*Agr. Nat. Resour. 56 (2022) 1091–1102*

AGRICULTURE AND NATURAL RESOURCES

Journal homepage: http://anres.kasetsart.org

Research article

# Molecular systematics of notable *Nymphaea* cultivars in Thailand based on multiple chloroplast DNA specific sites

**Chantana Khensri, Piriya Putanyawiwat, Vipa Hongtrakul***

*Department of Genetics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand*

## Article Info

**Article history:**
Received 4 May 2022
Revised 15 August 2022
Accepted 23 September 2022
Available online 19 December 2022

**Keywords:**
Chloroplast DNA,
Molecular systematics,
*Nymphaea*,
Waterlily

## Abstract

**Importance of the work**: The waterlily is a flowering plant in the genus *Nymphaea*, family Nymphaeaceae. It has an exceptionally beautiful and colorful flower. Recently, many new interspecific and intersubgeneric hybrids have been introduced; therefore, genetic information at the molecular level is needed.

**Objectives**: To generate DNA sequences for three plastid DNA regions (*rbc*L, *mat*K and *trn*H-*psb*A intergenic spacer) of elite cultivars to support cultivar classification and identification of waterlily.

**Materials & Methods:** The three plastid DNA regions from 29 notable cultivars were sequenced and then the data were analyzed using the BioEdit and MEGA (version 6.0) programs. Phylogenetic trees were constructed using the maximum likelihood method.

**Results**: The *trn*H-*psb*A intergenic spacer region had the highest sequence polymorphism, followed by the *mat*K and *rbc*L regions, respectively. A phylogenetic tree based on all three regions was consistent with the previous classification based on morphological characters. Three clearly defined clusters were identified, with Cluster I belonging to subg. *Nymphaea,* Cluster II belonging to subg. *Lotos* and Cluster III belonging to subg. *Brachyceras* and *Anecphya.* Five waterlily cultivars (*N.* 'Miss Siam', *N.* 'Mangala Ubol', *N.* 'Rojjana Ubol', *N.* 'Tuonta' and *N.* 'Chongkolnee') could be separated from other waterlily cultivars using the three gene regions combined.

**Main finding**: These DNA sequences markers should be useful as a tool for specific identification and for registration of future new releases of notable *Nymphaea* cultivars.

* Corresponding author.
E-mail address: fscivph@ku.ac.th (V. Hongtrakul)



https://doi.org/10.34044/j.anres.2022.56.6.04

## Introduction

Waterlilies, known as the 'Queen of aquatic plants', are in the family Nymphaeaceae and can be divided into six genera: *Victoria, Euryale, Nuphar, Barclaya, Ondinea* and *Nymphaea* (La-ongsri and Wattatana, 2010). Waterlilies in the genus *Nymphaea* can be categorized by origin into two major groups: tropical and hardy, with the native tropical waterlily found in Thailand, 'Chongkolnee' (*Nymphaea siamensis*), having large flowers with double petals that make its flower look like a chrysanthemum (Wasuwat and Wasuwat, 2004). Hardy waterlilies are native plants in temperate and frigid zones, with *N. mexicana* being a native hardy waterlily distributed in the USA in Florida, Texas, Georgia, Alabama, Louisiana, and in Mexico (Conard, 1905; Woods et al., 2005). A monograph of the genus *Nymphaea* was created by Conard (1905) and was confirmed based on a molecular technique by Borsch et al. (2007). Based on Conard (1905), the genus *Nymphaea* can be divided into five subgenera: *Hydrocallis* (American tropical night-blooming waterlily, 14 species); *Lotos* (tropical night-blooming waterlily, 3–4 species); *Nymphaea* (hardy waterlily, 8 species); *Brachyceras* (tropical day-blooming waterlily, 14–16 species); and *Anecphya* (Australian giant, tropical day-blooming waterlily, 7–10 species).

The waterlily has been introduced in Thailand for ornamental plant cultivation and to create new hybrids, many of which have been developed in Thailand, with *N.* 'Siam Blue Hardy', being an outstanding Thai intersubgeneric hybrid, having a blue-purple flower, which is a color that has never been seen in hardy waterlily previously (Songpanich and Hongtrakul, 2010). *N.* 'Mangala Ubol' is an intrasubgeneric hybrid that is considered the best hardy waterlily in the world and so it is in demand in international markets. It won the first prize in the International Waterlily and Water Gardening Society (IWGS) awards in 2004 (IWGS, 2016). Therefore, Thailand has been accepted as one of the outstanding, famous waterlily-producing countries that also has many native waterlilies. For example, *N.* 'Chongkolnee' is a beautiful Thai native waterlily that cannot now be found naturally and is in danger of extinction (Wasuwat and Wasuwat, 2004) and thus requires human intervention for its sustainable conservation. Genetic information at the molecular level of notable waterlily cultivars in Thailand is still limited. Study at the molecular level will provide reliable and detectable information for cultivar-specific identification and for hybrid detection which was previously considered based on co-morphological traits between parents.

Chloroplasts are organelles that conduct photosynthesis and the genome contains double stranded circular DNAs. Due to their low evolutionary rate (Chase et al., 2007), chloroplast genomes are popular as a tool for taxonomic study or for investigating phylogenetic relationships in plants. The *rbc*L and *mat*K genes in the chloroplast genome have become universal regions that are the most suitable for nucleotide sequencing and gene comparison (CBOL Plant Working Group, 2009). The *rbc*L DNA of about of 599 bp which is located at the 5' end of the gene (1,430 bp), is commonly used in studies, while the region of the *mat*K gene commonly used for study, is in the middle region of about 841 bp (position: 205–1,046 bp) which is in the intron of the *trn*K gene (Hollingsworth et al., 2011). The *mat*K gene has the highest evolutionary rate among chloroplast genes (Hilu and Liang, 1997). CBOL Plant Working Group (2009) suggested that using two or more universal regions will provide a more accurate result than using only one region for specific identification. In addition, the *trn*H-*psb*A intergenic spacer, non-coding region, which has high variation in plants (Shaw et al., 2007), is suitable for studying phylogenetic relationships and as a supplementary region that provides a clear result.

At the present, nucleotide sequencing of the chloroplast genome has been undertaken in various plants, such as *Nicotiana tabacum* (Ohyama et al., 1986), liverworts (Shinozaki et al., 1986) *Arabidopsis thaliana* (Dobrogojski et al., 2020) and *Nymphaea* plants, such as *N. mexicana* (Yang et al., 2014), *N. capensis* (Kim et al., 2019b), *N. lotos* (Kim et al., 2019a), *N. odorata*, *N. rubra*, *N. gigantea*, *N. potamophila*, *N. tetragona*, *N. colorata*, and *N. micrantha* (Sun et al., 2021). For specific region study in *Nymphaea* plants, the *trn*T-*trn*F region can be used to classify *Nymphaea* plants into three groups (Borsch et al., 2007): subg. *Nymphaea*, subg. *Brachyceras-Anecphya* and thirdly subg. *Hydrocallis* and subg. *Lotos*. However, the relationships between plants in these group are still not explicit. The internal transcribed spacer (ITS) region, the *trn*K intron and the *mat*K gene have been used for phylogenetic relationship study in subg. *Lotos* in India (Dkhar et al., 2012), where it was found that *N. petersiana* and *N. lotos* were closely related, while *N. pubescens* and *N. rubra* were classified into another group. Woods et al. (2005) succeeded in the classification of *N. odorata* subsp. *odorata* and *N. odorata* subsp. *tuberose* by analysis of the *rbc*L, *mat*K and *trn*H-*psb*A intergenic spacer regions. The analysis of genetic relationships and phylogenetic tree construction among notable *Nymphaea* cultivars in Thailand would support their classification and specific cultivar identification of waterlilies grown in Thailand.

## Materials and Methods

### *Plant materials*

In total, 29 waterlily cultivars grown in Thailand were used. Some samples were collected from the 'Pang U bon' Waterlily Garden and the Waterlily Germplasm Bank at the Rajamangala University of Technology Tawan-ok, Thailand. Some hybrid samples were obtained directly from Mr Pairat Songpanich, a waterlily hybridizer. Samples were obtained covering the native waterlily, hardy and tropical hybrids and Australian giant waterlily groups. Some cultivars had won the title of best new waterlily in the world. Some cultivars have interesting characters; for example, *N.* 'Chongkolnee' which is native to Thailand, is sterile and is considered unusual and rare, dating back more than 700 years ago. The 29 notable cultivars were in four subgenera: *Nymphaea*, *Brachyceras, Anecphya* and *Lotos* (Table 1).

**Table 1** Waterlilies used in this study

| No. | Cultivar name | Subgenus | Award/Characteristics | Source |
|---|---|---|---|---|
| 1 | *N.* 'Mangala Ubol' | *Nymphaea* | 2004 IWGS 1st place hardy waterlily | Pang U bon |
| 2 | *N.* 'Ploi Praow' | *Lotos* | 2010 IWGS 2nd best new tropical night–blooming waterlily | Pang U bon |
| 3 | *N.* 'Wanvisa' | *Nymphaea* | 2010 IWGS best new waterlily and best new hardy waterlily/ variegated flower with tones of orange, pink and yellow plus a heavily marbled leaf | Pang U bon |
| 4 | *N.* 'Chalong-Kwan' | *Brachyceras* | Most double-flowered waterlily created currently, with sterile and fragrant flowers. | Pang U bon |
| 5 | *N.* 'Ploi Jear' | *Lotos* | 2011 IWGS 2nd place night-blooming waterlily | Pang U bon |
| 6 | *N.* 'Chongkolnee' | *Brachyceras* | Native to Thailand, sterile and considered unusual and rare dating back more than 700 years during the Sukhothai period. | Pang U bon |
| 7 | *N.* 'Siam Purple1' | Hybrid (*Nymphaea* x *Brachyceras*) | 2011 IWGS best new waterlily and 1st place intersubgeneric waterlily | Pang U bon |
| 8 | *N.* 'Tan-khwan' | *Nymphaea* | 2006 IWGS best new hardy waterlily | Pang U bon |
| 9 | *N.* 'Suwanna' | *Brachyceras* | 2007 IWGS best new waterlily and best new tropical waterlily | Pang U bon |
| 10 | *N.* 'Rattana Ubol' | *Nymphaea* | 2009 IWGS best new hardy waterlily/pink white flowers with petal numbers of 34–45 and free flowering | Pang U bon |
| 11 | *N.* 'Supranee Pink' | *Nymphaea* | Pink flowers with petal numbers of 36–49 and free flowering | Pang U bon |
| 12 | *N.* 'Miss Siam' | *Nymphaea* | 2007 IWGS best new hardy waterlily | Pang U bon |
| 13 | *N.* 'Rojjana Ubol' | *Lotos* | 2011 IWGS 1st place night-blooming waterlily | Pang U bon |
| 14 | *N. gigantea* (blue) | *Anecphya* | Large cup-shaped flower and blue-purple petals | Pang U bon |
| 15 | *N.* 'Maeploi' | *Lotos* | 2010 IWGS best new tropical night-blooming waterlily/small bright red flowers | Pang U bon |
| 16 | *N.* 'Queen Sirikit' | Hybrid (*Nymphaea* x *Brachyceras*) | Petals white and purple at tip | Pang U bon |
| 17 | *N.* 'Siam Pink' | Hybrid (*Nymphaea* x *Brachyceras*) | 2010 IWGS best new intersubgeneric waterlily | Pang U bon |
| 18 | *N.* 'Tanpong' | *Nymphaea* | 2011 IWGS 2nd place hardy waterlily | Pairat Songpanich |
| 19 | *N.* 'Pink Ribbon' | *Nymphaea* | Named in support of breast cancer awareness because of the similarity to breast cancer ribbons. | Pairat Songpanich |
| 20 | *N.* 'Siam Purple2' | Hybrid (*Nymphaea* x *Brachyceras*) | 2014 IWGS 1st place best new intersubgeneric (ISG) waterlily /2nd best new waterlily/1st place intersubgeneric waterlily people choice awards | Pairat Songpanich |
| 21 | *N.* 'Siam Beauty' | *Nymphaea* | 2014 IWGS 2nd place best new hardy waterlily | Pairat Songpanich |
| 22 | *N.* 'Mayaranee' | *Anecphya* | 2013 IWGS 1st place *Anecphya* waterlily/petals of flowers change color during blooming from blue to light blue, purple and pink | Rajamangala U. of Technology Tawan-ok |
| 23 | *N.* 'Yaowalak' | *Anecphya* | 2013 IWGS 2nd place *Anecphya* waterlily/petals change color during 5 days blooming from blue to red-pink | Rajamangala U. of Technology Tawan-ok |

**Table 1** Continued

| No. | Cultivar name | Subgenus | Award/Characteristics | Source |
|---|---|---|---|---|
| 24 | *N.* 'Tuonta' | Hybrid (*Anecphya* x *Brachyceras*) | 2011 IWGS 2nd best new waterlily and 2nd place intersubgeneric waterlily/a fabulous bicolor waterlily with parentage from Australia | Rajamangala U. of Technology Tawan-ok |
| 25 | *N.* 'Chularat' | *Brachyceras* | 2011 People's Choice awards: 2nd place tropical day-blooming waterlily | Rajamangala U. of Technology Tawan-ok |
| 26 | *N.* 'See Som Tawan–ok' | *Brachyceras* | 2013 IWGS 2nd place tropical waterlily/beautiful orange flowers | Rajamangala U. of Technology Tawan-ok |
| 27 | *N.* 'Siam Nymph' | *Nymphaea* | 2011 People's Choice awards: 1st place hardy waterlily | Pang U bon |
| 28 | *N.* 'Joey Tomocik' | *Nymphaea* | Strongest and most vivid color of all yellow lilies available/ very free flowering | Pang U bon |
| 29 | *N.* 'Plai Fah' | *Brachyceras* | 2015 IWC 2nd place tropical new waterlily | Pang U bon |

Source: IWGS (2016); Songpanich (2016)

### *DNA extraction and amplification*

Total genomic DNA was extracted from fresh petals or young leaves following a CTAB method modified by Doyle and Doyle (1990). Total DNA was estimated using agarose gel electrophoresis and measured using a nanodrop spectrophotometer. Genomic DNA was diluted to 100 ng/µL before using for polymerase chain reaction (PCR) amplification. DNA fragments of three plastid DNA locations (the *rbc*L and *mat*K and *trn*H-*psb*A intergenic spacer regions) were amplified using the universal primers presented in Table 2.

The PCR reaction in the total of 30 µL contained 13.5 µL of sterile distilled water, 1× phusion HF buffer, 1.5 mM $MgCl_2$, 0.2 mM each dNTP, 0.5 µM of each primer, 0.6 units of phusion *Taq* DNA polymerase (Thermo Scientific, USA) and 120 ng of purified total DNA. The PCR thermal cycling program consisted of 30 s at 98 °C for pre-denaturing, 10 s at 98 °C for denaturing, 20 s at 52 °C for annealing and 30 s at 72 °C for extension. The denaturing to extension steps were repeated for 34 cycles. The reactions were completed at 72 °C for 5 min and held at 4 °C. All the PCR products were visualized using electrophoresis on 1.2% agarose gel in 1× tris-borate-ethylenediaminetetraacetic acid (TBE) buffer.

### *Sequencing strategy*

The PCR products of the *rbc*L and *mat*K and *trn*H-*psb*A intergenic spacer regions were cut from agarose gel and purified using a FavorPrep™ Gel/PCR Purification Kit (Favorgen, Taiwan). The purified PCR products were estimated for quality and quantity based on electrophoresis on 1.2% agarose gel in 1× TBE buffer before being sent for sequencing by 1st Base Company, Malaysia.

### *Sequence analysis, alignment and phylogenetic analysis*

The obtained sequences were confirmed for the three specific genes by comparison to the database in GeneBank (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Multiple sequences of each gene region were aligned using Clustal W in the Bioedit program and adjusted manually. Phylogenetic trees were constructed using the MEGA program (version 6.0; Tamura et al. (2013); https://www.megasoftware.net). Genetic distances and relationships were analyzed using the maximum likelihood (ML) method. To reinforce the position of branches in the phylogenetic tree, the bootstrap value was used 1,000 times and values were shown next to the branches.

**Table 2** Universal primers used in this study

| Region | Primer name | Sequence (5'–3') | Direction | Reference |
|---|---|---|---|---|
| *trn*H–*psb*A intergenic spacer | trnH2 | CGCGCATGGTGGATTCACAATCC | F | Tate and Simpson (2003) |
| | psbAF | GTTATGCATGAACGTAATGCTC | R | Sang et al. (1997) |
| *rbc*L | rbcLa–F | ATGTCACCACAAACAGAGACTAAAGC | F | Levin et al. (2003) |
| | rbcLa–R | GTAAAATCAAGTCCACCRCG | R | Kress et al. (2009) |
| *mat*K | matK–KIMIR | ACCCAGTCCATCTGGAAATCTTGGTTC | F | Fazekas et al. (2012) |
| | matK–KIM3F | CGTACAGTACTTTTGTGTTTACGAG | R | Fazekas et al. (2012) |

## Results and Discussion

### *Specific amplification DNA and nucleotide analysis*

DNA fragments with the expected sizes of the three specific gene regions (*rbc*L, *mat*K and *trn*H-*psb*A intergenic spacer) were successfully amplified in the 29 waterlily samples (Fig. S1). The PCR products of the *rbc*L, *mat*K and *trn*H-*psb*A intergenic spacer regions were in the ranges 554–581 bp, 815–940 bp and 531–574 bp, respectively. All 29 DNA sequences of each specific gene region were compared to the nucleotide database in Genebank to confirm the specific gene. The results indicated that the 29 nucleotide sequences of the 29 waterlily samples in the *rbc*L, *mat*K and *trn*H-*psb*A intergenic spacer regions were similar to the closest sequences of the three specific gene regions from the database with identity levels of 99%, 99–100% and 88–99%, respectively. Sequences of *N.* 'Mangala Ubol' at *rbc*L (534 bp), *mat*K (842 bp) and the *trn*H–*psb*A intergenic spacer (564 bp) region are shown in Figs. 1A–1C, respectively.

### *Nucleotide sequence analysis of rbcL*

From the nucleotide comparisons among the 29 waterlilies using the BioEdit program, 10 positions of variation were found in the 534 bp DNA sequence of the *rbc*L region (Table 3 and Fig. 2). All DNA sequences in the *rbc*L region were used to generate a phylogenetic tree using the MEGA program (version 6.0) and the ML procedure. The obtained genetic distance values were in the range 0.000–0.013. The results from phylogenetic tree based on the *rbc*L sequences could be used to separate two notable waterlilies cultivars (*N.* 'Mangala Ubol' and *N.* 'Rojjana Ubol') from other elite cultivars. The phylogenetic tree was consistent with morphological characters (Conard, 1905) which clearly clusters were defined: subg. *Nymphaea* cluster, subg. *Lotos* cluster and subg. *Brachyceras-Anecphya* cluster (Fig. 5A).

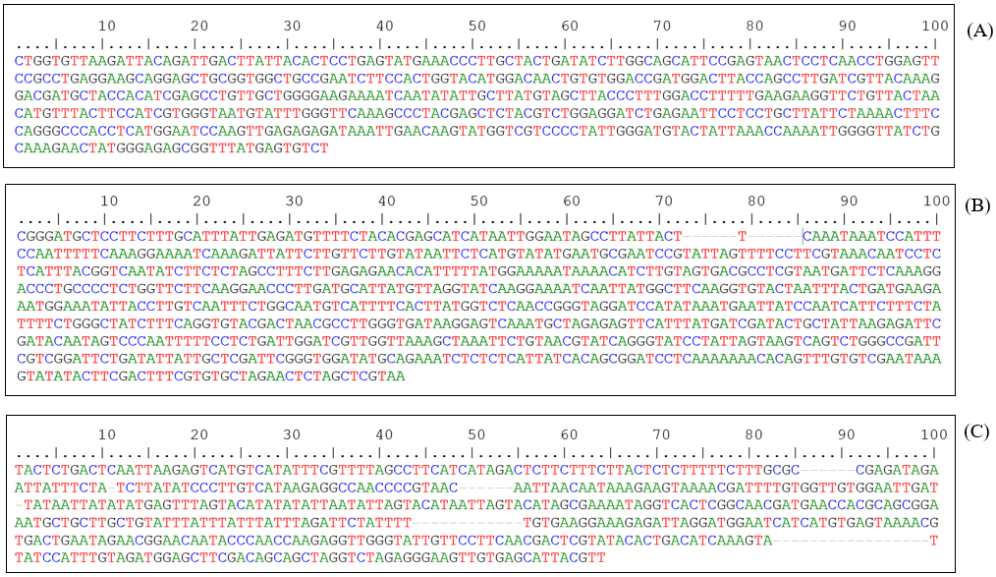


**Fig. 1** DNA sequences of *N.* 'Mangala Ubol' in three gene regions: (A) *rbc*L; (B) *mat*K; (C) *trn*H-*psb*A intergenic spacer, where number above nucleotide sequence is nucleotide position based on sequence of *N.* 'Mangala Ubol' after multiple sequence alignment of 29 waterlilies and gaps in sequences indicate the gap position for the best alignment.

**Table 3** Sequence variation found in three gene regions (*rbc*L, *mat*K and *trn*H-*psb*A intergenic spacer) of 29 waterlilies

| Variation pattern | Nucleotide position* | | |
|---|---|---|---|
| | *rbc*L gene | *mat*K gene | *trn*H–*psb*A intergenic spacer |
| Indel | - | 73–78, 80–91 | 86–91, 107–112, 149–154,187–192, 201, 229–241, 249–261, 294–302, 325–332, 344–355, 416–420 and 483–499 |
| Purine transition | 235, 283 and 409 | 3, 284, 324, 566, 589, 636 and 705 | 94,167, 268, 283, 375, 389 and 415 |
| Pyrimidine transition | 265, 316 and 499 | 42, 330, 414, 468, 518 and 771 | 34 and 73 |
| Transversion | 40, 103, 328 and 370 | 55, 56, 79, 104, 242, 261, 299, 352, 363, 420, 421, 474, 500, 519, 585, 648 and 825 | 66, 97, 99, 104, 113, 159, 173, 198, 221, 222, 273, 278, 337, 338, 357 and 431–434 |

* Nucleotide position based on sequence of *N.*'Mangala Ubol' after multiple sequence alignment of 29 waterlilies

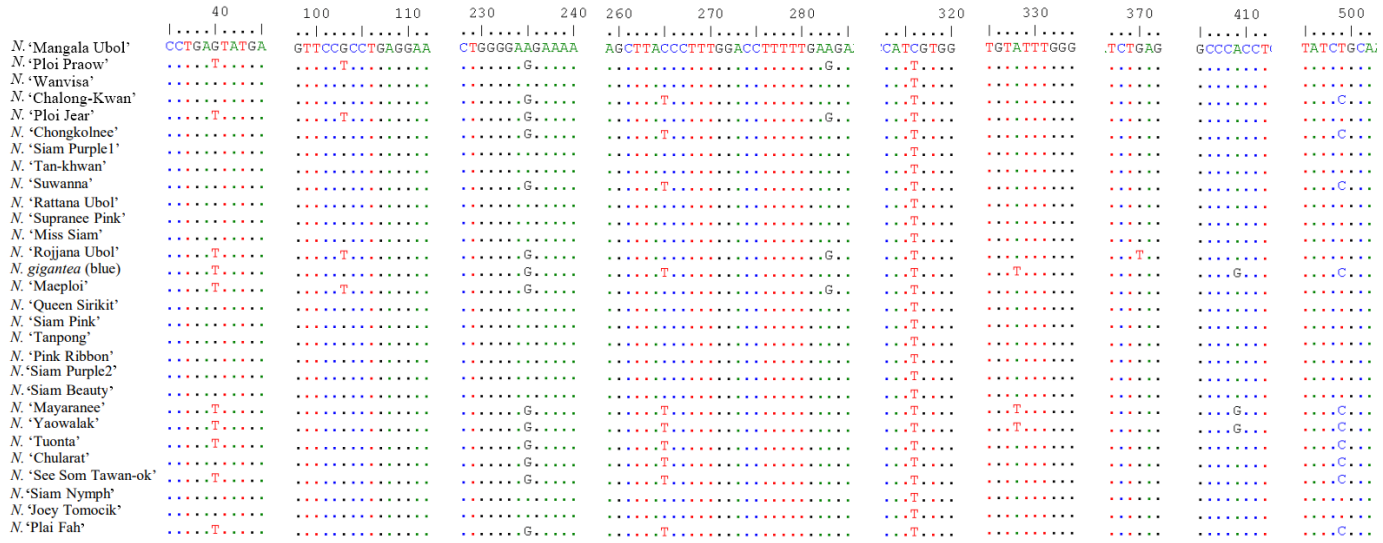


**Fig. 2** Variable position indicating base transition and transversion of *rbc*L gene sequences after multiple sequence alignment of 29 waterlilies

### *Nucleotide sequence analysis of matK*

In total, 48 variation positions were found in the 842 bp DNA sequence of the *mat*K region of the 29 waterlily samples (Table 3 and Fig. 3). All DNA sequences in this specific region were used to construct a phylogenetic tree. Genetic distance values were in the range 0.000–0.019 and five notable waterlily cultivars (*N.* 'Miss Siam', *N.* 'Mangala Ubol', *N.* 'Chongkolnee', *N.* 'Tuonta' and *N.* 'Rojjana Ubol') could be separated from the other elite cultivars using these *mat*K sequences. Th phylogenetic tree was consistent with morphological characters (Conard, 1905) for which three clusters were defined: Cluster I belonging to subg. *Nymphaea,* Cluster II belonging to subg. *Brachyceras-Anecphya* and Cluster III belonging to subg. *Lotos* (Fig. 5B).

### *Nucleotide sequence analysis of trnH-psbA intergenic spacer*

Variations in 130 positions were found among the *trn*H-*psb*A sequences, 564 bp in length, in the 29 waterlily samples (Table 3 and Fig. 4). All *trn*H-*psb*A sequences were used to generate a phylogenetic tree (Fig. 5C). Genetic distance values were in the range 0.000–0.048. The nucleotide sequence in the *trn*H-*psb*A region could be used to separate three notable cultivars (*N.* 'Mangala Ubol', *N.* 'Rojjana Ubol' and *N.* 'Tuonta') from the other elite cultivars. The phylogenetic tree was consistent with the previous classification based on morphological characters (Conard, 1905) which was the same as the trees constructed from the *rbc*L and *mat*K gene regions. The three clearly defined clusters were: subg. *Nymphaea*, subg. *Lotos* and subg. *Brachyceras-Anecphya*. The tree topology was supported by a good bootstrap value.

### *Nucleotide sequence analysis based on three regions combined (rbcL, matK and trnH–psbA intergenic spacer)*

Genetic relationships among 29 waterlily cultivars were estimated based on the three gene regions combined (*rbc*L, *mat*K and the *trn*H–*psb*A intergenic spacer). Altogether 188 variation positions were observed and a phylogenetic tree was generated using the MEGA program (version 6.0) and the ML procedure (Fig. 6). The separation of all 29 waterlilies based on the three gene regions was better than for any of the gene regions alone. The result indicated that five waterlily cultivars (*N.* 'Miss Siam', *N.* 'Mangala Ubol', *N.* 'Rojjana Ubol', *N.* 'Tuonta' and *N.* 'Chongkolnee') could be separated from the other waterlily cultivars. In addition, the different nucleotide patterns could be used for cultivar-specific identification (Table 4 and Figs. 2–4). However, the information from these three gene sequences could not be used to identify all 29 waterlilies in this study. The more gene regions and molecular markers that are included in the study, the better the separation result that is obtained. Considering the waterlily classification in the current study, the 29 waterlily cultivars were grouped into three clusters. Cluster I was composed of two subgroups. Subgroup 1 was hardy waterlilies, consisting of 10 waterlily cultivars in subg. *Nymphaea* (*N.* 'Miss Siam', *N.* 'Joey Tomocik', *N.* 'Siam Nymph', *N.* 'Siam Beauty', *N.* 'Pink Ribbon', *N.* 'Tanpong', *N.* 'Supranee Pink', *N.* 'Rattana Ubol', *N.* 'Tan–khwan', *N.* 'Wanvisa') and four intersubgeneric (ISG)

hybrids (*N.* 'Siam Purple 1', *N.* 'Siam Purple 2', *N.* 'Siam Pink' and *N.* 'Queen Sirikit'). These four ISG hybrids came from the cross between subg. *Nymphaea* and subg. *Brachyceras*. Although these four cultivars are ISG hybrids, based on their flowers and leaves, they are very much like waterlilies in subg. *Nymphaea.* Subgroup 2 consisted of the hardy waterlily, *N.* 'Mangala Ubol' which is a hybrid obtained from the cross *N. mexicana* x *N.* 'Perry's Fire Opal'. Both parents are in subg. *Nymphaea*. The hardy waterlily is a native plant that is found in temperate and frigid zones in Europe and America (Conard, 1905; Woods et al., 2005). It can subsist in cold weather and dormancy and abscission occur (Wasuwat and Wasuwat, 2004). Cluster II was tropical night-blooming waterlilies, consisting of four waterlily cultivars in subg. *Lotos* (*N.* 'Ploi Praow', *N.* 'Ploi Jear', *N.* 'Maeploi' and *N.* 'Rojjana Ubol'). Tropical night-blooming waterlilies are native plants in the tropical zone are not dormant in winter (Wasuwat and Wasuwat, 2004). Cluster III was tropical day-blooming waterlilies, consisting of waterlilies in subg. *Brachyceras* and subg. *Anecphya*, where subg. *Brachyceras* is composed of the six waterlily cultivars: *N.* 'Suwanna', *N.* 'Chularat', *N.* 'Chalong Kwan/ King of Siam', *N.* 'Chongkolnee' (a Thai native waterlily) and two open pollinated (OP) waterlily cultivars, *N.* 'See Som Tawan–ok' and *N.* 'Plai Fah'. Tropical day-blooming waterlilies are native plants in the tropical zone and are not dormant in winter, the same as tropical night-blooming waterlilies. Subg. *Anecphya* in cluster III consisted of three Australian giant waterlilies (*N. gigantea* (blue), *N.* 'Mayaranee' and *N.* 'Yaowalak'). The Australian giant waterlily has a cup-shaped flower and the flower bud is broadly cone-shaped with a tapered end, showing the structure above the water surface (Wasuwat and Wasuwat, 2004). *N.* 'Tuonta' is the ISG hybrid comes from the cross subg. *Anecphya* x subg. *Brachyceras*. Its appearance is very much like those in subg. *Anecphya*.

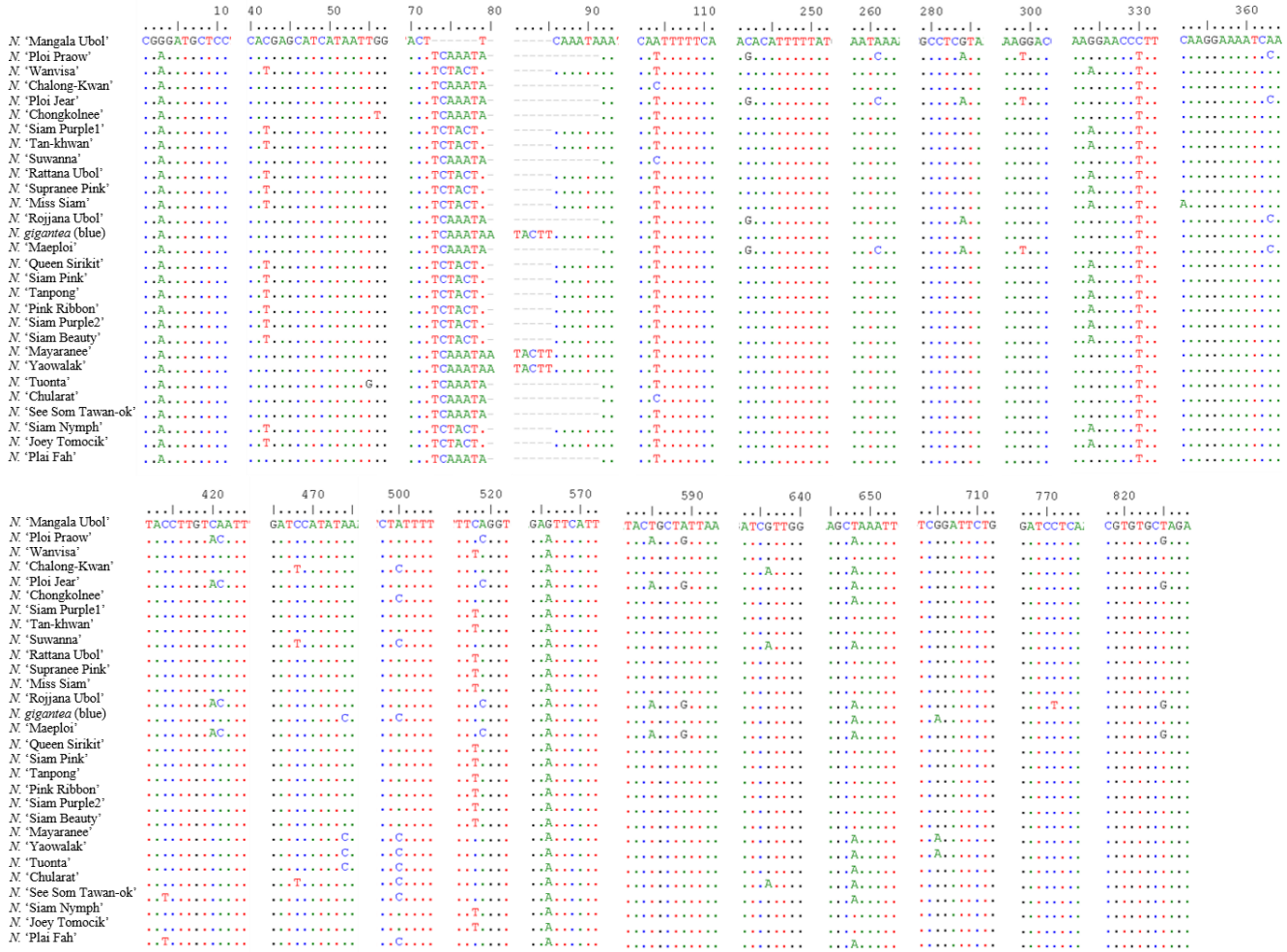

**Fig. 3** Variable position indicating base insertion/deletion, transition and transversion of *mat*K gene sequences after multiple sequence alignment of 29 waterlilies

**Fig. 4** Variable position indicating base insertion/deletion, transition and transversion of *trn*H–*psb*A intergenic spacer sequences after multiple sequence alignment of 29 waterlilies

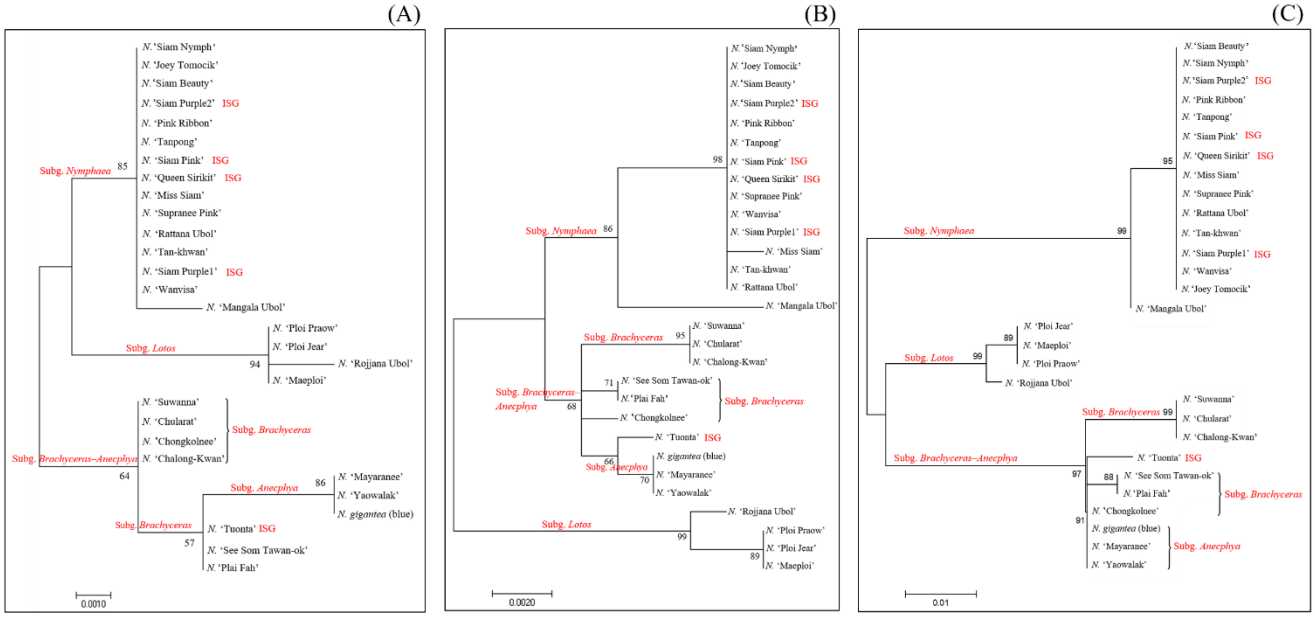


**Fig. 5** Maximum likelihood trees of 29 waterlily cultivars based on different chloroplast gene sequences: (A) *rbc*L; (B) *mat*K; (C) *trn*H-*psb*A intergenic spacer, where subg. = subgenus and ISG = intersubgeneric

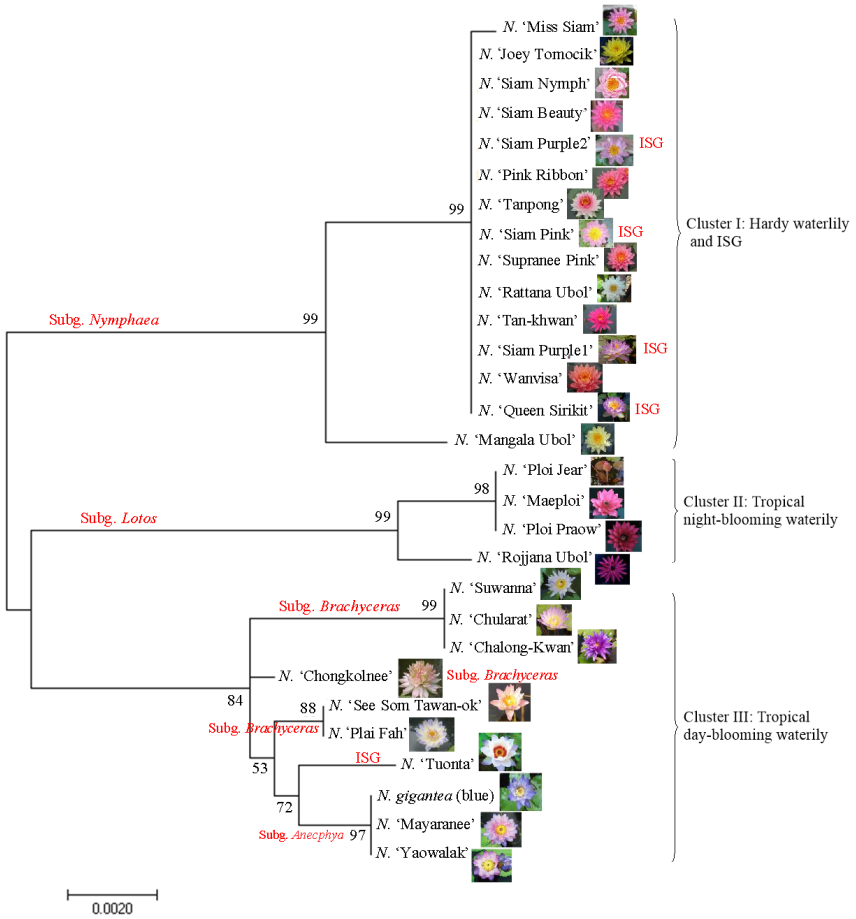


**Fig. 6** Maximum likelihood tree of 29 notable waterlily cultivars based on combined three gene regions (*rbc*L, *mat*K and *trn*H-*psb*A intergenic spacer), where subg. = subgenus and ISG = intersubgeneric

**Table 4** Specific base and position in three chloroplast gene sequences of individual waterlily cultivars compared to the other 28 waterlilies in this study for cultivar specific identification

| Cultivar | *rbc*L | | *mat*K | | *trn*H–*psb*A intergenic spacer | |
|---|---|---|---|---|---|---|
| | Position* | Base** | Position* | Base** | Position* | Base** |
| *N.* 'Mangala Ubol' | 316 | C (T) | 3 | G (A) | 325, 326 | T (–), T (–) |
| | | | 73, 74 | – (T), – (C) | 327, 328 | T (–), A (–) |
| | | | 76, 78 | – (A), – (T) | 329, 330 | T (–), T (–) |
| | | | 330, 566 | C (T), G (A) | 331, 332 | T (–), A (–) |
| *N.* 'Rojjana Ubol' | 370 | T (G) | 771 | T (C) | 278 | T (G) |
| *N.* 'Miss Siam' | – | – | 352 | A (C) | – | – |
| *N.* 'Chongkolnee' | – | – | 56 | T (G) | – | – |
| *N.* 'Tuonta' | – | – | 55 | G (T) | 273 | T (A) |

* Base position based on sequence of *N.* 'Mangala Ubol' after multiple sequence alignment of 29 waterlilies

** Base in parentheses indicates base found in other 28 waterlilies in this study.

All three gene specific regions (*rbc*L, *mat*K and the *trn*H–*psb*A intergenic spacer) are positioned in the chloroplast genome in which there is a unilateral genetic inheritance pattern. The plastid regions could be used as a tool to support taxonomic classification and to study the genetic relationship between mother and their offspring. In this study, it was found that the ISG hybrids were grouped in the same cluster as their mother's subgenus (Fig. 6). *N.* 'Siam Purple 1', *N.* 'Queen Sirikit', *N.* 'Siam Pink', and *N.* 'Siam Purple 2' are grouped in cluster I with other waterlilies in subg. *Nymphaea* because they are ISG hybrids between subg. *Nymphaea* (mother) and subg. *Brachyceras* (father). *N.* 'Ploi Jear' and *N.* 'Ploi Praow' are open pollinated hybrids and their mother is *N.* 'Maeploi'. The tree cultivars had no DNA sequencing differences and are grouped in cluster II with the other waterlilies (*N.* 'Rojjana Ubol') in subg. *Lotos*. As well as *N.* 'Tuonta' in cluster III of subg. *Brachyceras-Anecphya*, it is an ISG hybrid between subg. *Anecphya* (mother) and subg. *Brachyceras* (father). It is closely related with other waterlilies in subg. *Anecphya*, with which is grouped.

Considering waterlily classification based on morphological characters and their origins, waterlilies in the genus *Nymphaea* can be divided into two groups: hardy and tropical (Wasuwat and Wasuwat, 2004). In the current study, cluster I was composed of hardy waterlilies in subg. *Nymphaea,* whereas clusters II and III consisted of tropical waterlilies in the subgenera *Lotos, Brachyceras* and *Anecphya*. Waterlilies in the subgenera *Lotos, Brachyceras* and *Anecphya* are more closely related to waterlilies in subg. *Nymphaea* (Fig. 6). The same result was reported in waterlily classification by Borsch et al. (2007) in that waterlilies in genus *Nymphaea* were classified into three clusters based on sequences in the *trn*T-*trn*F region. The first separated cluster was subg. *Nymphaea* followed by the clusters of subg. *Brachyceras-Anecphya* and of subg. *Hydrocallis-Lotos*. Furthermore, Borsch et al. (2008) reported their study of nuclear ITSs and mitochondrial *mat*R indicated that the waterlilies in subg. *Brachyceras-Anecphya* were closely related and grouped and they were clearly separate from subg. *Nymphaea*. However, waterlilies in subg. *Nymphaea* should be closely related to subg. *Lotos*, as they both are in the syncarpiae group, which is a group of flowering plants whose ovaries consist of united carpels, whereas waterlilies in subg. *Brachyceras* and subg. *Anecphya* are in the apocarpiae group, whose ovaries consist of separate carpels (Conard, 1905).

Multiple sequence alignment of the three gene specific regions (*rbc*L, *mat*K and the *trn*H–*psb*A intergenic spacer) of the 29 waterlilies provided information of nucleotide variation patterns along the gene sequence (Tables 3 and 4). Indel (Insertion-deletion), base transition and transversion were detected (Table 3). The most variations of 130 positions were in the *trn*H-*psb*A intergenic spacer region, followed by 48 and 10 positions in the *mat*K and *rbc*L regions, respectively. The *trn*H-*psb*A intergenic spacer is a non-coding region that has been reported to have a high mutation rate in plants (Shaw et al., 2007) and remains the most variable for a single locus barcode for land plants (Kress and Erickson, 2007). Therefore, it has been used to study various plants (Gonzalez et al., 2009; Kress et al., 2010). However, the *mat*K region could be used for the specific identification of five notable waterlily cultivars from the 29 waterlilies used in the current study. The *mat*K region provided better information for cultivar specific identification than the other regions (the *trn*H-*psb*A intergenic spacer and *rbc*L) in the current study. The *mat*K region could be a potential marker for plant species identification due to it being a rapidly evolving, highly variant region of plant chloroplast DNA (Hilu and Liang, 1997; Biswal et al., 2012). In the current study, some waterlily cultivars could not be clearly identified based on each gene specific region; however, five waterlily cultivars could be clearly classified from the other cultivars using the three gene regions combined. CBOL Plant Working Group (2009) suggested using more specific regions to obtain the best results for cultivar-specific identification. Therefore, more specific gene regions should be used to clearly identify the notable waterlilies of Thailand. The current study was the first to classify and identify notable *Nymphaea* cultivars in Thailand using the combined information from the plastid *rbc*L and *mat*K coding regions and the *trn*H-*psb*A intergenic spacer region. These DNA sequences should be useful as a tool for specific identification and for registration of future new releases of notable *Nymphaea* cultivars.

## Conflict of Interest

The authors declare that there are no conflicts of interest.

## Acknowledgements

Ms. Primlarp Wasuwat of the 'Pang U Bon' Waterlily Garden, Dr. Nopchai Chansilpa of the Rajamangala University of Technology Tawan–ok Waterlily Germplasm Bank and Mr Pairat Songpanich, a waterlily hybridizer, provided petal

and leaf specimens of the waterlilies. The Kasetsart University Research and Development Institute (KURDI), Bangkok, Thailand and the Development and Promotion of Science and Technology Talented Project (DPST) of Thailand funded this research.

## References

Biswal, D.K., Debnath, M., Kumar, S., Tandon, P. 2012. Phylogenetic reconstruction in the Order Nymphaeales: ITS2 secondary structure analysis and *in silico* testing of maturase k (*mat*K) as a potential marker for DNA bar coding. BMC Bioinformatics 13: S26. doi.org/10.1186/1471-2105-13-S17-S26

Borsch, T., Lohne, C., Wiersema, J. 2008. Phylogeny and evolutionary patterns in Nymphaeales: Integrating genes, genomes and morphology. Taxon 57: 1052–1081. doi.org/10.1002/tax.574004

Borsch, T., Hilu, K.W., Wiersema, J.H., Lohne, C., Barthlott, W., Wilde, V. 2007. Phylogeny of *Nymphaea* (Nymphaeaceae): Evidence from substitutions and microstructural changes in the chloroplast *trnT-trnF* region. Int. J. Plant Sci. 168: 639–671.

CBOL Plant Working Group. 2009. A DNA barcode for land plants. Proc. Natl. Acad. Sci. USA 106: 12794–12797. doi.org/10.1073/pnas.0905845106

Chase, M.W., Cowan, R.S., Hollingsworth, P.M., et al. 2007. New trends in plant systematics: A proposal for a standardised protocol to barcode all land plants. Taxon 56: 295–299.

Conard, H.S. 1905. The Waterlilies: A Monograph of the Genus *Nymphaea*. The Carnegie Institute of Washington. Washington DC, USA.

Dkhar, J., Kumaria, S., Rao, S.R., Tandon, P. 2012. Sequence characteristics and phylogenetic implications of the nrDNA internal transcribed spacers (ITS) in the genus *Nymphaea* with focus on some Indian representatives. Plant Syst. Evol. 298: 93–108. doi.org/10.1007/s00606-011-0526-z

Dobrogojski, J., Adamiec, M., Lucinski, R. 2020. The chloroplast genome: A review. Acta Physiol. Plant 42: 98. doi.org/10.1007/s11738-020-03089-x

Doyle, J.J., Doyle, J.L. 1990. Isolation of plant DNA from fresh tissue. Focus 12: 13–15.

Fazekas, A.J., Kuzmina, M.L., Newmaster, S.G., Hollingsworth, P.M. 2012. DNA barcoding methods for land plants. In: Kress, W.J., Erickson, D.L. (Eds.). DNA Barcodes: Methods and Protocols. Humana Press. Totowa, NJ, USA, pp. 223–252.

Gonzalez, M.A., Baraloto, C., Engel, J., et al. 2009. Identification of Amazonian trees with DNA barcodes. PLoS One 4: e7483. doi.org/10.1371/journal.pone.0007483

Hilu, K.W., Liang, H. 1997. The *mat*K gene: Sequence variation and application in plant systematics. Am. J. Bot. 84: 830–839. doi.org/10.2307/2445819

Hollingsworth, P.M., Graham, S.W., Little, D.P. 2011. Choosing and using a plant DNA barcode. PLoS One 6: e19254. doi.org/10.1371/journal.pone.0019254

International Waterlily and Water Gardening Society (IWGS). 2016. IWGS new waterlily competitions. Greenville, VA, USA. https://iwgs.org/new-waterlily-competition/, 20 November 2016.

Kim, Y., Kwon, W., Song, M.J., Nam, S., Park, J. 2019a. The complete chloroplast genome sequence of the *Nymphaea lotos* L. (Nymphaeaceae). Mitochondr. DNA Part B: Resources 4: 389–390. doi.org/10.1080/23802359.2018.1547154

Kim, Y., Min, J., Kwon, W., Song, M.J., Nam, S., Park, J. 2019b. The complete chloroplast genome sequence of the *Nymphaea capensis* Thunb. (Nymphaeaceae). Mitochondrial DNA B: Resour. 4: 401–402. doi.org/10.1080/23802359.2018.1547171

Kress, W.J., Erickson, D.L. 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 2: e508. doi.org/10.1371/journal.pone.0000508

Kress, W.J., Erickson, D.L., Jones, F.A., Swenson, N.G., Perez, R., Sanjur, O., Bermingham, E. 2009. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. Proc. Natl. Acad. Sci. USA 106: 18621–18626. doi.org/10.1073/pnas.0909820106

Kress, W.J., Erickson, D.L, Swenson, N.G., Thompson, J., Uriarte, M., Zimmerman, J.K. 2010. Advances in the use of DNA barcodes to build a community phylogeny for tropical trees in a Puerto Rican forest dynamics plot. PLoS ONE 5: e15409.

La-ongsri, W., Wattatana, S. 2010. "Bua" in Thailand. The Botanical Garden Organization. Chiang Mai, Thailand. [in Thai]

Levin, R.A., Wagner, W.L., Hoch, P.C., Nepokroeff, M., Pires, J.C., Zimmer, E.A., Sytsma, K.J. 2003. Family–level relationships of Onagraceae based on chloroplast *rbc*L and *ndh*F data. Am. J. Bot. 90: 107–115. doi.org/10.3732/ajb.90.1.107

Ohyama, K., Fukuzawa, H., Kohchi, T., et al. 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322: 572–574. doi.org/10.1038/322572a0

Sang, T., Crawford, D.J., Stuessy, T.F. 1997. Chloroplast DNA phylogeny, reticulate evolution and biogeography of *Paeonia* (Paeoniaceae). Am. J. Bot. 84: 1120–1136. doi.org/10.2307/2446155

Shaw, J., Lickey, E.B., Schilling, E.E., Small, R.L. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. Am. J. Bot. 94: 275–288. doi.org/10.3732/ajb.94.3.275

Shinozaki, K., Ohme, M., Tanaka, M., et al. 1986. The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. EMBO J. 5: 2043–2049. doi.org/10.1002/j.1460-2075.1986.tb04464.x

Songpanich, P. 2016. Hardy waterlily hybridization. WGI Online Journal 4: 1–1. http://www.watergardenersinternational.org/journal/archive.html, 1 November 2016.

Songpanich, P., Hongtrakul, V. 2010. Intersubgeneric cross in *Nymphaea* spp. L. to develop a blue hardy waterlily. Sci. Hortic. 124: 475–481. doi.org/10.1016/j.scienta.2010.01.024

Sun, C., Chen, F., Teng, N., Xu, Y., Dai, Z. 2021. Comparative analysis of the complete chloroplast genome of seven *Nymphaea* species. Aquat. Bot. 170: 103353. doi.org/10.1016/j.aquabot.2021.103353

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30: 2725–2729. doi.org/10.1093/molbev/mst197

Tate, J.A., Simpson, B.B. 2003. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. Syst. Bot. 28: 723–737.

Wasuwat, P., Wasuwat, S. 2004. Waterlily in Thailand. Nation books Publishing. Bangkok, Thailand. [in Thai]

Woods, K., Hilu, K.W., Borsch, T., Wiersema, J.H. 2005. Pattern of variation and systematics of *Nymphaea odorata*: II. Sequence information from ITS and *trnL-trnF*. Syst. Bot. 3: 481–493.

Yang, J.B., Li, D.Z., Li, H.T. 2014. Highly effective sequencing whole chloroplast genomes of angiosperms by nine novel universal primer pairs. Mol. Ecol. Resour. 14: 1024–1031. doi.org/10.1111/1755-0998.12251