RESEARCH ARTICLE
doi: 10.2306/scienceasia1513-1874.2010.36.180

ScienceAsia 36 (2010): **180–186**

Leaf morphometric and genetic variation of *Butea superba* in Thailand

Jirattikarn Kaewmuangmoon[a], Trin Suwanvijitr[a], Wichai Cherdshewasart[b], Chanpen Chanchao[b,*]

[a] Program of Biotechnology, Faculty of Science, Chulalongkorn University, Bangkok 10330 Thailand
[b] Department of Biology, Faculty of Science, Chulalongkorn University, Bangkok 10330 Thailand

*Corresponding author, e-mail: chanpen@sc.chula.ac.th

Received 23 Jul 2009
Accepted 28 Jul 2010

ABSTRACT: Red kwao krua, *Butea superba* Roxb., is a herbal leguminous plant endemic in Thailand whose tuberous roots are used for male rejuvenation and the prevention of erectile dysfunction. Different populations from different provinces vary in their bioactivities, raising the need to evaluate the causes prior to optimal population selection and growth. Morphometric analysis of leaves collected from 34 populations from 24 provinces throughout Thailand using nine parameters for factor and cluster analyses were correlated with latitude and longitude revealing a clinal pattern. From the North to the South, leaf length increased in size in factor 1 but decreased in size in factor 2. Genetic analysis on the same samples, using DNA sequence analysis from the chloroplast *rbc*L, *trn*LF-cd, and *trn*LF-cf regions and maximum parsimony and neighbour-joining phylogenetic analysis, revealed essentially no genetic variation in the *rbc*L region, as expected. However, some between-population genetic variation was revealed by *trn*LF-cd and *trn*LF-cf sequences, suggesting potentially considerable genetic polymorphism. This was supported by preliminary RAPD analyses using five primers which indicated high genetic variation within and between populations.

KEYWORDS: clinal pattern, phylogeny, red kwao krua, chloroplast

INTRODUCTION

Butea superba Roxb. (Leguminosae: Papilionoideae (Faboideae)) is a large twining wood endemic in many parts of Thailand and is locally known as red kwao krua. Its tuberous roots have long been used in traditional medicine for male rejuvenation and the prevention of erectile dysfunction[1]. Morphometric and genetic analyses are almost always used to determine the differences between populations. They can be used to evaluate both fixed genetic and environmental plasticity effects upon traits. Previous morphometric studies have successfully used several plant organs such as leaf, seed, fruit, and flower. For example, Moreno-Sánchez used three parameters of leaf shape to compare plants in the genus *Archaeopteris*[2].

Geographic distance combined with habitat locality and terrain imposes both restricted gene flow and different biotic and abiotic selection pressures. This often leads to both fixed and population level genetic variation of organisms. The variation may further affect the bioactivity, characteristics, and classification of the organisms. Thus the evaluation of the genetic diversity is the first stage in understanding the different biotypes or characteristics within a population. To this end, many molecular techniques have been applied, ranging from those which do not require any specific sequence information (random screening of dominant markers), such as the somewhat unreliable rapid amplified polymorphic DNA (RAPD) analysis, the more informative amplified fragment length polymorphism analysis, and PCR restriction fragment length polymorphism (PCR-RFLP) analysis, through to the analysis of homologous DNA sequences using conserved region primers and PCR amplification, or codominant markers based upon known loci such as microsatellites, insertion sequences, and SNPs.

RAPD analysis has revealed a high level of genetic polymorphism both within and between five taxonomic groups of *Stylosanthes guianensis* which coincided with the pattern of morphological-agronomic characteristics such as seed protein pattern and pollen stainability[3]. In addition, variation in the DNA sequence of genes and intergenic regions of the chloroplast and mitochondria genome are widely used. For example, the large subunit of ribulose 1, 5 bisphosphate carboxylase/oxygenase (*rbc*L), is the most widely used gene fragment sequence in plant phylogenetic construction. Its conserved nature means it was chosen to generate a large molecular dataset of angiosperms and has been used in a large-scale analysis of green plants (asterias)[4]. In addition, the

chloroplast DNA transfer RNA-Leucine and phenylalanine region (*trn*LF) is also widely used in studying molecular phylogenetic relationships[5].

According to a large-scale survey on the distribution and diversity of kwao krua since 1998, *B. superba* was found to be widely distributed in deciduous forests in all but the southern provinces of Thailand, and is found in the same mountainous habitats as *Pueraria mirifica* (white kwao krua)[6]. Morphometric and genetic variation of kwao krua in Thailand has never been studied and remains unreported in the case of *B. superba*. We therefore aimed to determine the variation among and between *B. superba* populations in Thailand.

MATERIALS AND METHODS

Sample collection

From February 2005 to October 2006 leaves were collected from *B. superba* from 34 locations (in 24 provinces) in bamboo, dry evergreen, and mixed deciduous forests in mountainous regions in the northern, central, northeastern, eastern, and western Thailand, but not in the southern part (Table 1), in agreement with its previously reported distribution[6]. Populations were coded from the name of the province followed by a sequential number according to the locality within that province. Twenty-five mature leaves (about 2 m from the shoots) per population were picked for morphometry, and younger leaves were sampled for DNA extraction. In addition, leaves were also collected from *B. monosperma*, *P. mirifica*, and *P. lobata* for DNA extraction to use as outgroups in phylogenetic analyses.

Table 1 Locations and codes of the *B. superba* collection sites.

Code	Province	Latitude	Longitude
CT	Chantaburi	12.82	102.17
CB	Chonburi	13.3	101.31
RAT1	Ratchaburi 1	13.59	99.53
RAT2	Ratchaburi 2	13.59	99.53
RAT3	Ratchaburi 3	13.59	99.53
RAT4	Ratchaburi 4	13.59	99.53
CC	Chachoengsao	13.62	101.42
PB	Prachenburi	14.08	101.66
KC1	Kanchanaburi 1	14.63	99.09
KC2	Kanchanaburi 2	14.51	99.24
KC3	Kanchanaburi 3	14.44	99.49
SR	Saraburi	14.7	100.88
BR	Buriram	14.83	103.01
LB	Lopburi	14.94	100.77
NAK	Nakhon Ratchasima	15.01	102.17
NS1	Nakhon Sawan 1	15.7	100.09
NS2	Nakhon Sawan 2	15.7	100.09
NS3	Nakhon Sawan 3	15.7	100.09
CHY	Chaiyaphum	15.75	101.96
PC	Phetchaboon	16.26	101.08
KK	Khon Kaen	16.52	102.1
TK	Tak	16.6	98.84
PS1	Phitsanulok 1	16.9	100.48
PS2	Phitsanulok 2	16.9	100.48
PS3	Phitsanulok 3	16.9	100.48
NB	Nongbualamphu	17.25	102.25
SU	Sukhothai	17.26	99.66
SK	Sakhonnakorn	17.34	103.83
LY	Loei	17.4	101.49
UTT	Uttaradit	17.67	100.51
LP1	Lampang 1	18.33	99.51
LP2	Lampang 2	18.33	99.51
CR1	Chiangrai 1	19.92	99.93
CR2	Chiangrai 2	19.92	99.93

Morphometric parameters

Nine selected parameters were evaluated: petiole length (PL) and diameter (PD), rachis length (RL), petiolet length (PLL), terminal leaflet length and breadth (TLB), stipule length (SPL), the angle of first leaf border (AB), and the number of pairs of primary veins (NPV). All lengths were measured with a vernier calliper.

Statistical analysis

Principal component analysis (PCA) was first used on the raw data of each morphometric character from all the sampled populations (Factor; SPSS program for windows). Factor loading scores were obtained from the output and used to select characters, based upon an eigenvalue that was higher than 1.0, to provide a parsimonious reduction in the number of parameters. The selected parameters were then used for linkage cluster analysis between-groups (SPSS for windows 14.0) to investigate the relationship among populations. Finally, correlation was used to explore clinal patterns in the characteristics of *B. superba* leaves in Thailand.

DNA extraction

Genomic DNA was extracted from fresh young leaves using a DNeasy plant mini kit (Qiagen) or a NucleoSpin plant mini kit (Macherey-Nagel) according to the manufacturer's instructions. The extracted DNA was kept at $-20\,^{\circ}$C before use.

Primer design

For PCR, primers of *rbc*L were designed using the primer 3.0 program from the published sequence of

the *rbc*L gene from *Glycine max* (GenBank accession no. Z95552), with forward and reverse (329–348 and 649–630 bp, respectively, of reference sequence) primers being 5′-AGGTTCTGTTACTAACATGT-3′ and 5′-GGTCTCTCCAACGCATAAAT-3′, respectively. For amplification of a fragment of the non-coding intergenic spacer *trn*L-F region of chloroplast DNA, the forward *trn*L 5′ exon primer_c (B49317) and reverse *trn*F primer_f (A502702) from Taberlet et al were used. This was repeated using the *trn*L 3′ exon reverse primer_d[7].

For RAPD analysis, 5 primers (OPA-07, 5′-GAAACGGGTG-3′; OPA-12, 5′-TCGGCGATAG-3′; OPA-19, 5′-CAAACGTCGG-3′; OPC-15, 5′-GACGGATCAG-3′; OPD-2, 5′-GGACCCAACC-3′) were selected based on their ability to discern polymorphic bands in another leguminous plant[8]. All primers were synthesized by BioService Unit of the National Science and Technology Development Agency, Bangkok.

PCR conditions

For amplification of chloroplast loci, PCR reactions were carried out in 1 × PCR master mix solution (Intron Biotechnology), with 2 μM each of forward and reverse primer and 200 ng genomic DNA. PCR conditions for all amplifications were as follows: 94 °C for 2.5 min, 35 cycles of 94 °C for 1 min, 59 °C for 1 min, 72 °C for 3 min, and a final extension step at 72 °C for 10 min. After agarose electrophoresis, PCR amplicons of the expected sizes (300, 283, and 473 bp for *rbc*L, *trn*LF-cd, and *trn*LF-cf primers, respectively), were excised from the gel and purified using QIAquick PCR purification kit (Qiagen) as per the manufacturer's instructions. The purified PCR amplicons were then direct sequenced commercially by the BioService Unit or the Research Centre, Ramathibodi Hospital, Bangkok.

For RAPD analysis, PCR was performed as described above except for using 4 μM of one of the RAPD primers and cycle conditions were 94 °C for 2.5 min, 45 cycles of 94 °C for 1 min, 36 °C for 1.5 min, 72 °C for 3 min, and a final extension step at 72 °C for 10 min.

Phylogenetic analysis

All obtained DNA sequences were aligned using Clustal X. Phylogenetic analyses were performed by using neighbour-joining (NJ) and maximum parsimony (MP)[8]. In order to investigate the support for nodes estimated in a parsimony tree, bootstrap analysis with 1000 replicates were undertaken by UPGMA (PAUP*4.0b10).

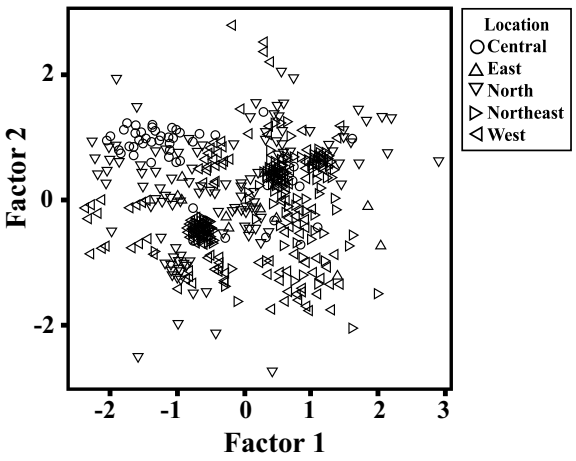

Fig. 1 Plots of factor score 1 and factor score 2 generated by PCA.

RAPD data analysis

Duplication of each reaction was performed and, after TBE-agarose electrophoresis, only bands that were reliably and accurately discriminated (i.e., showed a reproducible pattern) were scored as present or absent between samples on the same gel. For comparisons between gels, banding patterns were compared to markers and shared samples. The deduced R_f values were used to confirm band identities. In all cases scoreable bands were 0.1–1.5 kb in size. Neighbour-joining cluster analysis was performed to demonstrate the relationships among populations by considering the Nei-Li genetic distance (PAUP*4.0b10).

RESULTS AND DISCUSSION

Factor analysis

From the PCA factor analyses and factor loadings, 7 of the characters (PD, NPV, SPL, PL, RL, TLB, and AB) attained a factor loading greater than 0.6 and were selected for further analysis. The factor analysis divided these 7 parameters into three groups. The 1st factor group (PD, NPV, RL, and TLB), 2nd factor group (mainly SPL and PLL), and the 3rd factor (mainly AB) accounted for 34.2%, 17.2%, and 11.5% of the total variation, respectively. Interactive graphs were based on factors 1, 2, and 3. An example of PCA using factor 1 and 2 for the distribution of *B. superba* populations is shown in Fig. 1.

Cluster analysis

A dendrogram, constructed by a hierarchical cluster analysis of the squared Euclidean distances between values of factor scores, revealed no clear separation

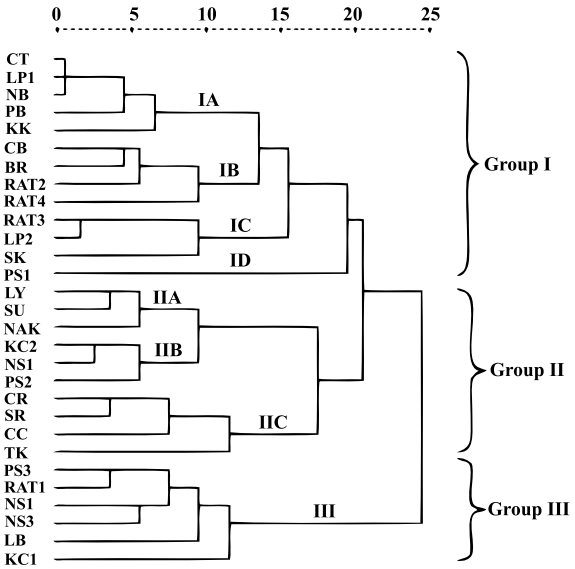

Fig. 2 A morphological dendrogram constructed by a cluster analysis.

of *B. superba* into distinct groups (Fig. 2). The dendrogram could broadly be divided into groups with the first two subdivided into four and three subgroups, respectively (Fig. 2). However, the support for such is equivocal since the factor analysis was provided by the PCA analysis. The 7 informative parameters of leaf length are not informative enough to separate the populations into groups, perhaps because only leaf parameters were used. More parameters derived from other developmentally independent organs such as seed, flower, and fruit tuber, and internodal distances parameters will be further examined to address this possibility. Alternatively, it may be possible that biotic and abiotic variations due to geography and season across these regions are not significantly different and so the morphology of the plants is not significantly affected.

Clinal patterns in the characteristic of *B. superba* in Thailand

To explore clinal patterns in the characteristics of *B. superba*, factor scores were plotted against latitude and longitude. A clear linear regression of geographic trends in morphometric characteristics of *B. superba* was revealed. Gradual transitions of characters from the South to the North and the West to the East were found (Fig. 3, Table 2). A distinct and highly significant slope ($P \leqslant 0.05$) was observed in latitude and longitude. In conclusion, from the North to the South, leaf length increases in size in factor 1 (PD, NPV, RL, and TLB) but decreases in size in factor 2 (SPL and PLL).

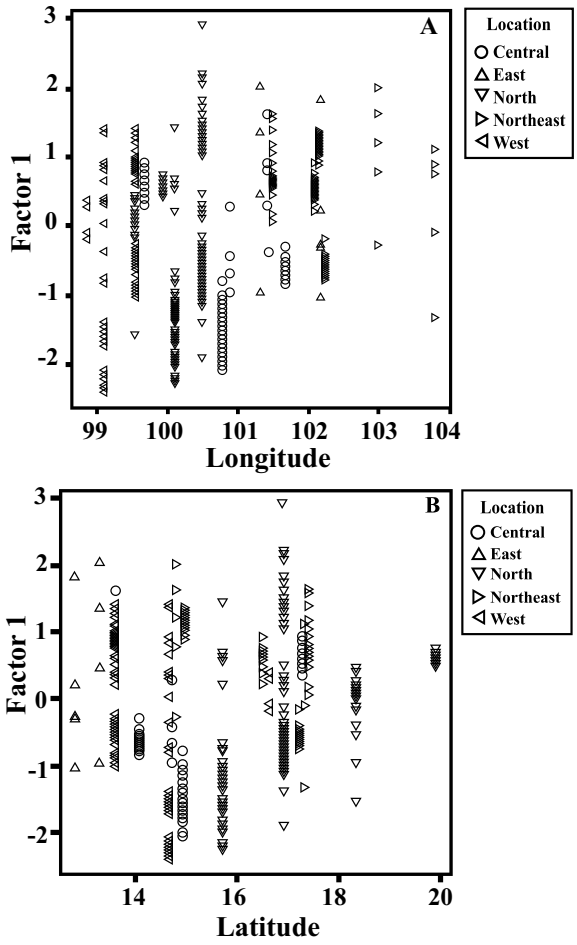

Fig. 3 Geographic trends in morphometric characters of *B. superba* in Thailand. Latitude (A) and longitude (B) plotted against factor score 1 as derived from PCA.

Table 2 Correlation of geographic trends in morphometric characters of *B. superba* in Thailand.

Factor	Latitude		Longitude	
	R value	P value	R value	P value
1	0.21*	0.00*	0.02	0.6
2	−0.24*	0.00*	0.06	0.12
3	−0.68	0.07	0.05	0.22

* Correlation is significant at the 0.01 level (2-tailed).

Genetic analysis

The expected PCR products amplified by *rbc*L, *trn*LF-cd, and *trn*LF-cf primers at 247 bp, 229 bp, and 410 bp, respectively, were obtained along with ampli-

Table 3 Summary of the amplified genes and their characteristics.

Characteristics	*rbc*L region	*trn*LF-cd	*trn*LF-cf
No. of samples analysed	32 (35)	29 (31)	25 (27)
Length range (bp)	247	229	410
AT content (%)	58.6	67.5	63.9
No. transversions	1	54	295
No. transitions	7	42	266
No. indels	0	62	102

Numbers in parentheses represent data including the outgroups which are one sample each of *B. monosperma* (only by *rbc*L), *P. mirifica*, and *P. lobata*.

cons from *B. monosperma*, *P. mirifica*, and *P. lobata* as outgroups, and DNA sequences of all PCR products were obtained by direct sequencing. Multiple alignments of the sequences were performed by Clustal X where comparisons revealed nucleotide variation in the form of single base pair substitution. The similarities in pairwise comparisons across samples of *rbc*L, *trn*LF-cd, and *trn*LF-cf sequences aligned by Clustal W ranged from 74–100%; 50–100%, and 45–100%, respectively (Table 3).

Phylogenetic analysis

Partial *rbc*L, *trn*LF-cd, and *trn*LF-cf sequences of *B. superba* plus the outgroups *B. monosperma* and *P. mirifica* and *P. lobata* were analysed to gain insight into the intergenetic and intragenetic relationships. The phylogenetic relationships were inferred by using the neighbour-joining (NJ) and maximum parsimony methods. Similar results were obtained from both methods (congruency of trees was tested by treemap; data not shown) so only the NJ phylogenetic trees are presented (Figs. 4–6).

Phylogenetic trees of *rbc*L from 32 populations of *B. superba* and three species of outgroup (*B. monosperma*, *P. mirifica*, and *P. lobata*) failed to resolve populations and even to separate *B. monosperma*, a different species within the same genus (Fig. 4) as expected and supporting the highly conserved (slow evolving) nature of this chloroplast coding gene.

The phylogenetic tree derived from *trn*LF-cd sequences from 29 *B. superba* populations and the two *Pueraria* outgroup species (Fig. 5) separated *B. superba* populations into two major groups composed of all but two samples (group I) and the SK plus CC samples (group II). Group I comprised of three minor groups (IA, IB, and IC). The IA minor group had highly similar sequences with low genetic distance. The IB and IC minor groups have higher genetic

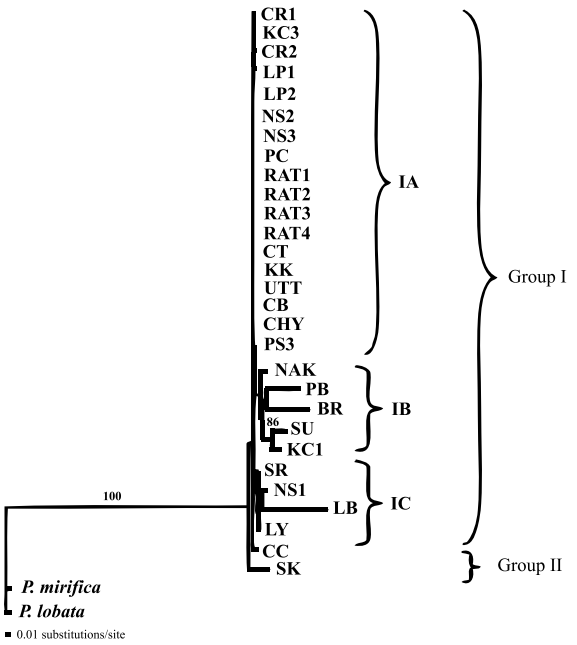

Fig. 4 A rooted phylogenetic tree of *rbc*L sequence inferred by the NJ method. No bootstrap values were above 50% for all branches. *B. monosperma* (BM), *P. mirifica*, and *P. lobata* are the outgroups.

variation. According to the IB minor group, the SU and KC1 populations formed a well separated clade (86% bootstrap support). In addition, minor splits were seen within the IC division but these had poor bootstrap support.

The phylogenetic tree derived from the larger *trn*LF-cf fragment, which encompasses all of the *trn*LF-cd fragment, derived, however, from only 25 *B. superba* populations and the two outgroups (Fig. 6) also separated *B. superba* populations into two groups. This time the second group only contained the SK population with none of the CC isolate rooted in the large 1A group. Additionally, the first group was only split into two minor groups, not three, despite the extra sequence length and number of informative characters. The failure to reform the clade 1C may perhaps be due to the lack of samples LB and LY from this analysis along with sample KC1 (from clade 1B). The two phylogenies are largely similar, as is expected given they are not only from the same gene but that *trn*LF-cf fragment contains all of the *trn*LF-cf sequence. The four minor changes (samples PS3 and LP1 shifting from clade 1A to 1B; samples SR and NS1 moving from the collapsed clade 1C to 1A) may be partly due to the missing samples, as mentioned above. However, the dramatic movement of sample

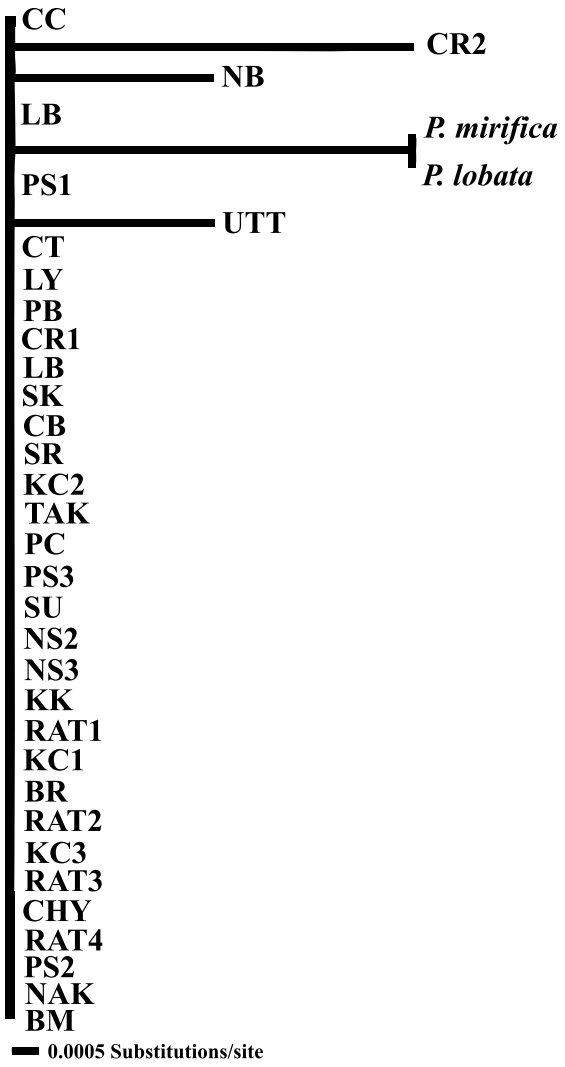

Fig. 5 A rooted phylogenetic tree of *trn*LF-cd sequence inferred by the NJ method. Bootstrap values are shown on the branches when they exceeded 50%. *P. mirifica* and *P. lobata* are the outgroups for tree rooting.

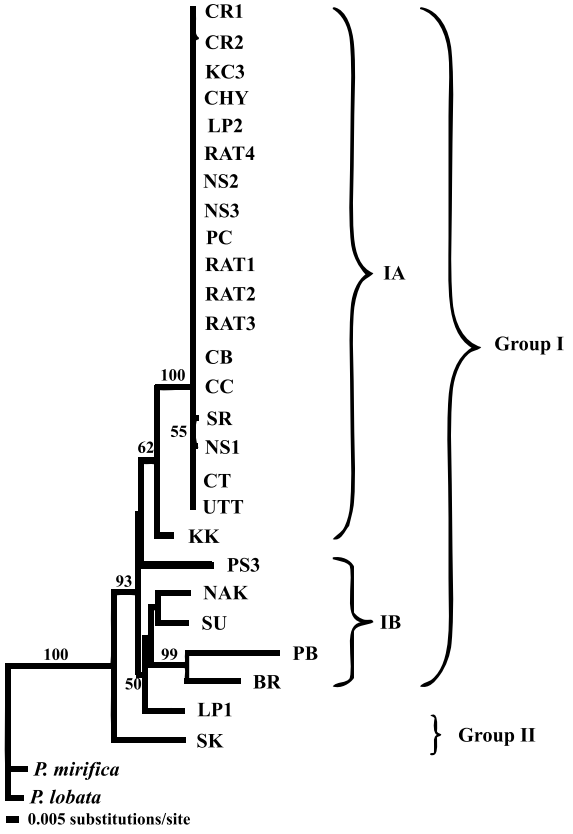

Fig. 6 A rooted phylogenetic tree of *trn*LF-cf sequences inferred by the NJ method. Bootstrap values are shown on the branches when they exceed 50%. *P. mirifica* and *P. lobata* are the outgroups for tree rooting.

CC from the distinct clade 2 (*trn*LF-cd) to clade 1A (*trn*LF-cf) is intriguing and perhaps suggests either recombination or chloroplast heterosis. Given that we direct sequenced the DNA and saw no evidence of multiple templates, to evoke some form of multiple sequence (heterosis) with two different copies of the *trn*LF region would require each copy to be defective at either the d or f primer site, which is possible, but unlikely.

Given the relatively conserved nature of the *trn*LF region, the data suggests that *B. superba* populations have potentially a relatively high level of genetic variation.

RAPD analysis

Preliminary RAPD analysis using 5 primers was used to evaluate if significant genetic diversity exists between *B. superba* populations as inferred from the results of DNA sequencing. Twenty three populations of *B. superba* were used. All 5 selected RAPD primers successfully amplified DNA from all tested *B. superba* populations producing reliably scoreable amplicons with the 5 primers which, interestingly, were all polymorphic across the 23 samples. Polymorphic bands amplified by each primer were 11, 10, 10, 10, and 7, respectively. No monomorphic bands were present suggesting a surprisingly high degree of genetic diversity. The results after scoring the 48 dominant markers were then analysed by using the neighbour-joining cluster to demonstrate the relationship among populations, using the Nei-Li genetic distance (PAUP*4.0)[8]. The derived phylogenetic tree

(Fig. 7) revealed two major separated groups (I and II). Group I (clade I) contained 19 populations of *B. superba*, had a 57% bootstrap value, and could be subdivided, albeit with only weak bootstrap support, into subclades 1A, 1B (TK and SR), and 1C (NAK, LY, and SV). This is congruent with the *trn*LF derived phylogenies. In contrast, there are 4 excepting samples. The first one is LB (clade 1C in *trn*LF to 1A in RAPD). The second one is the PB/PS3 pair which is still coresolved but moved from clade 1B (*trn*LF) to 1A (RAPD). The third one is the odd grouping of TK and SR in this RAPD based phylogeny. The last one is the unresolved KC1 sample. However, the KC1 sample does not belong to either Group I or II. Group II (Clade II) contains four populations of *B. superba* (KK, CR2, CT, and CHY) with a 65% bootstrap value which are all rather oddly from within the 1A clade of the *trn*LF phylogeny. It must be noted that this RAPD analysis not only lacks seven populations, but is also unrooted. This may account for some of the observed differences. Aside from these unresolved issues, it is obvious that group I has a relatively high genetic variation but this is not the case for those samples in group II. Also, with only five RAPD primers this preliminary study revealed nodes of closely related populations from the same province, in contrast to the chloroplast *trn*LF based sequence data. For example, the node of RAT1, RAT2, and RAT3 is close to the node of NS2 and NS3.

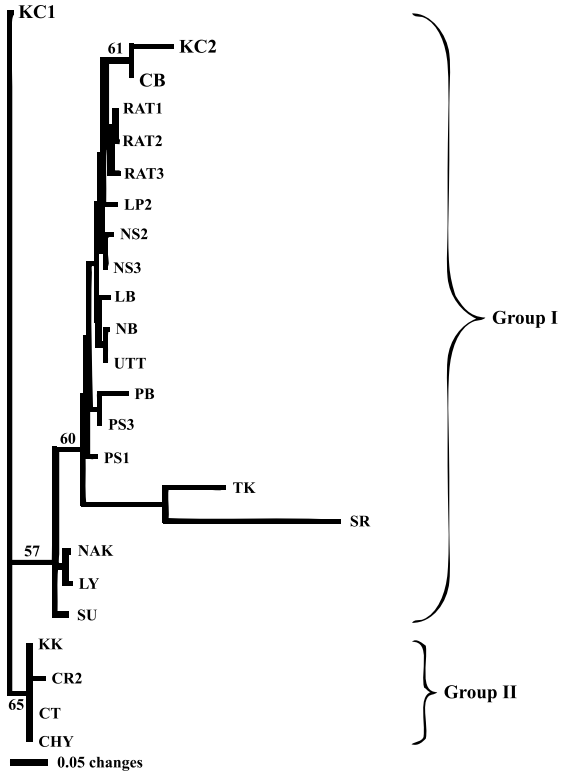

Fig. 7 A neighbour-joining tree of RAPD derived data using the Nei-Li genetic distance among 24 populations of *B. superba*.

These preliminary results should be useful in assessing the potential for genetic analysis for conservation strategies and correlating genetic relationships between geographic populations and bioactivity traits among Thai *B. superba* populations that may help to select the best population(s) of *B. superba* for pharmaceutical applications in the future.

Acknowledgements: This research was financially supported by following grants: Thailand Research Fund (RMU5180042), TRF/BIOTEC Special Program for Biodiversity Research and Training (BRT T_149006), the Research Program on Conservation and Utilization of Biodiversity of Chulalongkorn University and Asia Research Centre, Chulalongkorn University.

REFERENCES

1. Cherdshewasart W, Nimsakul N (2003) The clinical trial of *Butea superba* as an herbal alternative treatment for erectile dysfunction. *Asian J Androl* **5**, 243–6.
2. Moreno-Sánchez M (2004) Graphic approach for morphometric analysis of *Archaeopteria* leaves. *Ann Paleontol* **90**, 161–73.
3. Kazan K, Manners JM, Cameron DF (1993) Genetic variation in agronomically important species of Stylosanthes determined using random amplified polymorphic DNA markers. *Theor Appl Genet* **85**, 882–8.
4. Albach DC, Soltis PS, Soltis DE, Olmstead RG (2001) Phylogenetic analysis of asterias based on sequences of four genes. *Ann Missouri Bot Gard* **88**, 163–212.
5. Álvarez I, Wendel JF (2003) Ribosomal ITS sequences and plant phylogenetic inference. *Mol Phylogenet Evol* **29**, 417–34.
6. Cherdshewasart W, Subtang S, Dahlan W (2007) Major isoflavonoid contents of the phytoestrogen rich-herb *Pueraria mirifica* in comparison with *Pueraria lobata*. *J Pharmaceut Biomed Anal* **43**, 428–34.
7. Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol Biol* **17**, 1105–9.
8. Swofford DL (2000) *PAUP*. Phylogenetic analysis using Parsimony (* and other methods). Version 4.0b.4a*, Sinauer Associates, Sunderland, Massachusetts.