Journal of Medicinal Plants Research Vol. 6(37), pp. 5048-5058, 26 September, 2012
Available online at http://www.academicjournals.org/JMPR
DOI: 10.5897/JMPR11.1733
ISSN 1996-0875 ©2012 Academic Journals

Full Length Research Paper

Phylogenetic relationship and evolution analysis of the peony *Paeonia* species using multi-locus deoxyribonucleic acid (DNA) barcodes

Yan-Lin Sun[1,2] and Soon-Kwan Hong[2,3]*

[1]School of Life Sciences, Ludong University, Yantai, Shandong, 264-025, China.
[2]Department of Bio-Health Technology, Kangwon National University, Chuncheon, Kangwon, 200-701, Korea.
[3]Institute of Bioscience and Biotechnology, Kangwon National University, Chuncheon, Kangwon, 200-701, Korea.

Accepted 6 July, 2012

Paeonia **is a phylogenetically and taxonomically complex group. In this study, two chloroplast coding region of the** ***matK*** **and** ***rbcL*** **genes, two intergenic spacers, 5' *trnK-matK* and *matK*-3' *trnK*, the ribosomal RNA 18S and 5S genes, and the nuclear ribosomal DNA ITS gene were sequenced to study phylogenetic relationships of four** ***Paeonia*** **species. In the** ***matK*** **coding region, most nucleotide variations pointed at the divergence between** ***Paeonia suffruticosa*** **(MO) and other three species, strongly supporting the phylogenetic evolution of section** ***Paeonia*** **and section** ***Moutan*** **in the** ***Paeonia*** **genus; the variable sites of** ***Paeonia lactiflora*** **CHA1 accession compared to CHA accession were all identical to** ***Paeonia obovata*** **SHAN accession, indicating that CHA1 might be a hybrid result of CHA and SHAN. Another chloroplast coding region,** ***rbcL*** **showed lower variation rate and less phylogenetic information than the** ***matK*** **region in all four** ***Paeonia*** **species investigated in this study, while the 5S rRNA gene and nrDNA ITS region showed higher variation rate. However, combined phylogenetic relationship of all four peony species is more clearer than any single-locus one. This study would provide more sequence sources of peony species and help further understanding of the phylogenetic relationship.**

Key words: Deoxyribonucleic acid (DNA) barcoding, *matK*, nuclear ribosomal deoxyribonucleic acid internal transcribed spacer (nrDNA ITS), *Paeonia*, phylogeny, ribulose-bisphosphate carboxylase (*rbcL*), 5S ribosomal ribonucleic acid (rRNA).

INTRODUCTION

The traditional identification of herbal medicinal materials becomes difficult to satisfy increasing demands of safe use of medicinal herbs as a result of: (i) materials having similar morphological characters; (ii) materials sharing similar common names; (iii) the substitution of economically valuable materials with inexpensive herbs in wholesale markets; (iv) materials long fused and adulterated in folk; (v) processed materials or materials in powder form (Li et al., 2011). Therefore, an efficient, stable approach for clear identification of medicinal materials becomes undoubtedly beneficial. Nowadays, molecular identification sharing more advantages compared to traditional identification is believed to be a reliable alternative tool for accurate authentication of herbs (Joshi et al., 2004). Using DNA barcoding, the high-throughput standardized identification of biological samples, has become the latest move towards the generation of universal standards (Chen et al., 2010). It has been used as a forensic tool for rapidly identification based on DNA sequences in many plant species (Schindel and Miller, 2005).

Since the concept of applying DNA barcoding for the identification of global species was first proposed, considerable debate has been achieved. According to the Barcode of Life data systems (http://www.boldsystems.org),

*Corresponding author. E-mail: soonkwan@kangwon.ac.kr. Tel: 82-33-250-6476. Fax: 82-33-250-6470.

there have been more than 1,100,000 DNA barcode records of 95,000 species as of February, 2011. In NCBI GenBank database (http://www.ncbi.nlm.nih.gov), there have been over 180,000 DNA barcode records of plants as of February, 2011. The internal transcribed spacer (ITS) region of nuclear ribosomal DNA (nrDNA) comprising the ITS1 intergenic spacer, 5.8S ribosomal RNA (rRNA), and the ITS2 intergenic spacers (ITS1-5.8S-ITS2), is the most commonly used region for the barcoding and authentication of herbal medicinal materials (Coleman, 2003). Because of its high level of variation and species discrimination generally found in plant species, the ITS gene, particularly the ITS2 region with more informative variation sites, has been considered as a marker suitable for taxonomic classification over a wide range of levels (Coleman, 2003). However, the interspecific variation is sometimes low in some cases, and the cloning should be required in some cases because of the presence of multiple copies and the problems of secondary structure resulting in poor-quality sequence data (Álvarez and Wendel, 2003). Another rDNA regions, 5S and 18S genes are also considered to be one of the most variable regions commonly used in the identification of herbal medicinal materials (Chen et al., 2008; Meyer et al., 2010). Due to its high variability and discrimination ability, the 5S region has been show to successfully identify medicinal *Epimedium* L. species (Sun et al., 2004). However, this region has the significant problem of multiple copies, with molecular cloning needed in most cases.

Recently, the newly established Plant Working Group (PWG) of the Consortium for the Barcode of Life (CBOL) tested that a 2-locus combination of chloroplast coding region ribulose-bisphosphate carboxylase (*rbcL*) and maturase K (*matK*) has been found to be suitable DNA barcodes in terms of high universality, quality and coverage of sequence, and discriminatory ability (CBOL Plant Working Group, 2009). However, species discrimination of *rbcL* is sometimes believed to be insufficient, while successful amplification and sequencing of *matK* are not satisfactory due to non-universality of the amplification primers (CBOL Plant Working Group, 2009).

Compared to the coding regions of chloroplast DNA (cpDNA), non-coding region of cpDNA are presumably under less functional constraint and thus evolve more rapidly (Clegg et al., 1994). Thus, non-coding regions are usually used for phylogenetic studies, particularly at the lower taxonomical levels (Gielly and Taberlet, 1994). However, an accurate, universal, stable, and specific DNA barcode to identify the source species should possess (i) conserved flanking regions to enable routine amplification; (ii) sufficient internal variability to enable species discrimination; (iii) lengths short enough to routinely sequences even sub-optimal material; (iv) lack of heterozygosity enabling direct polymerase chain reaction (PCR) sequencing without cloning; (v) ease of alignment enabling the use of character-based data analysis methods; and (vi) lack of problematic sequence composition, that is, regions with several microsatellites, that reduces sequence quality (discussed in the review by Li et al., 2011). Until now, there has not been one potential DNA barcoding region found to fulfill all the DNA barcoding criteria listed earlier, so far. Thus, nowadays using a multiple combination for species discrimination has become the consequent choice by more and more scientists.

The genus *Paeonia* (Paeoniaceae), comprising ca. 35 species of shrubs and perennial herbs, is a phylogenetically and taxonomically complex group (Stebbins, 1938). Three sections, *Moutan*, *Paeonia*, and *Onaepia*, have been recognized within the genus. The largest section, *Paeonia*, consisting of ca. 27 herbaceous species, is believed to be taxonomically difficult as a result of hybridization. Section *Moutan* consisting of five woody species, distributed in central and western China. Section *Oneapia* containing two perennial herbaceous species, is endemic to Pacific North America (Sang et al., 1997a; Tank and Sang, 2001). The genus *Paeonia* has been placed in its own family, Paeoniaceae, and often in its own order, Paeoniales. Therefore, its broader relationships within angiosperms have been controversial (Keefe and Moseley, 1978). The phylogeny of *Paeonia* has been previously inferred based on nucleotide sequences from multiple genic and intergenic regions, including two loci of the low-copy nuclear gene alcohol dehydrogenase (*Adh1* and *Adh2*), the nuclear-encoded and chloroplast-expressed glycerol-3-phosphate acyltransferase (GPAT) gene, the cpDNA gene *mat*K, two intergenic cpDNA spacers (*trnL-trnF* and *psbA-trnH*), and the nrDNA ITS region (Sang et al., 1995; 1997a, b; Sang and Zhang, 1999; Tank and Sang, 2001). Despite the large amount of sequence data analyzed, some relationships remain unresolved, such as within subsection *Vaginitae* of section *Moutan*, and section *Paeonia* containing numerous hybrid species (Tank and Sang, 2001). Many molecular phylogenies support the monophyly of each of the three taxonomic sections of *Paeonia*, including section of *Paeonia* and *Moutan*, however, our previous result of phylogenetic relationship of Korean *Paeonia* accessions based on the nrDNA ITS sequence was not congruent with this suggestion (Sun and Hong, 2011). Specifically, our long-term aim is to find molecular means to identify species of this economically important group for conservation matters. In this study, we used DNA barcoding techniques and tested different loci including the chloroplast coding gene *rbcL* in combination with *matK*, *trnK-matK*, and *matK-trnK* intergenic spacer, and nrDNA ITS region, 5S rRNA and 18S rRNA, for further clearer phylogenetic relationship of *Paeonia*. We are also interested in (i) which of the DNA regions showed the greatest level of species discrimination; (ii) whether the monophyly of three taxonomic sections of *Paeonia* is geographical-specific; (iii) more source sequences for the accuracy of phylogeny analysis of *Paeonia*.

Table 1. Taxon, voucher specimen information, origin, collection number, and GenBank accession numbers for the investigated plant materials.

Taxon	Voucher specimen	Origin and collection number	GenBank accession number						
Section *Paeonia*			nrDNA ITS	5S rRNA	cpDNA *rbcL*	cpDNA *matK*	*trnK-matK* IGS	*matK-trnK* IGS	18S rRNA
Paeonia lactiflora pall	CHA	Korea, cult. and Hong	JN572147	JN680344	JN680348	JN712201	JN712201	JN712201	JN712205
Paeonia obovata Maxim.	SHAN	Korea, cult. and Hong (2)	JN572148	JN680345	JN680349	JN712202	JN712202	JN712202	JN712206
Paeonia lactiflora pall	CHA1	Korea, cult. and Hong (1)	JN572150	JN680347	JN680351	JN712204	JN712204	JN712204	JN712208
Section *Moutan*									
Paeonia suffruticosa Andr.	MO	Korea, cult. and Hong (3)	JN572149	JN680346	JN680350	JN712203	JN712203	JN712203	JN712207

MATERIALS AND METHODS

A total of four voucher specimens, consisting of three *Paeonia* species [*Paeonia lactiflora* (2 accessions), and *Paeonia obovata*, both belonging to section *Paeonia*, and *Paeonia suffruticosa* belonging to section *Moutan*] were included in this study. Leaves were collected from four vouch specimens, and deposited in Department of Bio-Health Technology, Plant Developmental and Engineering Laboratory (PDEL), Kangwon National University, Korea, for molecular study.

The names of these specimens are listed here together with the collectors' number have been deposited in the National Centre for Biotechnology Information (NCBI). The detailed information of the plant materials investigated in this study was shown in Table 1.

Isolation of DNA, polymerase chain reaction (PCR) amplification and sequencing

DNA extractions were performed by using the modified cetyltrimethylammonium bromide (CTAB) method described by Doyle and Doyle (1987). PCR amplification was conducted using this set of primers with the following program: 35 cycles of denaturation at 95°C for 1 min, annealing at their relevant Tms for 1 min, and a final extension step at 72°C for 1.5 min. The relevant Tms of different DNA markers were shown in Table 2. The amplification products were checked by electrophoresis through 1.0% agarose gel, and then purified before DNA sequence analysis using a QIAquick PCR Purification Kit (QIAGEN, Korea) or Gel Purification Kit (QIAGEN, Korea) according to the manufacturer's instructions. Purified PCR products were then sequenced at MACROGENE Advancing through Genomics (Korea, http://dna.macrogen.com/kor/).

Primers for amplifying and sequencing the *matK* coding region and two intergenic spacers are given in Table 2. The amplification products using the universal primers (Liang and Hilu, 1996) contained 5' *trnK-matK* intergenic spacer, complete *matK* coding gene, and 3' *trnK-matK* intergenic spacer, however, they were hard to sequence once because there are gaps between forward and reverse sequencing results. Thus, the internal primers were newly designed for use to fill gaps between sequences which were too short to overlap (Table 2).

Sequence editing and alignment

For editing and assembly of the complementary strands, the software program DNAMAN version 6.0 (Lynnon Biosoft Corporation, USA, www.lynon.com) was used. Analogue of our sequences and nucleotide sequence comparisons were detected with Basic Local Alignment Search Tool (BLAST) network services against databases (http://www.ncbi.nlm.nih.gov/). The multiple sequence alignment of ITS region (ITS1-5.8S-ITS2), 5S and 18S rRNA gene, chloroplast coding *rbcL* and *matK* genes, and *trnK-matK* and *matK-trnK* intergenic spacer of all the four *Paeonia* species were also performed using DNAMAN version 6.0 software, to detect single nucleotide polymorphisms.

Phylogenetic analysis

We assessed intra- and interspecific genetic divergences by using pairwise distance calculations (Meyer and Paulay, 2005). The phylogenetic relationships among the four *Paeonia* species was estimated after the construction of a phylogram based on multiple sequence alignment of various DNA sequences with the DNAMAN version 6.0 software (Lynnon Biosoft Corporation, USA, www.lynon.com). Phylogenetic analyses were performed initially using the 5S, and 18S rRNA, *rbcL*, and *matK* coding region, and the *trnK-matK*, *matK-trnK* intergenic spacer independently, and then using the combined data set from each plant material in these rRNA and cpDNA regions. Genetic distance (GD) was obtained with the help of MEGA software and mean GD of the intraspecific distance was calculated by sum of individual GD divide by number of samples.

RESULTS AND DISCUSSION

Due to the high primer universality and discriminatory power, the *rbcL* has been routinely used for phylogenetic studies (Schneider et al., 2004). In this study, a partial *rbcL* region with the length of from about 1250 bp (CHA) to 1280 bp (SHAN), was amplified from the *Paeonia* species using universal primers (Hasebe et al., 1994). With the analysis of sequence variation in the commonly existing sequence, a total of eight variable nucleotide sites were found, accounting for 0.64% of the whole observed sequences (Table 3). Among these variable nucleotide sites,

Table 2. Primer sequence, their relevant Tms and use.

Primer name		Primer sequence (5'-3')	Reference	Tm
trnK-matK				
MG1	a, r	CTACTACAGAACTAGTCGGATGGAGTAGAT	Liang and Hilu (1996)	61°c
MG15	a, f	ATCTGGGTTGCTAACTCAATG		
MG3	b, r	TACCTAACATAATGCATG	Newly designed in this study	47°c
MG17	b, f	TACGGATCATGACAATAA		
ITS nrDNA				
ITS4	r	TCCTCCGCTTATTGATATGC	White et al. (1990)	53°c
ITS5	f	GGAATAAAAGTCGTAACAAGG		
5S rRNA				
5SF	f	CGGTGCATTAATGCTGGTAT	Hizume (1993)	57°c
5SR	r	CCATCAGAACTCCGCAGTTA		
rbcL				
aF	f	ATGTCACCACAAACAGAGACTAAAGC	Hasebe et al. (1994)	65°c
cR	r	GCAGCAGCTAGTTCCGGGCTCCA		
18S rRNA				
18SF	f	AACCTGGTTGATCCTGCCAGT	Sogin (1990)	63°c
18SR	r	TGATCCTTCTGCAGGTTCACCTAC		

a: Primers used for PCR amplification prior to DNA sequencing; b: Internal primers used to fill gaps between sequences which were too short to overlap; f: forward primer; r: reverse primer; TM: temperature.

Table 3. Variable nucleotide sites occurring in the *rbcL* coding region among all the four *Paeonia* species.

Nucleotide site		1	4	612	662	776	999	1249	1251
Nucleotide sequence	CHA	T	T	C	T	C	A	T	C
	SHAN	-	A	A	-	-	T	-	.
	MO	C	A	A	C	G	T	C	.
	CHA1	-	-	-	-	-	-	-	.

'-' Means identity to the nucleotide site at the top row. '.'means nucleotide indel.

half of them (4) occurred between section *Moutan* species, *P. suffruticosa* (MO) and other section *Paeonia* species, *P. lactiflora* and *P. obovata* (CHA, SHAN, and CHA1); 37.5% (3) of them were found between *P. lactiflora* species and other both species; one variable site was caused by the nucleotide deletion of SHAN, MO, and CHA1 using the CHA sequence as model one. The phylogenetic relationship constructed by the *rbcL* sequence mainly showed the monophyly of section *Paeonia* and section *Moutan*, sharing the homology of 99% with each other (Figure 1). Although this result provided the evidence of supporting the monophyly in section level of genus *Paeonia*, the species discrimination was still doubted insufficient by our authors. The misgiving of *rbcL* discrimination power had previously aroused sympathy of many scientists (Kress et al., 2009), thus, to resolve the problem a more rapidly evolving region should be used in this work.

As is known, the sequence composition of MATK is more variable among different plant species than that of any other chloroplast-encoded protein (Vogel et al., 1997). Due to its high mutation rate, evolving about three times faster than *rbcL*, and lower structural conservation, *matK* gene sequence has been exploited as phylogenetic marker (Soltis et al., 1996). In this study, the entire *matK* coding region of *Paeonia* species is identical to be 1494 bp long, and 41 nucleotide substitutions were found among all the peony species, accounting for 2.74% of the full-length *matK* coding region (Table 4). Among these nucleotide sites with variation, 75.61% (31) of the nucleotide sites were obtained between section *Paeonia* species (CHA, SHAN, CHA1) and section *Moutan* (MO); 12.20% (5) were found between CHA and MO, and SHAN and CHA1; 9.76% (4) were found between SHAN,

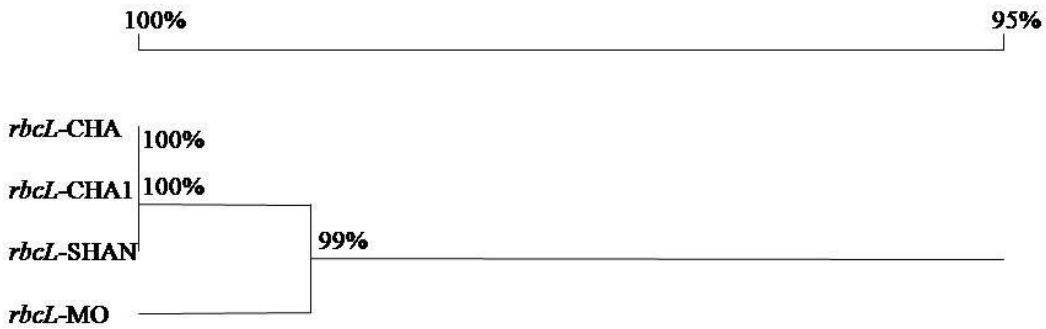

Figure 1. Phylogenetic relationships constructed using the *rbcL* chloroplast coding region.

Table 4. Variable nucleotide sites occurring in the *matK* coding region among all the four *Paeonia* species.

Nucleotide site		6	15	50	166	181	188	192	193	439	588	591	594	642	680	781	783	785	786	787	788	789
Nucleotide sequence	CHA	G	A	G	A	A	T	A	C	G	A	A	C	C	C	A	G	A	A	T	T	C
	SHAN	-	-	-	-	-	-	-	T	-	-	G	-	T	T	C	C	C	-	-	C	-
	MO	A	G	A	T	T	G	G	-	T	C	-	A	-	-	-	-	-	T	C	C	T
	CHA1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	C	C	-	-	C	-
Nucleotide site		790	791	792	793	828	850	871	873	874	880	889	932	960	991	992	1040	1084	1097	1172	1315	(bp)
Nucleotide sequence	CHA	C	T	T	T	C	C	A	C	A	G	A	C	C	C	G	C	T	C	T	A	
	SHAN	-	C	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	MO	T	-	C	-	T	T	C	T	C	A	C	T	T	A	T	T	C	T	C	G	
	CHA1	-	C	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	

'-' Means identity to the nucleotide site at the top row.

and other three species; 2.43% (1) were found between CHA and other three species. In addition, even between both different accessions of the same species, *P. lactiflora* (CHA and CHA1), there were also nucleotide substitutions observed. However, these nucleotide substitutions between CHA1 and CHA could just be identical to the nucleotide sequence of *P. obovata* (SHAN). This result indicated that CHA1 investigated in this study might be a hybridization result of *P. lactiflora* (CHA) and *P. obovata* (SHAN). This suggestion was also supported by the phylogenetic tree constructed by the *matK* region (Figure 2A). As known, hybridization behavior was not scarce the peony species, the reticulate evolution has been found in *Paeonia* species, especially section *Paeonia* (Sang et al., 1995): five *Paeonia* species including *Paeonia banatica*, *Paeonia russi*, *Paeonia emodi*, *Paeonia sterniana*, and *Paeonia peregrine* have been shown perfect or almost perfect additivity at all sites that are variable between two or three other species or species groups based on the internal transcribed spacers (ITS) sequence or nuclear ribosomal DNA (Sang

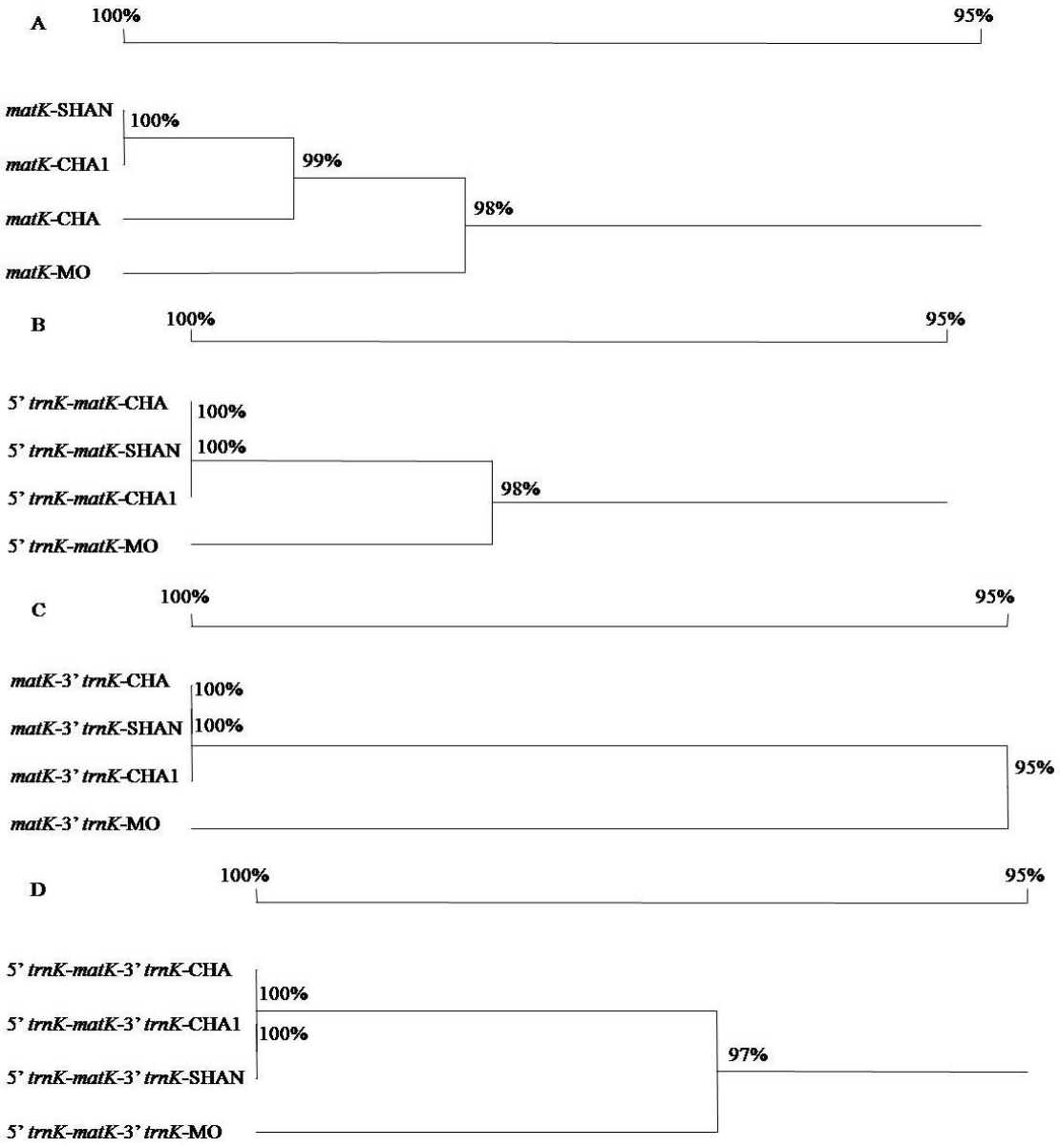

Figure 2. Phylogenetic relationships constructed using the *matK* coding region (**A**), noncoding 5' *trnK-matK* intergenic spacer (**B**), noncoding *matK*-3' *trnK* intergenic spacer (**C**), and the *trnK-matK-trnK* intron region (**D**).

et al., 1995); and the hybridization of *Paeonia* species was further investigated and evaluated using *matK* gene sequence with fixation of ITS sequence (Sang et al., 1997). Although no evidence has yet showed that *P. lactiflora* is a hybrid species, our result strongly supported CHA1 accession of *P. lactiflora* species was a hybrid species of *P. lactiflora* CHA accession and *P. obovata* SHAN accession.

To understand the phylogenetic relationship, the approach of using one coding region in combination with noncoding regions, that most likely evolve more rapidly than coding region, is widely accepted nowadays (de Groot et al., 2011). The *trnK* intron including the 5' *trnK-matK* and *matK*-3' *trnK* integenic spacer, was therefore selected and investigated for the purpose of phylogenetic analysis in *Paeonia* species. With this straightforward sequence alignment of 5' *trnK-matK* intergenic spacer, a total of 23 variable sites (including nucleotide substitutions and indels) occurred among these species, of which 30.43% (7) of variable sites were caused by nucleotide indels (Table 5). However, all the variable sites were found between section *Moutan* species, *P. suffruticosa* (MO) and section *Paeonia* species, *P. lactiflora* and *P. obovata* (CHA, SHAN, and CHA1). For the *matK*-3' *trnK* intergenic spacer, 16 variable sites (including nucleotide substitutions and indels) were detected, of which 6 sites were because of nucleotide indels (Table 6). The nucleotide variations occurring in

Table 5. Variable nucleotide sites occurring in the noncoding 5' *trnK-matK* intergenic spacer among all the four *Paeonia* species.

Nucleotide site		214	323	353	443	444	445	446	447	464	471	484	498	504	517	536	551	559	564	570	580	630	633	696
Nucleotide sequence	CHA	T	C	.	T	T	T	C	T	G	T	T	G	T	T	G	A	G	T	T	T	T	C	A
	SHAN	-	-	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MO	C	T	T	.	.	.	.	.	A	G	A	C	G	G	T	C	A	G	G	G	G	A	.
	CHA1	-	-	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

'-' Means identity to the nucleotide site at the top row. '.'means nucleotide indel.

Table 6. Variable nucleotide sites occurring in the noncoding *matK*-3' *trnK* intergenic spacer among all the four *Paeonia* species.

Nucleotide site		30	40	67	86	88	119	145	146	147	148	149	150	151	152	159	225
Nucleotide sequence	CHA	A	T	T	G	C	.	T	G	A	T	T	T	T	C	A	G
	SHAN	-	-	-	-	-	.	-	-	-	-	-	-	-	-	-	-
	MO	G	C	C	A	T	C	C	T	G	.	.	.	.	.	C	T
	CHA1	-	-	-	-	-	.	-	-	-	-	-	-	-	-	-	-

'-' Means identity to the nucleotide site at the top row. '.'means nucleotide indel.

this intergenic spacer were also owed to the sequence divergence of *P. suffruticosa* (MO) and other three species, similar to that the 5' *trnK-matK* intergenic spacer shown. According to the above results, it suggested that both noncoding intergenic spacers investigated in this study provided few nucleotide variations and less phylogenetic information than that *matK* coding region did, however, the nucleotide variations occurring in these both noncoding intergenic spacers more clearly denoted the phylogenetic relationship in the section level, while the nucleotide variations occurring in the *matK* coding region likely showed more hybridization information. This suggestion was also supported by the phylogenetic tree constructed by both intergenic spacers independently (Figures 2B and C).

Combined with the obtained sequencing results of *trnK- matK-trnK* intron region, three section *Paeonia* specieswere found to have nearly 100% homology with each other, and 97% homology with one section *Moutan* species, *P. suffruticosa* (MO, Figure 2D). The recognized monophyly of section *Paeonia* and section *Moutan*, that was not congruent with our previous study using only nrDNA ITS region (Sun and Hong, 2011), was strongly supported here by the phylogenetic analysis based on the *trnK* intron region.

The rRNA gene sequences are found easy to access due to highly conserved flanking regions allowing for the use of universal primers (Chenuil, 2006). The 18S rRNA gene is a part of the ribosomal functional core and is exposed to similar selective forces in all living beings (Moore and Steitz, 2002). Therefore, the 18S gene is widely used in phylogenetic studies and biodiversity screening due to the primer university (Chenuil, 2006; Meyer et al., 2010). In this study, the 18S rRNA gene was also used for analyzing the phylogenetic relationships of *Paeonia* species, but the low species discrimination within all the *Paeonia* species was found. Among all the species, only one variable site and five indels were detected (Table 7), accounting for 0.35% of all observed 18S gene sequence. Because of little phylogenetic information, phylogenetic reconstruction was not performed for this region alone. The low sequence divergence value was also found within the *Prasiolales* based on the 18S rRNA gene, ranging from 0.4% to 3.8% (Sherwood et al., 2000).

Another rRNA gene, 5S residing as independent tandem array in a plant genome, is used globally as molecular markers for resolving species-level phylogenetic relationships (Persson, 2000). Using universal primers reported by Hizume (1993), the length of the 5S rRNA varied from 528 bp (CHA1) to 557 bp (SHAN). Among about 529 bp of commonly existing sequences, a total of 170

Table 7. Variable nucleotide sites occurring in the 18S rRNA region among all the four *Paeonia* species.

Nucleotide site		7	1043	1715	1720	1722	1723
Nucleotide sequence	CHA	.	.	.	A	.	.
	SHAN	A	.	T	G	A	G
	MO	.	C	T	-	.	.
	CHA1	A	.	T	-	.	.

'-' means identity to the nucleotide site at the top row. '.'means nucleotide indel.

Figure 3. DNA alignment of 5S rRNA gene sequence of four accessions of three peony species. – means identity to the top row sequence; means nucleotide indel in this site; means no additional nucleotide sequence.

variable sites (including nucleotide substitutions and indels) occurred among these species (Figure 3). This high nucleotide variation rate (32.14%) of the whole observed 5S rRNA sequence showed high potential to provide more phylogenetic information than that 18S rRNA sequence did, indicating that 5S rRNA gene was more suitable in phylogenetic study of *Paeonia* species. Among all the variable sites, only 23.33% (35) showed nucleotide diversity between CHA and CHA1, 30% (51) occurred between *P. lactiflora* (CHA and CHA1) and other species (SHAN and MO), 29.41% (50) between *P. obovata* (SHAN) and other species, 14.71% (25) between *P. suffruticosa* (MO) and other species, and other 5.29% (9) among three species. Accessibly, both accessions from *P. lactiflora* (CHA and CHA1) showed relatively high homology of 94% with each other compared to with other species (Figure 4). However, *P. suffruticosa* (MO) was more closely near to *P. lactiflora* (CHA and CHA1) than *P. obovata* (SHAN) that belongs to the same section *Paeonia* as *P. lactiflora* but not *P. suffruticosa*. The similar result had also beenobtained from our previous phylogenetic study in *Paeonia* species based on the nrDNA ITS region (Sun and Hong, 2011). In the present study, the occurrence of this result was also understandable when aligning the 5S rRNA sequence among all the species. SHAN had even two times of specific, variable nucleotide sites in the 5S rRNA sequence than MO, and SHAN also shared so many nucleotide substitutions (51 variable sites) with MO. For the nrDNA ITS region, a total of 28 variable sites was found, of which 14 were located in the ITS1 region, 13 in the ITS2 region, only one in the 5.8S rRNA region (Table 8). Among these variable nucleotide sites, 42.85% (12) occurred between *P. lactiflora* and other species, 32.14% (9) were MO-specific variable nucleotide sites, 21.43%(6) were SHAN-specific ones. Compared with the nrDNA ITS region of both accessions of *P. lactiflora*, only one nucleotide variation occurred, located in the ITS1 region. Due to insufficient divergence information included in the nrDNA ITS region, whether CHA1 was the hybrid species of CHA and SHAN could not be obtained, based on this region only. Therefore, using multi-locus DNA barcode as phylogenetic markers becomes necessary.

To further understand the phylogenetic relationship of all four *Paeonia* species, three regions sharing relatively high nucleotide variation rates and more phylogenetic information were selected and combined to use for

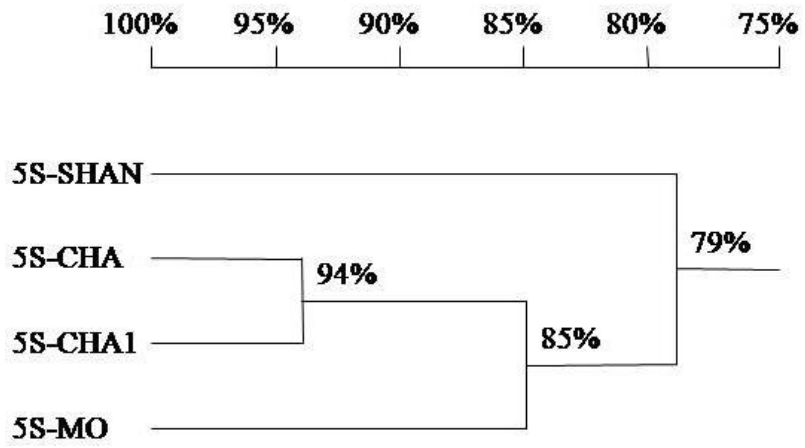

Figure 4. Phylogenetic tree of genus *Paeonia* generated using 5S rRNA gene sequence.

Table 8 Variable nucleotide sites occurring in the nrDNA ITS region among all the four *Paeonia* species.

Nucleotide site		45	49	65	78	84	89	104	121	127	134	135	165	222	223	383	435	453	457	464	508	509	518	605	617	624	638	644	649
Nucleotide sequence	CHA	A	C	T	T	G	G	T	A	A	A	G	C	A	G	G	C	A	T	T	T	A	T	G	G	A	T	A	T
	SHAN	G	-	C	-	-	-	G	C	C	T	C	A	G	A	A	T	C	G	C	C	-	C	T	-	-	-	-	C
	MO	C	G	C	C	A	A	G	C	-	-	-	A	G	A	-	-	C	G	C	C	G	C	T	A	G	C	.	-
	CHA1	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

'-' Means identity to the nucleotide site at the top row. '.'means nucleotide indel.

constructing the phylogenetic tree, including 5S rRNA, *trnK* intron (5' *trnK*-matK-3' *trnK*), and nrDNA ITS (Figure 5). The combined phylogenetic tree illustrated that CHA, SHAN, and CHA1 all belonging to section *Paeonia* were monophyletic, supporting the hypothesis of the monophyly of section *Paeonia*; MO belonging to section *Moutan*, shared 95% homology with *Paeonia* monophyletic group. This combined result also implied that our previous suggestion evaluated by nrDNA ITS only was insufficient and unilateral.

Conclusion

This work attempted seven DNA barcodes (*rbcL* chloroplast coding region, *matK* coding region, 5' *trnK-matK* and *matK*-3' *trnK* intergenic spacers, 18S and 5S rRNA genes, nrDNA ITS region) to study the phylogenetic relationship of four peony species. Among these DNA barcodes, *matK* coding region, 5S rRNA and nrDNA ITS shared relatively more variation information than other DNA barcodes. More importantly, the *matK* coding region expressed not only the phylogenetic relationship in section level but the hybridization information. The evolution of all four *Paeonia* species was shown more clearly in combination with multi-locus DNA barcodes than that the single-locus on e did. Thus, this work not only

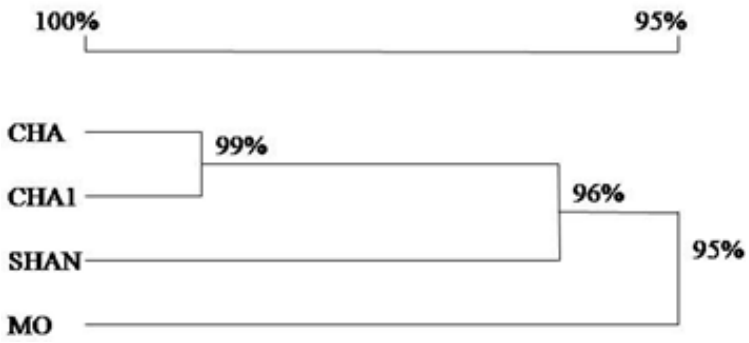

Figure 5. Phylogenetic tree of genus *Paeonia* generated using combined regions with 5S rRNA, *trnK* intron, and nrDNA ITS region.

provided more sequence sources of the peony species, but helped further understand the phylogenetic relationship of the taxonomically complex *Paeonia* species.

ACKNOWLEDGEMENT

This work was supported by Nutraceutical Bio Brain Korea 21 Project Group.

REFERENCES

Álvarez I, Wendel JF (2003). Ribosomal ITS sequences and plant phylogenetic inference. Mol. Phyl. Evol. 29:417-434.

CBOL Plant Working Group (2009). A DNA barcode for land plants. PNAS, USA 106:12794-12797.

Chen F, Chan HYE, Wong KL, Wang J, Yu MT, But PPH, Shaw PC (2008). Authentication of *Saussurea lappa*, an endangered medicinal material, by ITS DNA and 5S rRNA sequencing. Planta Med. 74:889-892.

Chen SL, Yao H, Han JP, Liu C, Song JY, Shi LC, Zhu YJ, Ma XY, Gao T, Pang XH, Luo K, Li Y, Li XW, Jia XC, Lin YL, Leon C (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS, ONE 5:e8613.

Chenuil A (2006). Choosing the right molecular genetic markers for studying biodiversity: from molecular evolution to practical aspects. Genetica 127:101-201.

Clegg MT, Gaut BS, Learn GHJr, Morton BR (1994). Rates and patterns of chloroplast DNA evolution. Proc. Natl. Acad. Sci. USA 91:6795-6801.

Coleman AW (2003). ITS2 is a double-edged tool for eukaryote evolutionary comparisons. Trends Genet. 19:370-375.

de Groot GA, During HJ, Maas JW, Schneider H, Vogel JC, Erkens RHJ (2011). Use of *rbcL* and *trnL-F* as a two-locus DNA barcode for identification of NW-European ferns: an ecological perspective. PLoS, ONE 6:e16371.

Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19:11-15.

Gielly L, Taberlet P (1994). The use of chloroplast DNA to resolve plant phylogenies: noncoding versus *rbc*L sequences. Mol. Biol. Evol. 11:769-777.

Hasebe M, Omori T, Nakazawa M, Sano T, Kato M, Iwatsuki K (1994). *rbcL* gene sequences provide evidence for the evolutionary lineages of leptosporangiate ferns. Proc. Natl. Acad. Sci. USA 91:5730-5734.

Hizume M (1993). Chromosomal localization of 5S rRNA genes in *Vicia faba* and *Crepis capillaries*. Cytologia 58:417-421.

Joshi K, Chavan P, Warude D, Patwardhan B (2004). Molecular markers in herbal drug technology. Curr. Sci. 87: 159-165.

Keefe JM, Moseley MFJr (1978). Wood anatomy and phylogeny of *Paeonia* section *Moutan*. J. Arnold Arbor. 59:274-297.

Kress WJ, Erickson DL, Jones FA, Swenson NG, 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.

Li M, Cao H, But PPH, Shaw PC (2011). Identification of herbal medicinal materials using DNA barcodes. J. Syst. Evol. 49:271-283.

Liang HP, Hilu KW (1996). Application of the *matK* gene sequences to grass systematics. Can. J. Bot. 74:125-134.

Meyer A, Todt C, Mikkelsen NT, Lieb B (2010). Fast evolving 18S rRNA sequences from Solenogastres (Mollusca) resist standard PCR amplification and give new insights into mollusk substitution rate heterogeneity. BMC Evol. Biol. 10:70.

Meyer CP, Paulay G (2005). DNA barcoding: error rates based on comprehensive sampling. PLoS Biol. 3:e422.

Moore PB, Steitz TA (2002). The involvement of RNA in ribosome function. Nature 418:229-235.

Persson C (2000). Phylogeny of the neotropical *Alibertia* group (Rubiaceae), with emphasis on the genus *Alibertia*, inferred from ITS and 5S ribosomal DNA sequences. Am. J. Bot. 87:1018-1028.

Sang T, Crawford DJ, Stuessy TF (1995). Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. Proc. Natl. Acad. Sci. USA 92:6813-6817.

Sang T, Crawford DJ, Stuessy TF (1997a). Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). Am. J. Bot. 84:1120-1136.

Sang T, Donoghue MJ, Zhang D-M (1997b). Evolution of alcohol dehydrogenase genes in peonies (*Paeonia*): phylogenetic relationships of putative nonhybrid species. Mol. Biol. Evol. 14:994-1007.

Sang T, Zhang DM (1999). Reconstructing hybrid speciation using sequences of low copy nuclear genes: hybrid origins of five *Paeonia* species based on *Adh* gene phylogenies. Syst. Bot. 24:148-163.

Schindel DE, Miller SE (2005). DNA barcoding a useful tool for taxonomists. Nature 435:17.

Schneider H, Russell SJ, Cox CJ, Bakker F, Henderson S, Rumsey F, Barrett J, Gibby M, Vogel JC (2004). Chloroplast phylogeny of asplenioid ferns based on *rbcL* and *trnL-F* spacer sequences (Polypodiidae, Aspleniaceae) and its implications for biogeography. Syst. Bot. 29:260-274.

Sherwood AR, Garbary DJ, Sheath RG (2000). Assessing the phylogenetic position of the Prasiolales (Chlorophyta) using *rbcL* and 18S rRNA gene sequence data. Phycologia 39:139-146.

Sogin ML (1990). Amplification of ribosomal RNA genes for molecular evolution studies. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ eds. PCR protocols. A guide to methods and applications. New York: Academic Press. pp. 307-322.

Soltis DE, Kuzoff RK, Conti E, Gornall R, Ferguson K (1996). *matK* and *rbcL* gene sequence data indicate that Saxifraga (Saxifragaceae) is

polyphyletic. Am. J. Bot. 83:371-382.
Stebbins GLJr (1938). Notes on the systematic relationships of the Old World species and of some horticultural forms of the genus *Paeonia*. University of California Publications in Botany. 19:245-266.
Sun Y (2004). Molecular authentication of *Eimedium* species and the anti-osteoporosis effect of herba *Epimedi*. PhD dissertation, The Chinese University of Hong Kong.
Sun YL, Hong SK (2011). Genetic diversity and phylogenetic analysis of genus *Paeonia* based on nuclear ribosomal DNA ITS sequence. J. Plant Biotechnol. 38: 234-240.
Tank DC, Sang T (2001). Phylogenetic utility of the glycerol-3-phosphate acyltransferase gene: evolution and implications in *Paeonia* (Paeoniaceae). Mol. Phyl. Evol. 19:421-429.
Vogel J, Hübschmann T, Börner T, Hess WR (1997). Splicing and intro-internal RNA editing of *trnK-matK* transcripts in barley plastids: support for MatK as an essential splice factor. J. Mol. Biol. 270:179-187.
White TJ, Bruns T, Lee S, Taylor J (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ eds. PCR protocols-a guide to methods and applications. San Diego, Calif: Academic Press. pp. 315-322.