gamma distribution was selected. Bayesian analysis was run for $2 \times 10^6$ generations with a sampling frequency of 100 generations. Tracer v.1.3 (Rambut and Drummond, 2004) was used for visual inspection of the point where the log likelihood was stationary. Trees sampled before this point were discarded as burn-in. The remaining trees of each run were included in posterior probability calculations. All phylogenetic analyses used the sequence of *C. singularis* Gagnep. (GenBank under the accession number JQ409872 for ITS and JQ409716 for *mat*K) as the outgroup. This outgroup species was selected because a previous study found that this species belonged to the species group that was closest to the group of *C. gracillima* and *C. parviflora* (Záveská *et al.*, 2012).

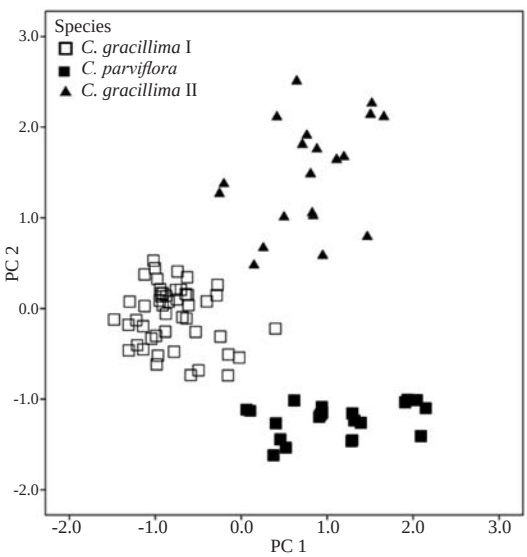


**Figure 3** Plot of the first two principal component scores (PC1 and PC2) derived from principal components analysis of 15 morphological characters among 90 specimens of *Curcuma gracillima* and *C. parviflora* from Thailand.

## RESULTS

### Morphological variations

In total, 90 specimens of *C. gracillima* (n = 70) and *C. parviflora* (n = 20) were measured for 15 morphological characters (Table 2). The average and range of each morphological character are shown in Table 2. PCA revealed that two PCs had eigenvalues higher than 1.0. The two PCs together accounted for 79.91% of the total variance. PC-1 accounted for 65.09% of the variables and explained 65.09% of the total variance. PC-2 explained 14.81% of the total variance. A plot of these two axes revealed three groups (Figure 3). Specimens of *C. parviflora* clustered together and formed a separate group. Specimens of *C. gracillima* were separated into two distinct groups (I and II). These groups were associated with geographic origins. Specimens from Pa Hin Ngam National Park and Sai Thong National Park, Chaiyaphum province were in group I and specimens from Kongchiam district, Ubon Ratchathani province were in group II.

### DNA sequence variation

In total, 38 sequences were obtained from *C. gracillima* and *C. parviflora* (Table 1). The sequence length for the ITS fragment was 652 bp and for the *mat*K gene was 754 bp. The ITS contained more variable sites (21) than the *mat*K sequences (13). Intraspecific genetic divergences (Table 3) based on ITS sequences of *C. gracillima* ranged between 0 and 2.61% with a mean of 0.70%. *C. parviflora* possessed much lower genetic diversity with ranges of intraspecific genetic divergences between 0% and 0.99% and a mean of 0.27%. Sequence divergences based on the *mat*K sequences revealed similar patterns. The

intraspecific genetic divergence for *C. gracillima* ranged between 0% and 1.59% with a mean of 0.31%. Intraspecific genetic divergence based on the *mat*K sequences of *C. parviflora* ranged from 0% to 0.66% with a mean of 0.29%.

**Phylogenetic relationship**

All three phylogenetic methods (NJ, MP and Bayesian) from both ITS and combined data revealed similar tree topologies (Figures 4 and 5); thus, only MP trees are shown. The MP tree from the ITS sequences revealed three clades (I, II and III). Specimens of *C. gracillima* from Ubon Ratchathani province and one specimen of *C. gracillima* cultivated in the greenhouse formed clade I, but with moderate support; the remaining specimens of *C. gracillima* comprised clade II but with no support. *C. parviflora* comprised clade III with moderate support. Thus, this species is monophyletic.

The combined data revealed similar tree topologies with the ITS tree. There were three clades (I, II and III). Clade I comprised specimens from Ubon Ratchathani province and one specimen of *C. gracillima* germinated in the greenhouse. The remaining specimens of *C. gracillima* formed clade II although with low support. Clade III comprised all the specimens of *C. parviflora* plus one specimen of *C. gracillima*, but with no support. All specimens of *C. parviflora* were at the derived position; thus, this species is monophyletic with strong support (Figure 5).

**DNA barcoding**

Distributions of the intraspecific and interspecific genetic divergences based on ITS, *mat*K sequences and combined data are shown in Figure 6. Sequences of the ITS region revealed a lesser extent of overlapping compared with *mat*K. The accuracy of ITS and *mat*K for DNA barcoding was assessed using the "best match" and "best close match" methods in TaxonDNA (Meier *et al*., 2006). The ITS sequences performed much better than the *mat*K sequences for differentiating *C. gracillima* and *C. parviflora*. According to the best match method, the ITS sequences were nearly completely (94.73%) successful with only one specimen that had ambiguous identification (Table 5). The *mat*K sequences could only correctly identify 57.89% of the specimens with 15.78% misidentification and 26.31% ambiguous identification. The same patterns were recovered based on the best close match method (Table 5). The combined data (ITS+*mat*K) did not increase the correct identification rate in both methods as the results revealed slightly lower percentages of correct identifications (Table 5).

## DISCUSSION

*C. gracillima* and *C. parviflora* are commercially important ornamental plants in Thailand. However, these species are difficult to identify morphologically. A single diagnostic character use to separate these species is the flower lip (Maknoi, 2006). *C. gracillima* has a deep bi-lobed labellum without a fringed edge while *C. parviflora* has a shallow lobed labellum with a fringed edge. Nonetheless, the results of the current study revealed that the two species have additional distinct morphologies. Among the 15 morphological characters in the morphometric analysis, 14 were significantly different between the two species. Among these characters, the coma bract numbers could be used to differentiate these species. *C. gracillima* had the larger size (1.0–3.0 cm) and number (5–20) than *C. parviflora* (size < 0.5 cm, number < 4).

*C. gracillima* showed greater morphological variation. Two distinct morphological groups were found. Specimens from Ubon Ratchathani province (Group I) were clearly different from other areas. Samples from this area were larger compared to *C. gracillima* from other areas (Table 2). Separation of these two morphological groups was also supported by genetic data. Phylogenetic analyses based on both ITS sequences and combined (ITS and *mat*K)

**Table 3** Means and ranges of intraspecific and interspecific genetic divergences for internal transcribed spacer (ITS) and *mat*K sequences of *Curcuma gracillima* and *C. parviflora* from Thailand.

| Species | ITS (n = 16) | | *mat*K (n = 16) | | ITS+*mat*K (n = 12) | |
|---|---|---|---|---|---|---|
| | Range of intraspecific genetic divergence (mean) % | Range of interspecific genetic divergence (mean) % | Range of intraspecific genetic divergence (mean) % | Range of interspecific genetic divergence (mean) % | Range of intraspecific genetic divergence (mean) % | Range of interspecific genetic divergence (mean) % |
| *C. gracillima* | 0–2.61 (0.70) | 0.33–1.66 (0.53) | 0–1.59 (0.31) | 0–0.11 (0.55) | 0–2.73 (0.56) | 0.15-1.31 (0.54) |
| *C. parviflora* | 0–0.99 (0.27) | | 0–0.66 (0.29) | | 0–0.52 (0.176) | |

n = Number of sequences.

**Table 4** DNA barcode statistics for the internal transcribed spacer (ITS) and *mat*K gene of *Curcuma gracillima* and *C. parviflora* from Thailand. Percentage correct identification is based on best-match methods in TaxonDNA (Meier *et al.*, 2006).

| Region | Best match | | | Best close match | | | |
|---|---|---|---|---|---|---|---|
| | Successfully identified (%) | Ambiguous (%) | Misidentified (%) | Successfully identified (%) | Ambiguous (%) | Misidentified (%) | Threshold (%) |
| ITS | 18 (94.73) | 0 | 1 (5.26) | 18 (94.73) | 0 | 1 (5.26) | 1.75 |
| *mat*K | 11 (57.89) | 5 (26.31) | 3 (15.78) | 11 (57.89) | 5 (26.31) | 3 (15.78) | 1.32 |

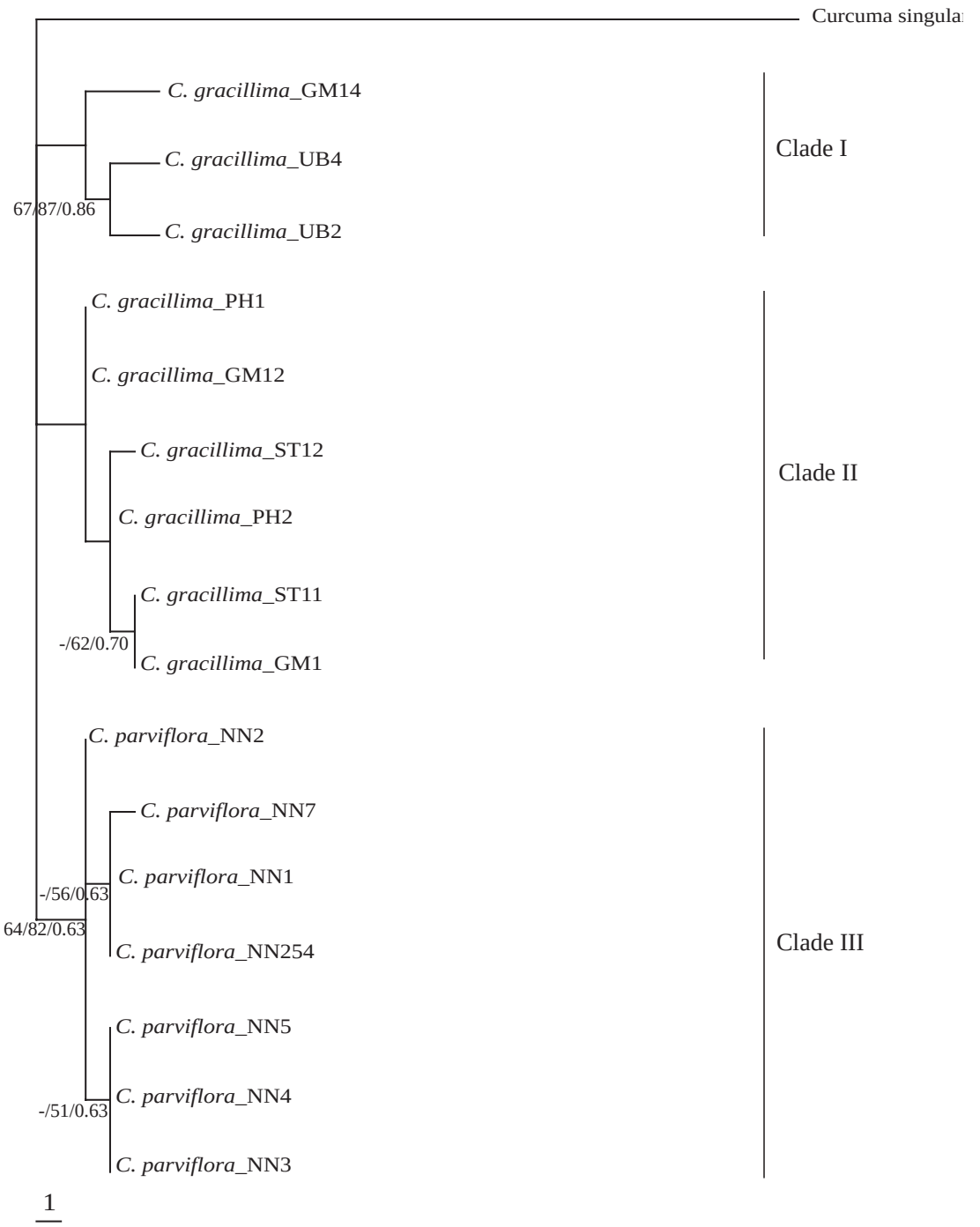


**Figure 4** Maximum parsimony tree calculated from 16 internal transcribed spacer (ITS) sequences of *Curcuma gracillima* and *C. parviflora* from Thailand. The ITS sequence of *C. harmandii* was used as an outgroup. Bootstrap support for neighbor-joining, parsimony and posterior probability based on the likelihood ratio test, is shown above or near the branch. Scale bar represents one substitution.

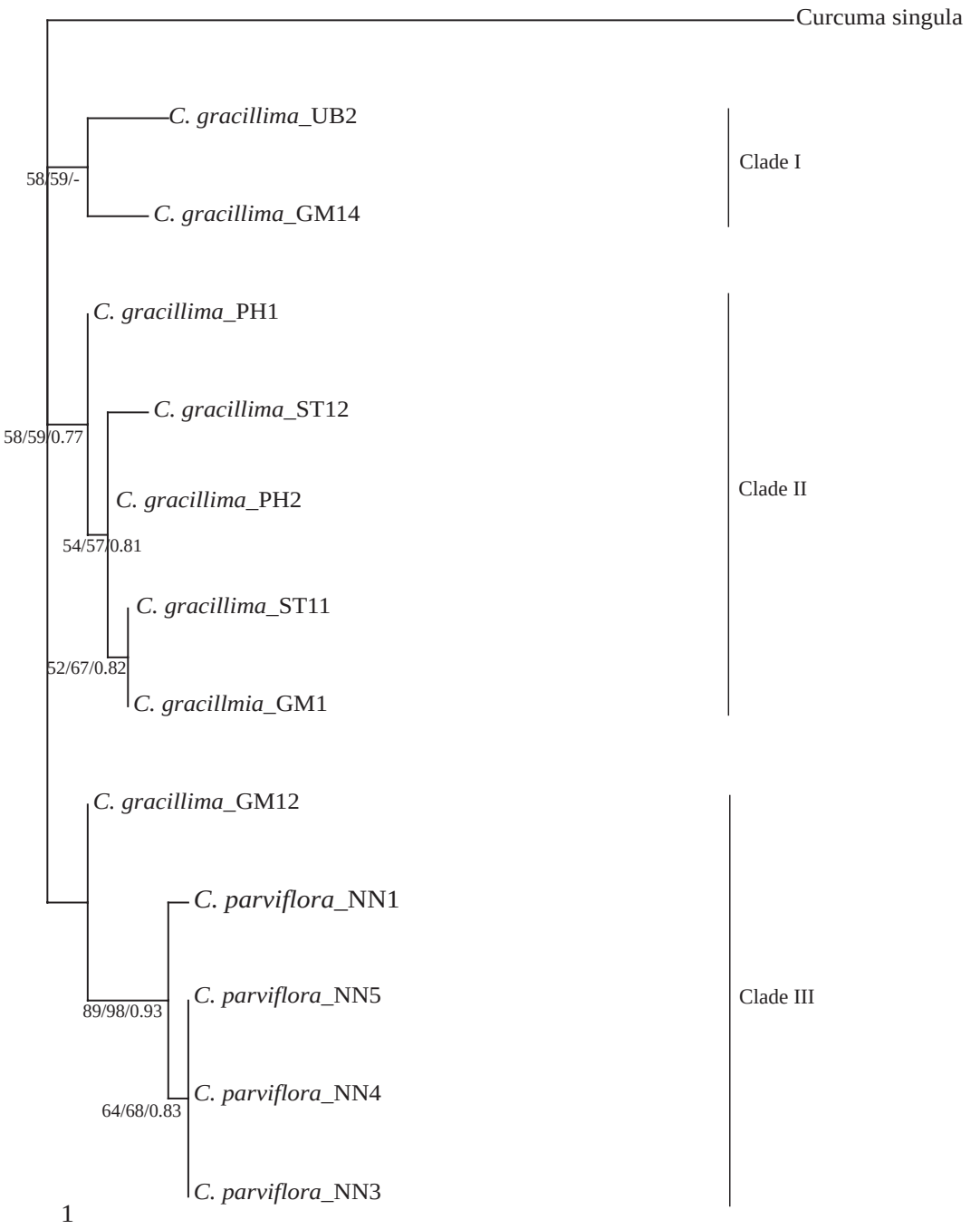


**Figure 5** Maximum parsimony tree calculated from 12 combined data (internal transcribed spacer + *mat*K) sequences of *Curcuma gracillima* and *C. parviflora* from Thailand. The sequence of *C. harmandii* was used as an outgroup. Bootstrap support for neighbor-joining, parsimony and posterior probability based on the likelihood ratio test, is shown above or near the branch. Scale bar represents one substitution.

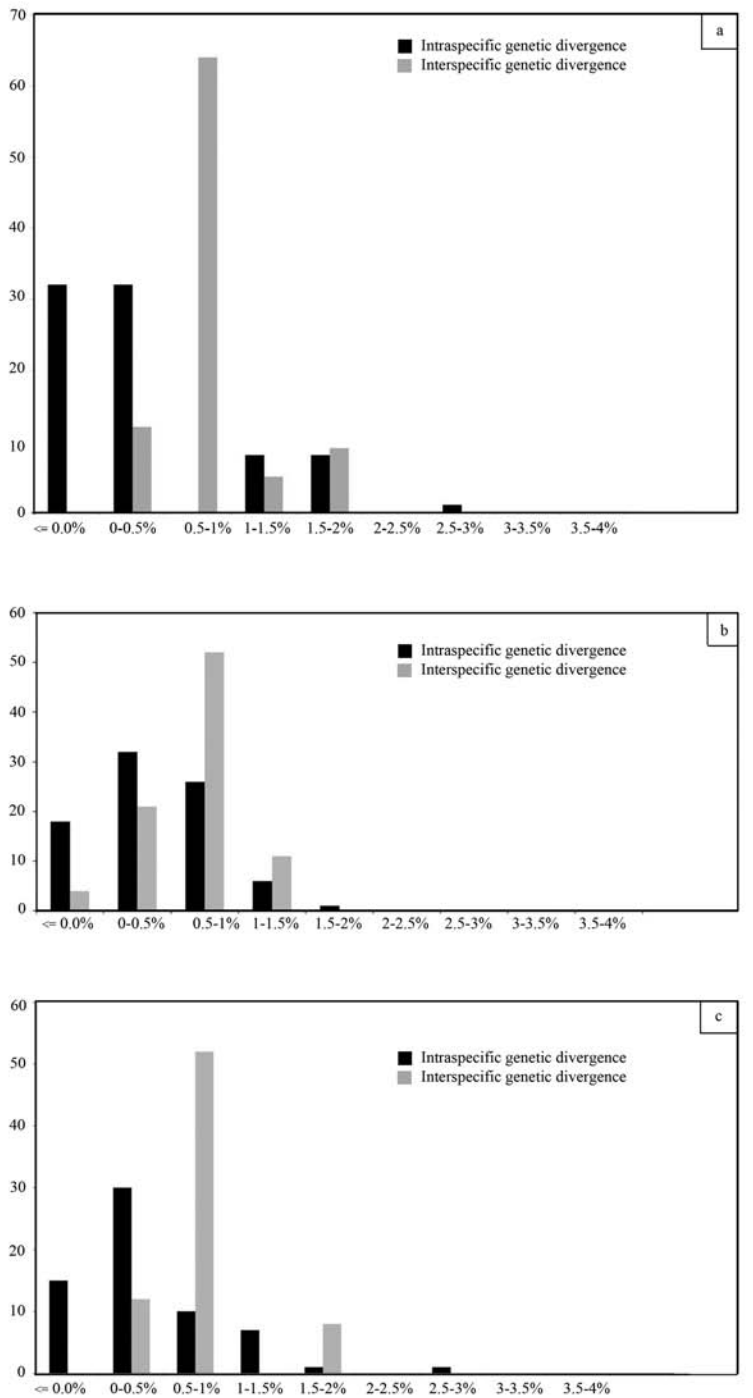


**Figure 6** Distributions of intraspecific and interspecific genetic distances based on: (a) Internal transcribed spacer sequences; (b) *mat*K sequences; (c) Combined data for *Curcuma gracillima* and *C. parviflora* from Thailand.

data found that specimens from Ubon Ratchathani province formed a distinct clade. Thus, the results strongly suggested that the specimens from Ubon Ratchathani province identified here, according to the key to species of *Curcuma* in Thailand (Maknoi, 2006), most likely represent different species. Further study should clarify this hypothesis.

Members of the genus *Curcuma* of the family Zingiberaceae are difficult to distinguish using traditional taxonomy because of the great morphological variation due to extensive species hybridization (Záveská *et al*., 2012). Thus, the development of a molecular marker for species identification is necessary for these medicinal and economically important plants. DNA barcoding has been proposed as a method for species identification using short DNA sequences (Hebert *et al*., 2003). The principal aim is to assign an unknown specimen to a species based on the level of genetic divergence (Hebert *et al*., 2003). The DNA fragments proposed as the barcoding regions in plants have recently been under debate (Hollingsworth *et al*., 2011). The Consortium for the Barcode of Life (CBOL) proposed *rbc*L and *mat*K as the barcode sequences in plants (CBOL Plant Working Group, 2009). However, later studies indicated that the internal transcribed spacer of the nuclear ribosomal DNA (nrDNA ITS) outperform the *rbc*L and *mat*K sequences for discriminating plant species. For example, Chen *et al.* (2010) compared the efficiency of the seven candidate DNA barcode regions including *psb*A-*trn*H, *mat*K, *rbc*L, *rpo*C1, *ycf*5, ITS2 and ITS in plants from 4,800 species of 753 genera. They found that the ITS2 sequences produced the highest identification success with 92.7% correct identification at the species level. Shi *et al.* (2011) compare the performance of six DNA sequence regions—ITS2, *rbc*L, *mat*K, *psb*K-*psb*I, *trn*H-*psb*A and *rpo*B—for discriminating 260 species of the family Zingiberaceae in China. They found that the ITS2 region was the most effective for species identification with 99.5 and 73.1% correct identification at the genus and species levels, respectively.

The current results support the previous finding that ITS sequences are more useful for plant species identification. The current study found that more than 94% of the specimens of the two closely related species were successfully identify based on ITS compared to 58% for *mat*K sequences. Thus, the results support ITS sequences as effective barcoding sequences for *Curcuma* species.

In conclusion, the current results indicate hidden diversity in *C. gracillima* in Thailand. Both the morphological and molecular data suggest the existence of divergent lineages in this species. Therefore, further study is needed to clarify the species status of these lineages. The ITS sequences effectively discriminated the closely related *Curcuma* species; thus, this sequence should be used in further DNA barcoding study of the genus *Curcuma*.

## ACKNOWLEDGEMENTS

This work was financially supported by a grant from Mahasarakham University (grant year 2013). The authors would like to thank Dr. Jolyon Dodgson for comments and English language improvement of an earlier draft of the manuscript.

## LITERATURE CITED

Apavatjrut, P., S. Anuntalabhochai, P. Sirirugsa and C. Alisi. 1999. Molecular markers in identification of some early flowering *Curcuma* (Zingiberaceae) species. **Ann. Bot.** 84: 529–534.

CBOL Plant Working Group. 2009. A DNA barcode for land plants. **Proc. Natl. Acad. Sci. USA.** 106: 12794–12797.

Chen, S., H. Yao, J. Han, C. Liu, J. Song, L. Shi, Y. Zhu., X. Ma, T. Gao, X. Pang, K. Luo, Y. Li, X. Li, X. Jia, Y. Lin and C. Leon. 2010.

Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. **PLoS ONE** 5: e8613.

Farombi, E.O., S. Shrotriya, H.K Na, S.H. Kim and Y.J. Surh. 2008. Curcumin attenuates dimethylnitrosamin-induced liver injury in rats through Nrf2-mediated induction of heme oxygenase-1. **Food Chem. Toxicol.** 46: 1279–1287.

Fischer, A.J., G. Omar, J. Eubanks, C.R. McGuire, B.D. Dierks and T.A. Reh. 2004. Different aspects of gliosis in retinal Muller glia can be induced by CNTF, insulin, and FGF2 in the absence of damage. **Mol. Vis.** 10: 973–986.

Gantait, A., T. Barman and P.K. Mukherjee. 2011. Validated method for estimation of curcumin in turmeric powder. **Indian J. Tradit. Knowl.** 10: 247–250.

Hatcher, H.R., R. Planalp, J. Cho, F.M. Torti and S.V. Torti. 2008. Curcumin: From ancient medicine to current clinical trials. **Cell. Mol. Life Sci.** 65: 1631–1652.

Hebert, P.D.N., A. Cywinska, S.L. Ball and J.R. deWaard. 2003. Biological identifications through DNA barcodes. **Proc. R. Soc. B.** 270: 313–321.

Heywood, V.H. 1985. **Flowering Plants of the World.** Croom Helm. London, UK. 336 pp.

Hollingsworth, P.M. 2011. Refining the DNA barcode for land plants. **Proc. Natl. Acad. Sci. U.S.A.** 108: 19451–19452.

Hollingsworth, P.M., S.W. Graham and D.P. Little. 2011. Choosing and using a plant DNA barcode. **PloS One** 6: e19254.

Huelsenbeck, J.P. and F.R. Ronquist. 2001. MRBAYES: Bayesian inference of phylogeny. **Bioinformatics** 17: 754–755.

Kress, W.J., K.J. Wurdack, E.A. Zimmer, L.A. Weigt and D.H. Janzen. 2005. Use of DNA barcodes to identify flowering plants. **Proc. Natl. Acad. Sci. USA.** 102: 8369–8374.

Kress, W.J. and D.L. Erickson. 2007. A two-locus global DNA barcode for land plants: the coding *rbc*L gene complements the non-coding *trn*H-*psb*A spacer region. **PLoS One** e508.

Kumar, S.G., H. Nayaka, S.M. Dharmesh and P.V. Salimath. 2006. Free and bound phenolic antioxidants in amla (*Emblica officinalis*) and tumeric (*Curcuma longa*). **J. Food Comp. Anal.** 19: 446–452.

Larsen, K., J.M. Lock and P.J.M. Maas. 1998. Zingiberaceae, pp. 474–495. *In* K. Kubitzki, (ed.). **The Families and Genera of Vascular Plants,** Vol. 4. Springer. Berlin, Germany.

Mahadtanapuk, S., N. Topoonyanont,T. Handa, M. Sanguansermsri and S. Anuntalabhochai. 2006. Genetic transformation of *Curcuma alismatifolia* Gagnep. Using retarded shoots. **Plant Biotech.** 23: 233–237.

Majeed, M., V. Badmaev, U. Shivakumar and R. Rajendran. 1995. **Turmeric and the Healing Curcuminoids.** Keats Publishing. New Cannan, CT, USA. 47 pp.

Maknoi, C. 2006. **Taxonomy and Phylogeny of the Genus *Curcuma* L. (Zingiberaceae) With Particular Reference to Its Occurrence in Thailand.** Ph.D. dissertation, Prince of Songkla University. Songkhla, Thailand. 179 pp.

Meier, R., S. Kwong, G. Vaidya and P.K.L. Ng. 2006. DNA barcoding and taxonomy in Diptera: A tale of high intraspecific variability and low identification success. **Syst. Biol.** 55: 715–728.

Purseglove, J.W. 1974. **Tropical Crops Monocotyledons.** Longman Group Ltd. London, UK. 719 pp.

Rambut, A. and A. J. Drummond. 2004. **Tracer 1.3.** [Available from: http://beastbioedacuk/Tracer ]. [Sourced: 27 May 2010].

Shi, L.C., J. Zhang, J.P. Han, J.Y. Song and H. Yao. 2011. Testing the potential of proposed DNA barcodes for species identification of Zingiberaceae. **J. Syst. Evol.** 49: 261–266.

Sirirugsa, P. 1996. The genus *Curcuma* of Thailand, pp: 39–46. *In* T.L. Wu, Q.G. Wu and Z.Y. Chen, (eds.). **Proceedings of the 2nd**

**Symposium on the Family Zingiberaceae.** Zhongshan University. Guangzhou, People's Republic of China.

Smith, R.M. 1981. **Synoptic Keys to the Genera of Zingiberaceae Pro Parte.** Royal Botanical Garden Edinburgh. Scotland, UK. 28 pp.

Srimal, R.C. and B.N. Dhawan. 1973. Pharmacology of diferuloyl methane (curcumin), a non-steroidal anti-inflammatory agent. **J. Pharm. Pharmacol.** 25: 447–452.

Swofford, D. L. 2002. PAUP*: **Phylogenetic Analysis Using Parsimony** (*and other methods). Version 4.b.10.: Sinauer Associate. Sunderland, MA, USA.

Tamura, K., J. Dudley, M. Nei and S. Kumar. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. **Mol. Biol. Evol.** 28: 2731–2739.

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin and D.G. Higgins. 1997. The ClustalX Windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Res.** 24: 4876–4882.

Venkateswara R.G., T. Annamalai and T. Mukhopadhyay. 2008. Nardal, a new sesquiterpene aldehyde from the plant of *Nardostachys jatamansi*. **Indian J. Chem.** 47: 163–165.

White, T.J., T. Bruns, S. Lee and J. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics, pp. 315–322. *In* M.A. Innis, D.H. Gelfand, J.J. Sninsky and T.J. White, (eds.). **PCR Protocols: A Guide to Methods and Applications.** Academic Press. New York, NY, USA.

Wilken, R., S.V. Mysore, B.W. Marilene and S.S. Eri. 2011. Curcumin: A review of anti-cancer properties and therapeutic activity in head and neck squamous cell carcinoma. **Mol. Cancer Res.** 10: 1–19.

Wu, T.L. and K. Larsen. 2000. Zingiberaceae, pp. 359–362. *In* Z.Y. Wu and P.H. Raven, (eds.), **Flora of China.** Science Press. Beijing, People's Republic of China.

Yu, J., J.H. Xue and S.L. Zhou. 2011. New Universal *mat*K primers for DNA barcoding angiosperms. **J. Syst. Evol.** 49: 176–181.

Záveská, E, T. Fér, O. Šída, K. Krak, K Marhold and J. Leong-Škorničková. 2012. Phylogeny of *Curcuma* (Zingiberaceae) based on plastid and nuclear sequences: Proposal of the new subgenus *Ecomata*. **Taxon** 61: 747–763.

Zhi, D. Y., A. Qimike, C.F. Yang, J.F. Chen and Q.F. Wang. 2011. Testing four barcoding markers for species identification of Potamogetonaceae. **J. Syst. Evol.** 49: 246–251.