

Original article

Analysis of phylogenetic relationship between sacred lotus and woody plants using nucleotide sequences of *rbcL* and its adjacent genes

Akhom Khatfan¹

Nopadol Riablershirun¹

Vithaya Sathornviriyapong²

Niran Juntawong^{1*}

¹Department of Botany, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand

²Department of Horticulture, Faculty of Agriculture, Kasetsart University, Bangkok 10900, Thailand

*Corresponding Author, Email: fscinrj@ku.ac.th

Received: Nov 12, 2013

Accepted: Jan 7, 2014

ABSTRACT

Sacred lotus (*Nelumbo nucifera* Gaertn.) was morphologically considered the same family of water lilies, Nymphaeaceae, in spite of its difference in palynology and floral development. The further investigation genetic variation was to determine different in polymorphism. The chloroplast *rbcL* and adjacent genes were selected from three wild-type Thai lotus cultivars, Klong Yong (KY), Nong Nan (NH) and Rachine (RC). The contiguous 3.74 kb of chloroplast DNA was obtained by PCR using three primer pairs comprising of *atpB-rbcL* intergenic spacer (IGS), *rbcL* gene, *rbcL-accD* IGS and partial *accD* gene. These four regions were compared within *Nelumbo* species and with ten selected dicotyledonous species. The Neighbor-Joining phylogenetic tree based on *rbcL* genes and *atpB-rbcL* and *rbcL-accD* intergenic spacers indicated that sacred lotus was closely related to woody plants rather than water lilies.

Keywords: Intergenic spacer, *Nelumbo*, Rubisco, chloroplast, Proteales

INTRODUCTION

Sacred lotus or Asian lotus (*Nelumbo nucifera* Gaertn.), a large perennial aquatic plant growing in shallow waters, is considered as a multipurpose plant by Asian people; sacred lotus is mainly used for foods, cosmetics, medical herbs, worship and ornamental plants (Mukherjee *et al.*, 1996).

Before 2003, sacred lotus was considered the same family of water lilies, Nymphaeaceae, according to its habit of aquatic herb and

possession of a hook-shaped appendage in stamens and petaloid staminodes (Schneider and Buchanan, 1980; The Angiosperm Phylogeny Group, 2003). Nevertheless, sacred lotus was distinguished by many characters such as leaf shape, carpel structure and pollen grain shape (Kreunen and Osborn, 1999; Bodhipadma *et al.*, 2013). Unlike the water lilies which pollen grains were monosulcate, sacred lotus pollen grains were triclopatate which was unique for eudicot plants (Judd and Olmstead, 2004).

Therefore sacred lotus was re-classified in a new family, Nelumbonaceae, and was related to Proteales which the members were characterized by protandry and highly self-incompatibility (The Angiosperm Phylogeny Group, 2003). However, some sacred lotus characteristics were contrary to those of Proteales such as possession of progyny and degree of stigma-pollen self-incompatibility (Kreunen and Osborn, 1999; Khatfan *et al.*, 2014).

Genetic polymorphism was considered the new technique to classify organisms based on genetic information. This polymorphism was firstly studied in early 1990s (Les *et al.*, 1991). Consequently, both nuclear and plastid polymorphisms were studied in various plants (Matsuda *et al.*, 2005; Tian *et al.*, 2008; Moore *et al.*, 2010; Sangin *et al.*, 2010), resulting in re-classification of various organisms including sacred lotus (*Nelumbo nucifera*) which was eventually classified to the new family, Nelumbonaceae based on genetic polymorphism (Les *et al.*, 1991; Moore *et al.*, 2010).

The large sub-unit of the Rubisco gene was the target in this study due to several documents depicted of high reliability in polymorphism (Matsuda *et al.*, 2005; Shaw *et al.*, 2005; Sangin *et al.*, 2010). The aim of this study was to include phylogenetic information of *rbcL-accD* intergenic spacer to the phylogenic database. The information was expected to be useful on molecular breeding of sacred lotus; the genetics information of woody plants in Proteales should be involved

in molecular breeding of sacred lotus, or *vice versa*.

MATERIALS AND METHODS

Sites of sampling

Three Thai lotus cultivars—namely, Klong Yong (KY), Nong Han (NH) and Rachine (RC) which were naturally grown at $-13^{\circ} 59' N 100^{\circ} 23' E$, $17^{\circ} 15' N 104^{\circ} 10' E$ and $12^{\circ} 57' N 99^{\circ} 58' E$, respectively. KY is the white petal cultivar, while NH and RC are the red petal cultivars.

Samples preparation

The leaf samples were collected from three individuals of each cultivar. The samples were frozen in $-20^{\circ}C$ deep freeze until extraction. About 0.5 g of leaf tissue was ground in liquid nitrogen to powder and was extracted for DNA using modified protocol from Doyle and Doyle (1987). Ground powder was extracted in 5 ml of 2% (w/v) Cetyl Trimethyl Ammonium Bromide (CTAB, Sigma-Aldrich) buffer. The suspension was partitioned in chloroform: isopentanol (24:1, v/v) and then was precipitated in isopropanol. The DNA pellet was re-suspended in 500 μL of TE buffer. For RNA removal, RNase 100 ng μL^{-1} was added and followed by 30 min $37^{\circ}C$. RNases was removed by adding the same volume of phenol: chloroform: isopentanol (25:24:1, v/v). The extract solution was re-partitioned in the solvent of chloroform: isopentanol (24:1, v/v) and then was precipitated in absolute ethanol. The pellet was kept in 20 μL of TE buffer.

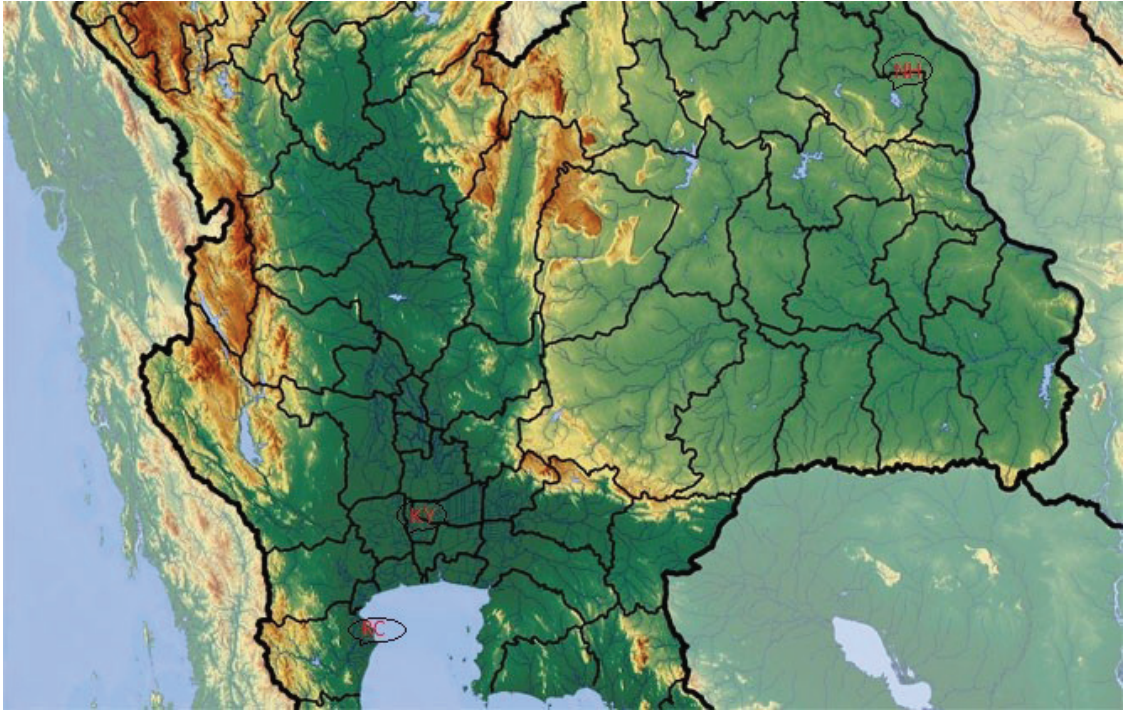


Figure 1 Map indicating sites of lotus cultivar sampling.

Primer design and DNA amplification

Six primers were designed for amplifying and sequencing three selected regions of cpDNA

fragment. The melting temperature (t_m) of each primer was between 44 - 52°C. The primers were synthesized by BioDesign™, Thailand.

Table 1 List of primers designed in this study.

Primer	Sequence	Direction	T_m (°C) ^a
P1	5'- TCGACCCTTAATTACCAAAG - 3'	Forward	51.9
R1	5'- GGATTCAAAGCTGG - 3'	Forward	44.7
R2	5'- TGGTAAGTCCATCG - 3'	Reverse	44.7
R3	5'- TTACCAAGGATGATGAG - 3'	Forward	47.7
R4	5'- TGTAGTACGGAATCATC - 3'	Reverse	47.7
D4	5'- CTGAACTACTCATTTC - 3'	Reverse	45.3

Note: ^a T_m was calculated based on von Ahsen *et al.* (2001) at the $[Na^+]$ of 0.050 M.

DNA samples were amplified in the reaction mixture containing 1 unit of *Taq* polymerase (Fermentas, Canada) was utilized in 40 μ L reaction mixture. The mixture contained 500 mM KCl, 100 mM

Tris-HCl, pH 8.0, and 1.0% (v/v) Triton X-100, 2.0 mM $MgCl_2$, 1 μ M of each primer, and 200 μ M each dNTP (Fermentas). PCR amplification was carried out in 30 cycles of 94°C denaturing for 45 s, 48°C

annealing for 1 min 15 s and 72°C extension for 1 min 15 s, followed by 72°C extension for 10 min and 4°C for storing. The PCR products were purified by electrophoresis and were recovered with gel extraction kit (Fermentas).

DNA Sequence analysis

The samples were sequenced at BioDesign™ Thailand. The sequences were manually edited and were aligned using ClustalW (version 2.1, 2014) programed by Kyoto University Bioinformatics Center, Kyoto, Japan. The sequences were compared to the *N. nucifera* (JQ336993 and KF009944) and *N. lutea* (JQ336992) together with the other dicotyledonous plants which were downloaded from GenBank. The aligned sequences were used for creating the Neighbor-Joining dendrogram by the MEGA4 program (Tamura *et al.*, 2007).

RESULTS AND DISCUSSION

DNA Fragments and contiguous sequences comparison

Three fragments (A, C, D) were amplified into the size of 1.1, 1.2 and 2.3 kilobase (kb), respectively (Figure 2). A contiguous sequence of 3.74 kb was generated from these three DNA fragments by using ClustalW (version 2.1, 2014), comprising three genic portions which were 0.15 kb of *atpB*, 1.43 kb of *rbcL* and 0.81 kb of *accD* gene, and two intergenic spacers (IGS) which were 0.74 kb of *atpB-rbcL* and 0.69 kb of *rbcL-accD* IGS (Figure 3). The contiguous sequence located the full length *rbcL* gene. The location of *rbcL* gene between *atpB* and *accD* genes indicated that *Nelumbo* was typically a dicotyledonous plant (Harris *et al.*, 2013).

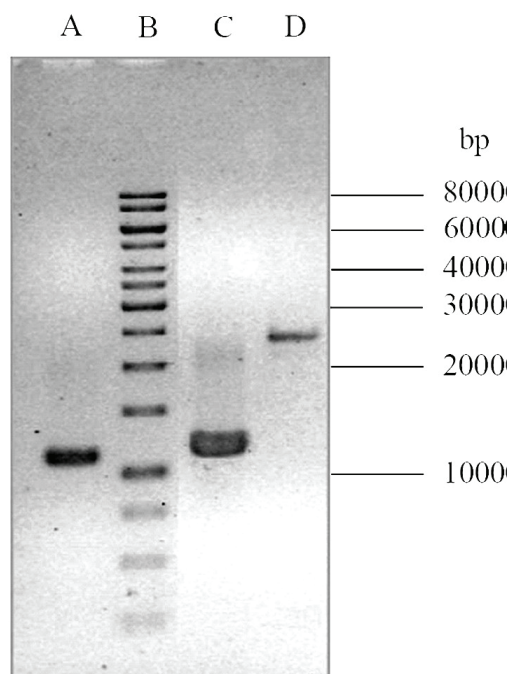


Figure 2 Gel electrophoresis of three contiguous fragments.

Based on alignment, the contiguous sequences of three lotus cultivars (NH, RC and KY) were similar and corresponded to the nucleotide portion of the *Nelumbo nucifera* cpDNA (Xue *et al.*, 2012) indicating that three lotus cultivars, in this study, were closely related to *N. nucifera*. The contiguous sequences of three cultivars together with the *N. nucifera* and *N. lutea* sequences studied by Xue *et al.* (2012) were similar on the portion of *atpB* gene, *atpB-rbcL* intergenic spacer (IGS), *rbcL* and *accD* gene. Only the portion of *rbcL-accD* IGS was different within two species of *Nelumbo*

(Figure 4). The similarity in *rbcL* of *Nelumbo* was also reported the study of Les *et al.* (1991) who documented that the *rbcL* sequence was almost similar within two species of *Nelumbo*, except a nucleotide substitution at the position 61887 which the base “G”, in *N. nucifera*, was replaced by the base “A”, in *N. lutea*. This substitution resulted in the 95th amino acid residue switching from Ser in *N. nucifera*, into Asn, in *N. lutea*, which commonly occurred in nature and did not affect plant photosynthesis (Wilson and Finlay, 1997; Betts and Russell, 2003).

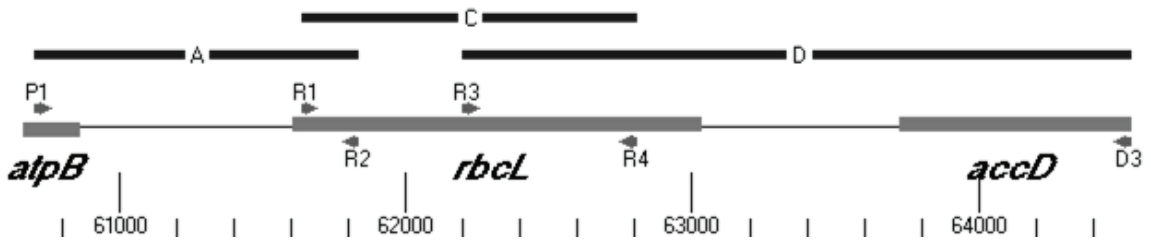


Figure 3 Position of three fragments in this study.

On *rbcL-accD* IGS, only three points of nucleotide insertion/deletion/substitution were observed within three cultivars of *N. nucifera*, indicating that intergenic sequences were evolved quicker than genic sequences. The result was similar to the study of Matsuda *et al.* (2005) who studied the sequence variation in Moraceae. In dicotyledonous plants *rbcL*

and *atpB* genes were more conserved than *accD* gene which lined adjacent to those two genes in the flowering plants (Diekmann *et al.*, 2009). The *accD* gene was also located next to the *rbcL* gene in both monocotyledonous and dicotyledonous plants; however this gene was not function within monocotyledon (Guisinger *et al.*, 2010; Harris *et al.*, 2013).

KY were native to the central part of Thailand which was about 700 km. apart from the NH which was indigenous to the north-eastern part of Thailand. The insertions or deletions

of chloroplast nucleotide which caused by geologic separation was also reported by other citations (Yasui and Ohnishi, 1998; Matsuda *et al.*, 2005; Xue *et al.*, 2012).

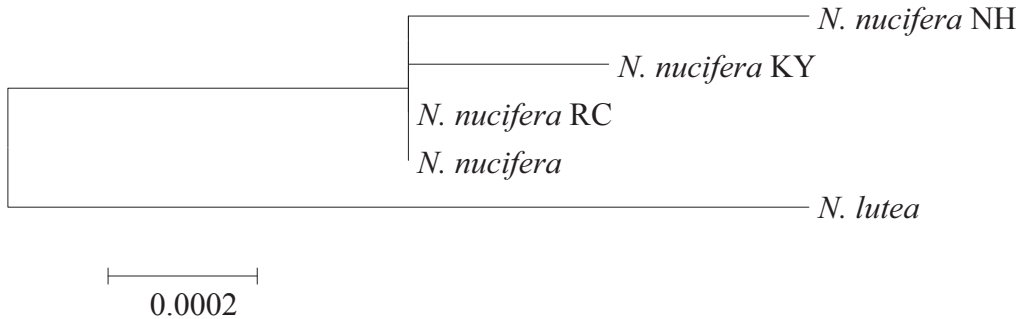


Figure 5 Rooted phylogenetic trees of three Thai sacred lotus cultivars generated by Neighbor-Joining method based on *atpB-rbcL* IGS, *rbcL* gene and *rbcL-accD* IGS. The numbers indicating bootstrap percentages were not shown due to the values being lower than 50.

One more substitution of nucleotide was also found in KY *accD* gene at position 63730 (Figure 4). This substitution resulted in the change of the 3rd amino acid residue from Lys to Asn. The nucleotide substitution in this position has not been reported elsewhere before. Therefore, there was insufficiency explanation regarding to this substitution. However, another nucleotide substitution at *accD* gene was also reported in *N. lutea* but at the position 63757 (Xue *et al.*, 2012) which did not result in the amino acid substitution at the corresponding residue (the 12th amino acid). These two nucleotide substitutions might be more or less involving to the petal color; the whitish color of petal in KY and the yellowish color of petal in *N. lutea*. There was no further study result of nucleotide substitutions at this position of *accD* gene. However, the Lys-Asn modification at the 3rd residue of *accD* gene might not affect the ability of Zn²⁺ binding

of this enzyme since this position was quite far from the 27th serine which played an important role in Zn²⁺ binding efficiency in lipid metabolism (Meades *et al.*, 2010).

Phylogenetic relationship of Nelumbonaceae and other dicotyledonous plants

As previously mention, the chloroplast genic sequences (such as *atpB*, *rbcL* and *accD*) were too similar and gave too little information to generate to phylogenetic tree with the *Nelumbo* taxons. Hence, the polymorphisms of chloroplast genic regions such as, *atpB*, *rbcL* and *accD* were insufficient for classification of Nelumbonaceae as it was mentioned in the study of Les *et al.* (1991) and Xue *et al.* (2012). The nucleotide sequences of three lotus cultivars were required to be compared to other dicotyledonous. The corresponding location from other ten plant species, including two water lilies, *Nuphar advena* and *Nymphaea*

alba, were selected to create phylogenetic trees based on both genic and intergenic sequences.

The details of plants, including the Accession numbers, were indicated in Table 2.

Table 2 Selected dicotyledonous plants whose sequences were used to generate phylogenetic trees.

Binomial name	Habit	Family	Accession numbers
<i>Buxus microphylla</i>	Shrub	Buxaceae	EF380351
<i>Eucalyptus globulus</i>	Tree	Myrtaceae	NC_008115
<i>Hevea brasiliensis</i>	Tree	Euphorbiaceae	HQ285842
<i>Liriodendron tulipifera</i>	Tree	Magnoliaceae	DQ899947
<i>Magnolia yunnanensis</i>	Tree	Magnoliaceae	KF753638
<i>Nelumbo nucifera</i>	Herb	Nelumbonaceae	FJ754269, JQ336993, KM655836
<i>Nelumbo lutea</i>	Herb	Nelumbonaceae	JQ336993
<i>Nuphar advena</i>	Herb	Nymphaeaceae	DQ354691
<i>Nymphaea alba</i>	Herb	Nymphaeaceae	AJ627251
<i>Platanus occidentalis</i>	Tree	Platanaceae	DQ923116
<i>Tectona grandis</i>	Tree	Lamiaceae	NC_020098
<i>Tetracentron sinense</i>	Tree	Trochodendraceae	KC608752

Both gene sequences and intergenic spacers (IGS) were separately used to create the phylogenetic tree. Phylogenetic trees based on *rbcL* gene, *atpB-rbcL* and *rbcL-accD* IGS indicated that *Nelumbo* was closely related to a woody plant, *Platanus* rather than water lilies genera such as *Nymphaeae* and *Nuphar* (Figure 6, 8 and 9). The similar results based on plastid DNA sequences were reported and indicated that *Nelumbo* was more genetically related to basal eudicots like *Buxus*, *Magnolia*, *Platanus* and *Trochodendron* (Les *et al.*, 1991; Moore *et al.*, 2010). *Platanus*, *Trochodendron* and *Buxus*

shared similar morphologic characteristics to sacred lotus like flower shape, monoecism (except *Buxus*), fruit morphology and leaf venation (except *Trochodendron*). Although, based on *accD* gene, *Nelumbo* was more closely related to water lilies (Figure 7) which were considered the early-diverging angiosperms. Hence, the genetic polymorphism based on *accD* gene might be more or less involving the common ancestors of basal eudicots and water lilies and was required to be farther studied.

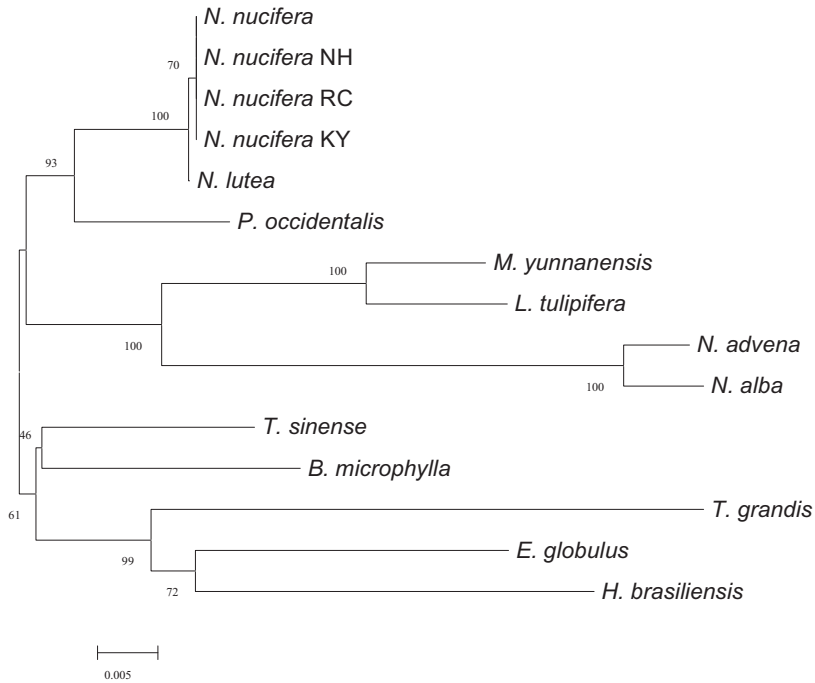


Figure 6 Rooted phylogenetic trees of sacred lotus and other dicotyledonous plants generated by Neighbor-Joining method, based on *rbcL* gene. Numbers indicate bootstrap percentages.

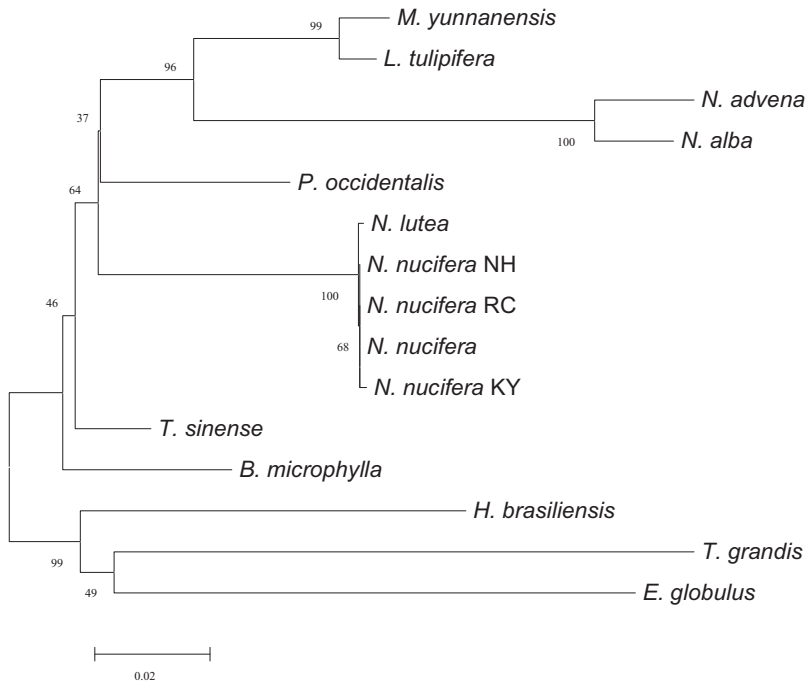


Figure 7 Rooted phylogenetic trees of sacred lotus and other dicotyledonous plants generated by Neighbor-Joining method, based on *accD* gene. Numbers indicate bootstrap percentages.

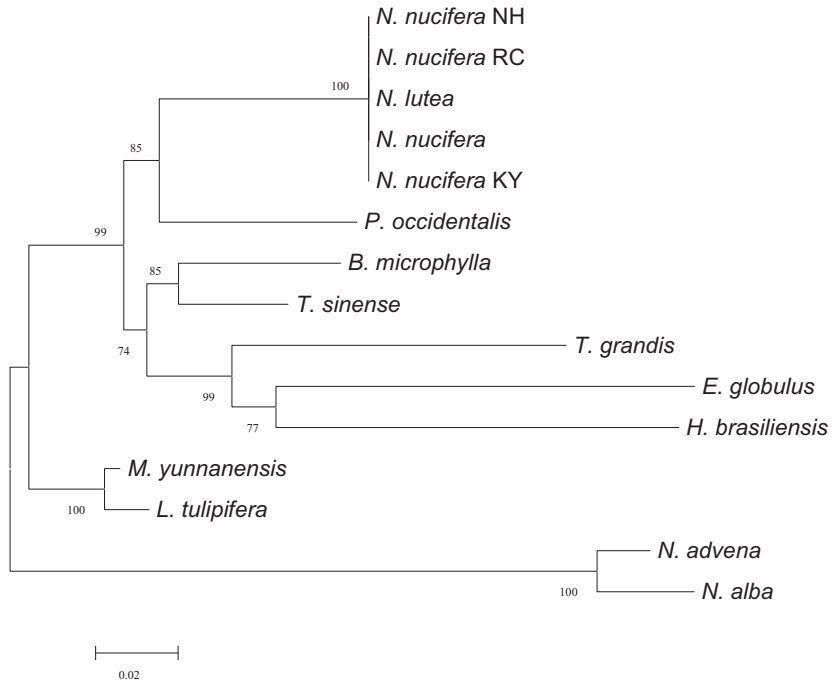


Figure 8 Rooted phylogenetic trees of sacred lotus and other dicotyledonous plants generated by Neighbor-Joining method, based on *atpB-rbcL* IGS. Numbers indicate bootstrap percentages.

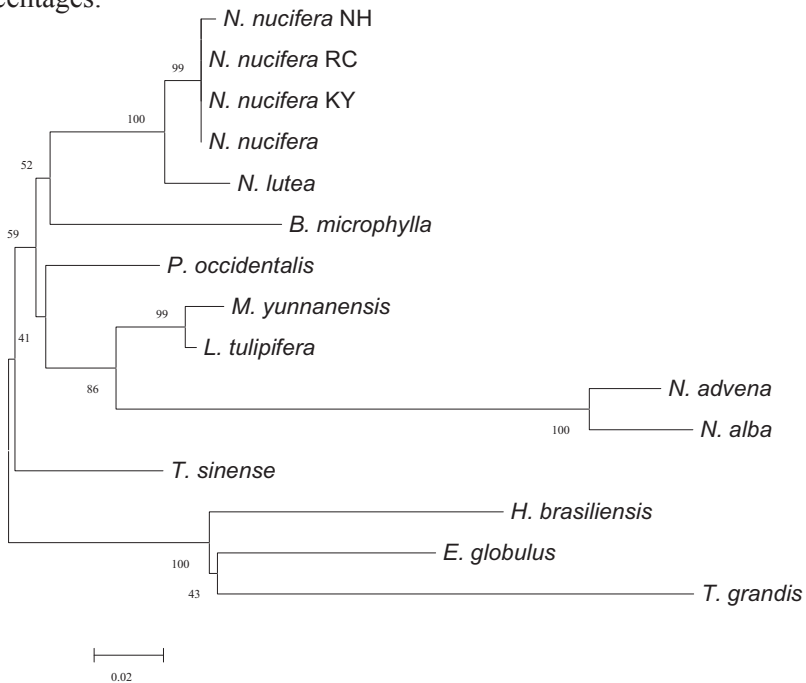


Figure 9 Rooted phylogenetic trees of sacred lotus and other dicotyledonous plants generated by Neighbor-Joining method, based on *rbcL-accD* IGS. Numbers indicate bootstrap percentages.

The phylogenetic tree based on *rbcL* (Figure 6) was less diverged than those based on *accD* gene, *atpB-rbcL* and *rbcL-accD* IGS (Figure 7, 8 and 9), indicating that *rbcL* gene was more conserved than other regions in chloroplast. The conservatism of *rbcL* gene might play critical roles on plant metabolism (Parry *et al.*, 2003). Moreover, the genic regions were typically lesser difference at than that of the intergenic region as both in nuclear and plastid genes (Shaw *et al.*, 2005; Walsh and Hoot, 2001).

Genetic polymorphism study would be one of the methods to determine plant evolution which depicted ecological change during evolution. To adapt to the harsh environment like global warming and for the better in both agricultural and forestry aspect, the knowledge of plant evolution was required to modern plant breeders in order to select the suitable plants to grow in the certain areas in order to get better on both economic and ecological benefit.

CONCLUSION

Chloroplast *rbcL* and *accD* genes were conserved among the Nelumbonaceae species and were almost the same size within dicot plants. Based on nucleotide comparison of the corresponding locations between sacred lotus species and other dicotyledonous plant species, sacred lotus was more closely related to basal eudicot plants like *Buxus microphylla*, *Platanus occidentalis* and *Tetracentron sinense* than the early-diverging angiosperms like *Nuphar advena* and *Nymphaea alba*, although the latter group was similar in habit and morphology.

ACKNOWLEDGEMENTS

The authors would like to thank the Kasetsart University Research and Development Institute (KURDI), the Graduate School of Kasetsart University and The Higher Education Research Promotion and National Research University Project of Thailand, Office of the Higher Education Commission for the financial support to carry out the project.

REFERENCES

- Betts, M. J. and R. B. Russell. 2003. Amino acid properties and consequences of substitutions, pp. 289-316. *In* Barnes MR and Gray IC (eds.). **Bioinformatics for Geneticists**. John Wiley & Sons, Ltd.
- Bodhipadmaa, K., S. Noichindaa, P. Thaiyantoo, and D. W. M. Leung. 2013. Morphology, viability, and germinability of pollen from two forms of *Nymphaea nouchali* var. *versicolor*, a day-blooming waterlily. **ScienceAsia** 39: 214-218.
- Diekmann, K., T. R. Hodkinson, K. H. Wolfe, R. van den Bekerom, P. J. Dix and S. Barth. 2009. Complete chloroplast genome sequence of a major allogamous forage species, perennial Ryegrass (*Lolium perenne* L.). **DNA Res.** 16: 165-176.
- Doyle, J. J. and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. **Phytochem. Bull.** 19: 11-15.
- Guisinger, M. M., T. W. Chumley, J. V. Kuehl, J. L. Boore and R. K. Jansen. 2010.

- Implications of the plastid genome sequence of *Typha* (Typhaceae, Poales) for understanding genome evolution in Poaceae. **J. Mol. Evol.** 70: 149-166.
- Harris, M. E., G. Meyer, T. Vandergon and V. O. Vandergon. 2013. Loss of the Acetyl-CoA Carboxylase (*accD*) gene in Poales. **Plant Mol. Biol. Rep.** 31: 21-31.
- Judd, W. S. and R. G. Olmstead. 2004. A survey of tricolpate (Eudicot) phylogenetic relationships. **Am. J. Bot.** 91 (10): 1627-1644.
- Khatfan, A., Z. Li, L. Q. Chen, N. Riablershirun, V. Sathornviriyapong and N. Juntawong. 2014. Pollen viability, germination, and seed setting of *Nelumbo nucifera*. **ScienceAsia** 40: 384-392.
- Kreunen, S.S. and J.M. Osborn. 1999. Pollen and anther development in *Nelumbo* (Nelumbonaceae). **Am. J. Bot.** 86(12): 1662-1676.
- Les, D. H., D. K. Garvin and C. F. Wimpee. 1991. Molecular evolutionary history of ancient aquatic angiosperms. **Proc. Natl. Acad. Sci. USA** 88 (22): 10119-10123.
- Matsuda, Y., H. Yoshimura, H. Kanamoto, T. Ujihara, K. Tomizawa, Y. Sugimura and S. Kitajima. 2005. Sequence variation in the *rbcL-accD* region in the chloroplast genome of Moraceae. **Plant Biotechnol.** 22 (3): 231-233.
- Meades, G. Jr, B. K. Benson, A. Grove and G. L. Waldrop. 2010. A tale of two functions: enzymatic activity and translational repression by carboxyltransferase. **Nucleic Acids Res.** 38 (4): 1217- 1227.
- Moore, M. J., P. S. Soltis, C. D. Bell, J. G. Burleigh and D. E. Soltis. 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of Eudicots. **Proc. Natl. Acad. Sci. USA** 107 (10): 4623-4628.
- Mukherjee, P. K., R. Balasubramanian, K. Saha, B. P. Saha and M. Pal. 1996. A review on *Nelumbo nucifera* Gaertn. **Anc. Sci. Life** 15 (4): 268-276.
- Parry, M. A. J., P. J. Andralojc, R. A. C. Mitchell, P. J. Madgwick and A. J. Keys. 2003. Manipulation of Rubisco: the amount, activity, function and regulation. **J. Exp. Bot.** 54: 1321-1333.
- Sangin, P., A. J. Lindstrom, G. Kokubugata, M. Chaiprasongsuk and M. Mingmuang. 2010. Phylogenetic relationships within Cycadaceae inferred from non-coding regions of chloroplast DNA. **Kasetsart J. (Nat. Sci.)** 44: 544-557.
- Schneider, E. L. and J. D. Buchanan. 1980. Morphological studies of the Nymphaeaceae. XI. The floral biology of *Nelumbo pentapetala*. **Am. J. Bot.** 67: 189-193.
- Shaw, J., E. Lickey, J. T. Beck, S. B. Farmer, W. Liu, J. Miller, K. C. Siripun, C. T. Winder, E. E. Schilling and R. L. Small. 2005. The tortoise and the hare II: relative utility of 21 non coding

- chloroplast DNA sequences for phylogenetic analysis. **Am. J. Bot.** 92: 142-166.
- Tian, H. L., X. Q. Chen, J. X. Wang, J. H. Xue, J. Wen, G. Mitchell and S. L. Zhou. 2008. Development and characterization of microsatellite loci for lotus (*Nelumbo nucifera*). **Conserv Genet.** doi 10.1007/s10592-007-9503-z
- Tamura K., J. Dudley, M. Nei and S. Kumar. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. **Mol. Biol. Evol.** 24: 1596-1599.
- The Angiosperm Phylogeny Group. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. **Bot. J. Linn. Soc.** 141: 399–436.
- Von Ahsen, N., C. T. Wittwer and E. Schütz. 2001. Oligonucleotide melting temperatures under PCR conditions: Nearest-neighbor corrections for Mg²⁺, deoxynucleotide triphosphate, and dimethyl sulfoxide concentrations with comparison to alternative empirical formulas. **Clinical Chemistry** 47: 1956-1961.
- Walsh, B. M. and S. B. Hoot. 2001. Phylogenetic relationships of *Capsicum* (Solanaceae) using DNA sequences from two noncoding regions: The chloroplast *atpB-rbcL* spacer region and nuclear waxy introns. **Int. J. Plant Sci.** 162 (6): 1409-1418.
- Wilson, D. R. and B. B. Finlay. 1997. The “Asx-Pro turn” as a local structural motif stabilized by alternative patterns of hydrogen bonds and a consensus-derived model of the sequence Asn-Pro-Asn. **Prot. Eng.** 10: 519-529.
- Xue, J. H., S. Wang and S. L. Zhou. 2012. Polymorphic chloroplast microsatellite loci in *Nelumbo* (Nelumbonaceae). **Am. J. Bot.** 99: e240-e244. doi:10.3732/ajb.1100547
- Yasui, Y. and O. Ohnishi. 1998. Interspecific relationships in *Fagopyrum* (Polygonaceae) revealed by the nucleotide sequences of the *rbcL* and *accD* genes and their intergenic region. **Am. J. Bot.** 85 (9): 1134-1142.
-