ScienceAsia 29 (2003): **13-20**

Molecular Characterization of the Divergence of Rare Species of the Genus *Afgekia* (Papilionoideae, Tribe Tephrosieae) by RAPD Markers and Nucleotide Sequences Analysis

Preecha Prathepha[a,*] and Visut Baimai[b]

[a] Department of Biotechnology, Faculty of Technology, Mahasarakham University, Mahasarakham 44000 Thailand.

[b] Department of Biology, Faculty of Science, Mahidol University, Bangkok 10400 Thailand.

* Corresponding author, E-mail: preecha.p@techno.msu.ac.th

Received 22 Feb 2001
Accepted 7 Oct 2002

ABSTRACT The two taxa of the genus *Afgekia*, namely *Afgekia sericea* and *A. mahidolae* are endemic to Thailand. The former species is widely distributed in northeastern Thailand, while the latter has been found only in western Thailand. We previously reported on morphological and cytological data and proposed that these two taxa resulted from allopatric speciation of an ancestral species in different geographical populations. This work provides further characterization of the relationships between them at the DNA level.

According to RAPD data, *A. sericea* and *A. mahidolae* exhibited a high value of genetic distance (D= 0.55). Furthermore, the nucleotide sequences of the internal transcribed spacer of nuclear ribosomal DNA (nrDNA ITS1 and ITS2) and the chloroplast *rbcL* gene of the two taxa were compared. For the ITS regions, the sequence divergence of these spacers between the taxa was very high. This indicated that these two species are probably not recently diverged. For the partial nucleotide sequence of *rbcL* gene, both the divergence at the nucleotide and amino acid level were higher than expected (5.6% for nucleotide, 5.1% for amino acid). These data are likely to reflect the evolutionary process which led to the present patterns of phenotypic and genetic variation between the two taxa.

KEYWORDS: *Afgekia*, Leguminosae, molecular evolution, rare plant.

INTRODUCTION

The genus *Afgekia* was proposed by WG Craib in 1927 when he first discovered *Afgekia sericea* Craib in northeastern Thailand. Then, in 1972 Burtt and Chermsirivathana found the second species, *A. mahidolae* in western Thailand[1]. Moreover, *Adinobotrys filipes* was revised by Geesink[2] and he placed it in the genus *Afgekia*. At present time, the genus *Afgekia* is still a small genus consisting of these three species. Prathepha and Baimai[3] reported that *A. sericea* and *A. mahidolae* are likely to be allied species. In addition, the first two species are morphologically similar while the other species are quite remote from them. These two species are wild, woody climbers with showy inflorescence. *A. sericea* and *A. mahidolae* differ in a large number of characters, most of which pertain to differences in floral structures.[3] Until now, these two taxa have been reported only in Thailand.

Rabinowitz[4] recognized that species exhibiting the most vulnerable form of rarity are characterized by a narrow geographic range, high specificity to habitat, and low population densities. The two taxa of interest of the genus *Afgekia* are species that fit this form of rarity. For *A. mahidolae*, the rarity is indicated by its endemic distribution in western Thailand, its habitat is restricted to limestone soils and its low population densities.[3] *A. sericea* is not characterized by specificity to habitat, but it has low population densities due to deforestation in northeastern Thailand. Based on this evidence, this species could be recognized as a rare species.[5] These two species are diploid (2x=16) with similar chromosome characteristics.[3,6] Molecular evolution studies of the two taxa have not been reported. Thus, this study is the first preliminary molecular research on this genus.

Molecular markers and molecular sequences contain useful information about evolutionary history.[7,8] Recently, the use of random amplified polymorphic DNAs (RAPDs) has become popular. The arbitrarily primed polymerase chain reaction (or RAPD) amplifies anonymous fragments of DNA from any genome.[9,10] The size distribution of amplified fragments varies among species. However, closely related species have similar fragment distribution, while distantly related ones are more divergent.[11] Thus, RAPD bands

(fragments) distribution contain considerable phylogenetic information.[12-15] The internal transcribed spacer (ITS) of nuclear ribosomal DNA (nrDNA) have been widely used to resolve phylogenetic relationships for many plant taxa.[16-18] The ITS has several advantages that make it ideal to sequence for phylogenetic analysis of congeneric species such as its rate of evolution is appropriate for studies at the species levels, and it is phylogenetically interpretable, *ie*, the sequences are relatively easy to align because there tends to be very little length variation at the generic levels in flowering plants. The ITS are also quite variable among closely related species of animals.[19]

In addition, analysis of chloroplast DNA (cpDNA) sequence variation represents one of the most powerful tools used by scientists to infer phylogenetic relationships of plants.[20,21] An effective effort has emerged to produce a detailed phylogeny of flowering plants based on sequences of the chloroplast-encoded large subunit of ribulose-1,5-bisphosphate carboxylase (*rbcL*).[7] The sequences of the *rbcL* gene have been used to reconstruct the phylogeny of plants.[22] Since the *rbcL* gene exists in a single copy in the chloroplast genome, it is not necessary to be concerned about intragenic recombination and inappropriate comparisons among paralogous loci.[23] Based on the comparative sequence of *rbcL*, Kaess and Wink[24] provided interesting information on the relation among the Papilionoideae.

In this study we describe the genetic relationships between of *A. sericea* and *A. mahidolae* by using RAPD data and nucleotide sequence analysis of the internal transcribed spacer regions and *rbcL* gene.

Materials and Methods

Plant Materials

The plant materials used in this study were collected from natural habitats. Nine individuals were sampled per species from one population in Chaiyaphum province northeastern Thailand, for *A. sericea* and Kanchanaburi, western Thailand for *A. mahidolae*. Voucher specimens were deposited in a greenhouse of Mahasarakham University.

Total DNA Extraction

DNA was isolated from fresh leaves according to the CTAB method of Doyle and Doyle.[25] DNA was isolated from nine individuals plants of each species.

RAPD analysis

The DNA samples were used for random amplification according to Prathepha and Baimai.[3] Nine random 10-mer primers (OPAS-03,04,05,06, 07, 09, 10, 12 and 19) shown in Table 1 were used for the experiment. The 18 plants examined of the two taxa yielded 83 bands. These scored bands were strong and reproducible.

Table 1. Sequences of random primers used in RAPD analysis of *A. sericea* and *A. mahidolae*.

Code	Sequence (5' to 3')
OPAS-03	ACGGTTCCAC
OPAS-04	GTCTTGGGCA
OPAS-05	GTCACCTGCT
OPAS-06	GGCGCGTTAG
OPAS-07	GACGAGCAGG
OPAS-09	TGGAGTCCCC
OPAS-10	CCCGTCTACC
OPAS-12	TGACCAGGCA
OPAS-19	TGACAGCCCC

Analysis of RAPD Data

The presence / absence of bands was recorded in binary (0,1) form. Genetic relationships between the two taxa was determined by using the index of genetic distance (D), (D = 1-S). S values were obtained by using a method based on the theory of Nei and Li.[26] This method involved estimating the fraction of (shared fragments (S) between the two taxa, using the formula: $S = 2m_{xy} / (m_x + m_y)$, where m_{xy} is the number of RAPD bands shared by the two taxa and m_x and m_y represent the total number of RAPD bands present in each taxon. S can take any value between 0 and 1, with 0 meaning that no common bands exist, and 1 meaning that all bands are identical between these two taxa. In addition, S value was calculated for each primer and mean S value obtained from the average over all primers.

Amplification of ITS1 and ITS2 Regions

The complete ITS regions was amplified using ITS primers designed by Mummenhoff *et al*[27] as shown in Fig 1. Amplifications were performed in a volume of 50 ml containing 10 mM Tris-HCl (pH 9.0), 50 mM KCl and 0.1% Triton®x-100, 1.5 mM $MgCl_2$, 40 mM each of dNTPs, 10 mM of each primer, and approximately 150 ng genomic DNA, and 1 U of *Taq* polymerase (Promega). Amplification was performed in a Hybaid thermal cycler, using the following parameters: 94°C for 3 min, followed by 35 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 2 min.

Amplification of *rbcL* Gene

Primers for the *rbcL* gene were used in this study according to Kaess and Wink.[24] The primer pairs are: *rbcL*N, 5′ ATG TCA CCA CAA ACA GAA ACT AAA GC 3′ and *rbcL*R, 5′ TAT CCA TTG CTG GGA ATT CAA ATT TG 3′. Primer pairs used for PCR flank the beginning

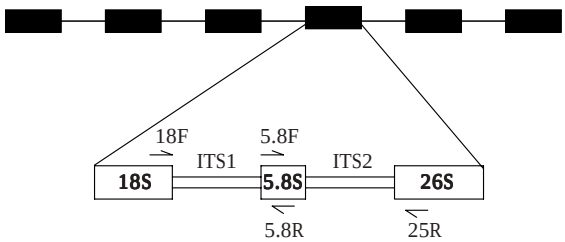

Fig 1. Schematic diagram of the repeated unit of 18S, 5.8S and 26Snuclear ribosomal DNA showing the internal transcribed spacer (ITS) regions. The positions of primers used for DNA amplification and sequencing of the ITS1 and ITS2 regions are indicated by arrows. Primer sequence are as follows:

18F:	5′	GGA	AGG	AGA	AGT	CGT	AAC	AAG	G 3′
5.8F:	5′	CTT	CTG	GCC	GAG	GGC	ACG	TC	3′
5.8R:	5′	GCT	ACG	TTC	TTC	ATC	GAT	GC	3′
25R:	5′	TCC	TCC	GCT	TAT	TGA	TAT	GC	3′

and the end of the *rbcL* gene.The PCR reaction mixture contained approximately 150 ng of total DNA, 20 picomoles primer pairs *rbcL*-N/ *rbcL*-R, 1.5mM $MgCl_2$, 100mM dNTPs, and 2 Units *Taq* polymerase (Promega). PCR cycle parameters were set as follows after an initial denaturation step for 3 min at 94°C: denaturation of template DNA for 1 min at 94°C, primer annealing for 1 min at 45°C, primer extension for 2 min at 72°C. After 35 cycles a final extension step of 5 min at 72°C was added to allow completion of unfinished strands. Ten microliters of PCR products of ITS 1 and 2 regions and *rbcL* gene were checked for successful amplification and quantified by agarose gel electrophoresis at 75 V for 2 h in a 1.4% agarose (FMC) gel using 0.5x TBE as the gel buffer.

Sequencing Protocol

The PCR products were purified prior to sequencing using QIAquick Gel extraction kit (Qiagen). Direct sequencing was performed in both directions from the double-stranded DNA fragment using the amplification primers. The Taq Dye Terminator Cycle Sequencing kit (Applied Biosystems) was used as recommended by the manufacturer. The PCR products were analyzed on an ABI 737A autosequencer. For each taxon, forward and reverse sequencing reactions were performed for sequence confirmation.

Sequence Analysis

Nucleotide sequences of the PCR products were aligned with each other and sequence from other species for the beginning and the end of ITS1 and 2 and the *rbcL* gene fragment using the Clustal W computer program (Figs 3-4).[28] GC or AT content was manually calculated. The *rbcL* genes were translated to amino acid sequence using the plant plastid code and aligned using Clustal W (Figs 5-6).

RESULTS

RAPD analysis

Nine random primers used in this experiment illustrated that the two taxa could be distinguished by RAPD markers. The RAPD markers generated by some primers are shown in Fig 2a-d. These primers generated from 7 to 14 bands, ranging from 0.35 to 3.2 kb in size. Either monomorphic or species-specific bands were found between the two taxa. For example, two species-specific bands, OPAS03-950 bp and OPAS03-1200 bp were produced by the OPAS03 primer providing marker bands for *A. sericea* and *A. mahidolae*, respectively (Fig 2a). Genetic distance (D) between these two species varied from 0.67 to 0.36 operated by OPAS10 and OPAS12 , respectively. The average genetic distance of the two species was 0.55.

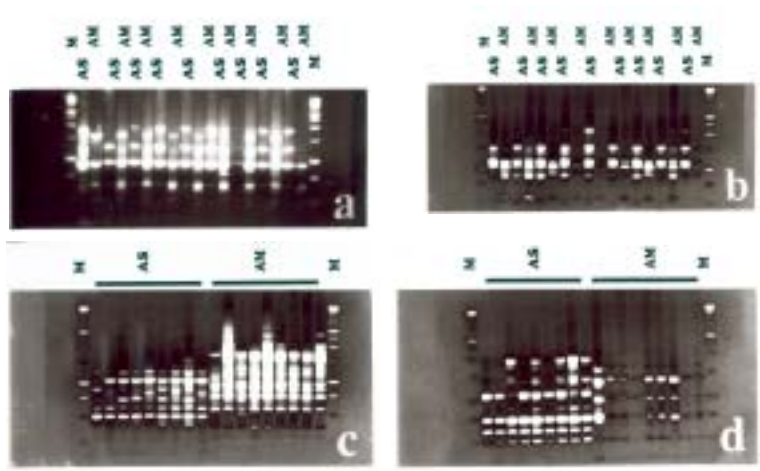

Fig 2. RAPD band profiles of *A. sericea* (AS) and *A. mahidolae* (AM) generated by primers OPAS03, a; and OPAS04, b; OPAS07, c; OPAS19, d. M = molecular size market (1 kb ladder, Promega).

Sequence Analysis of ITS Regions

Aligned ITS1 and ITS2 sequences of the two taxa are shown in Figs 3-4. Sequences from adjacent coding regions including the 5.8S rDNA were also obtained from this study (data not shown). The complete ITS sequences of *A. sericea* and *A. mahidolae* were nearly the same length, 311 to 319 bp in ITS1 and 213 to 224 bp in ITS2, respectively. The length of ITS1 of the two taxa were larger than the species of the subfamily Papilionoideae that were reported by Kaess and Wink.[24] However, the length observed for ITS2 sequences of *A. mahidolae* falls well within the range reported for other species of Papilionoideae (219-224 bp),[24] while the length of ITS2 of *A. sericea* is shorter than the given range. The nucleotide sequences data reported in this study have been deposited with the GenBank Data Library with accession number AF 378104-378107.

Nucleotide composition of the ITS1 was 50.8% G+C, for *A. sericea* and 54.4 % for *A. mahidolae*. The two sequences had 54.3% identity. Meanwhile, little nucleotide variation was found within each taxon. A report on legumes by Kaess and Wink[24] indicated that the ITS regions of legumes showed little nucleotide variation. Aligned ITS1 sequences between the two taxa resulted in 128 variable characters including 120 nucleotide substitutions and 8 insertions and/or deletions (Fig 3). The transition/transversion ratio was 0.63.

The ITS2 sequence showed less variation between the two taxa than ITS1. For nucleotide data of ITS2, the sequence identity between the two taxa was 71.9%. Nucleotide composition was 47% G+C for *A. sericea*, and 50% G+C for *A. mahidolae*. The pairwise alignment of ITS2 region resulted in 49 variable characters (38 nucleotide substitutions and 11 insertions and/or deletions) (Fig 4). The transition/transversion ratio was 1.4. Based on this study, ITS1 sequences are more variable. Generally, the ITS1 sequences were difficult to resolve unambiguously. These regions were highly G+C rich, resulting in compression and multiple bases at each position.[29]

rbcL Gene – Sequence Divergence and Nucleotide Site Variation

The *rbcL* gene of the two taxa was about 1.4 kb in size, but only 834 nucleotides were read for *A. sericea* and *A. mahidolae* (Fig 5). The sequence divergence between the two taxa was 5.6% . Alignment of all 834 bp of *rbcL* sequence resulted in 6 classes of nucleotide substitution and required gaps for 2 nucleotides . Most transitions were T to A (or a reciprocal change). Nucleotide substitution includes 11 transitions (TS) and 36 transversions (TV).Thus, the transition/ transversion ratio was 0.31. Sequence divergence of the two taxa was compared to that of two legumes, *Millettia japonica* and *Tephrosia grandiflora*. The sequence divergence was 7.9% and 11.6%, respectively, for *A. sericea*, and 9.2% and 11.6%, respectively, for *A. mahidolae*. Sequence variation was very little within the nine examined plants of each taxon.

When the nucleotide sequences of the *rbcl* genes of the two taxa were translated to amino acid sequence by inferring from the typical *rbcL* protein of the two legumes and an outgroup species, *Tectona grandis*, the first codon of the two nucleotide sequences was GTT, which encodes valine (V). Consequently, the nucleotide sequence obtained from this study gave a protein containing 278 amino acids (Fig 6). Since nucleotide substitution occurred in the *rbcL* gene of the two taxa, this can lead to either different amino acids or same amino acids (silent mutation). The two *rbcL* proteins of the two taxa shared most amino acid sequence (s) (264 out of 278 or 95%). In addition, amino acid sequence divergence was also observed in comparing the proteins among species used as a reference in this study. The partial sequences of the *rbcL* genes of the two taxa has been deposited in GenBank with accession numbers AF 378793 and AF 378794.

Discussion

We previously reported that the rare Thai plant species *Afgekia sericea* and *A. mahidolae* were closely related species. In our previous studies, geographical, morphological and cytological evidence suggest that *A. sericea* and *A. mahidolae* arose as a result of allelic changes and/or chromosomal rearrangements in allopatric populations of the ancestral species.[3] Hence, the results obtained from this study will have contributed to of our understanding of speciation and evolutionary relationships in these two species.

In the present study, RAPD fragments that are polymorphic between these two species could be used as RAPD markers to distinguish the two species. The markers can be used as genetic markers for species identification. The sizes of ITS1 and ITS2 regions of *A. sericea* and *A. mahidolae* reported here are complete ITS sequences (with combined length of 524 bp for *A. sericea* and 543 bp for *A. mahidolae*). The ITS sequence of the genus *Afgekia* was longer than other species of the Papilionoideae which were reported by Kaess and Wink (ca. 460 bp).[24]

Although the internal transcribed spacers are thought to be important in post-transcriptional processing.[29] In this study, this region has been shown to be very useful for resolving phylogenetic relationships within plant genera. The percentage of sequence divergence in the internal transcribed spacers 1 and 2 (ITS 1 and 2) between the two taxa was very high, while the nucleotide sequences also showed little variation

	1 15	16 30	31 45	46 60	61 75	76 90	91 105
A. sericea	TCGATGCCTTACGAG	CAG-TTCAACCCGTG	AATCTATTTGAGCGT	TGCTCATGACTAGCT	GGAGGTGTTTAG-CA	GCACCTCCACCACCC	TTAGGTTGGGAGGGG
A. mahidolae	TCGATGCCTTACGAG	AAGGTCCAACCCTGG	AACCTGTTTAGGGCC	TGCT-ATGACGCTCT	GAAGGGGACCGAACA	TGACCTTG--CCCTC	TCAGGCGACCTCCGG

	106 120	121 135	136 150	196 210	211 225	226 240	241 255
A. sericea	GA-CTGTTTAGCACA	TTCCCCTATTGG-TC	AAACTCAACAACCCT	CAACATGGAATGTGT	TAAGGAACTAAACTA	AAATTT-ATTTGGGG	CACCCCCGTCGACTT
A. mahidolae	AAACCTTGTCGCGGA	ATGCCTTAATGGGAC	TGAATTAAAATTCGT	CAAGGGCACACGCGA	CTACGAACTAACGGT	AGCTTCTATTCCGCG	GTCCCCGGGTGCGCC

	256 270	271 285	286 300	301 315	316 322
A. sericea	AAAGACAATGTTTGT	GCAGGGAGGCCTTTG	GGGG-TGAAGCGT--	AATGAAT--CTTTGC	AA-CAAA
A. mahidolae	GTAGATAATCTCTTT	TAATAACTGAATTTA	TATGCTTTTGCATTA	AATGAACAACGTCCC	ATCAATA

Fig 3. Aligned ITS1 sequences of *A. sericea* and *A. mahidolae*. Gaps are indicated by a "-".

	1 15	16 30	31 45	46 60	61 75
A. sericea	CTTCTTT--ACTACC	-GTGGCCCCACCTA-	CCTAG--GTCTATCG	TGAGCATCCGCAGAT	GCACATG-GGTTACA
A. mahidolae	CTTCTTTGAACCTCC	GGTGGCCTGGTCTAA	CCT-GAAGTCTAACG	TGAGCATCCGCAGAT	-CACGTCTGGTTATA

	76 90	91 105	106 120	121 135	136 150
A. sericea	AAGCTGGAGTTTAGT	AGGGCAGGCATGTGA	TCAGTCTCGAGCATC	CCTCGTTCAACATGG	CACAC--CCTACTAC
A. mahidolae	GAACCCGAGTTTAGA	AAGGCAG-CATACGA	TCGTTCTCGAGCATC	CGTCGTACACCTTGC	ATCACGACCTACTCC

	151 162	163 177	178 192	193 207	208 222	223 231
A. sericea	A----------GT	CTCACTTTTCAACAG	GCCCATGGAGACAGA	ACAACTCGTGGGGAG	GTCAACATTTCACCC	AGCACGCGA
A. mahidolae	ACTACAGACAGT	CTCACTTTTCAACAA	GCC--CGGAGATGGA	GCAACTCACGGG-AG	GCGAACATT-CACCC	AGCACACGA

Fig 4. Aligned ITS2 sequences of *A. sericea* and *A. mahidolae*. Gaps are indicated by a "-".

	1 15	16 30	31 45	46 60	61 75	76 90
1. *Afgekia sericea*	GTTGGGTTCAAAGAT	GGTGTTAAAGATTAT	AAATTAACTTATTAT	ACTCCTGAGTATGAA	ACCAAAGATACTGAT	ATCTTGGCAGCATTC
2. *Afgekia mahidolae*	GTTGGGTTCAAGGCT	GGTGTGGAAGATTAT	CAATTAACTTATTAT	ACTCCTGAGTCTGAA	ACCAAAGATACTGAT	ATCTTGGCAGCATTC
3. *Millettia japonica*	GTTGGGTTCCAAGCT	GGTGTTAAAGATTAT	AAATTGAATTATTAT	ACTCCTGACTATGAA	ACCAAAGATACTGAT	ATCTTGGCAGCATTC
4. *Tephrosia grandiflora*	GTTGGGTTCAAGGCT	GGTGTTAAAGATTAT	AAATTGACTTATTAT	ACTCCTGACTATGAA	ACCAAAGATACTGAT	ATCTTGGCAGCATTC
5. *Tectona grandis*	GTTGGATTCAAAGCG	GGTGTTAAAGAGTAC	AAATTGACTTATTAT	ACTCCTGAATACGAA	ACCAAAGATACTGAT	ATCTTGGCAGCATTC

	91 105	106 120	121 135	136 150	151 165	166 180
1. *Afgekia sericea*	CGAGTAACTCCTCAA	CCTGGAGTTCCCCCT	GAAGAAGCAGGTGCT	GCGGTAGCAGCCGAA	TCTTCTACTGGGACA	TGGACAACCGTGTGG
2. *Afgekia mahidolae*	CGAGTAACTCCTCAA	CCTGGAGTTCCCCCT	GAAGAAGCAGGTGCT	GCGGTAGCAGCCGAA	TCTTCTACTGGGACA	TGGACAACCGTGTGG
3. *Millettia japonica*	CGAGTATCTCCTCAA	CCTGGAGTTCCGCCT	GAAGAAGCAGGTGCT	GCGGTAGCAGCCGAA	TCTTCCACTGGGACA	TGGACAACTGTGTGG
4. *Tephrosia grandiflora*	CGAGTAACTCCTCAA	CCTGGAGTTCCGCCT	GAAGAAGCAGGTGCT	GCAGTAGCTGCCGAA	TCTTCTACTGGTACA	TGGACAACTGTGTGG
5. *Tectona grandis*	CGAGTAACTCCTCAA	CCTGGAGTTCCGCCT	GAAGAAGCAGGGGCC	GCGGTAGCTGCCGAA	TCTTCTACTGGTACA	TGGACAACTGTGTGG

Fig 5. Aligned *rbcL* gene sequences of the two species of the genus *Afgekia* and other species. Gaps are indicated by a "-".

	181	195	196	210	211	225	226	240	241	255	256	270
1. *Afgekia sericea*	ACCGATGGGCTTACC		AGTCTTGATCGTTAT		AAAGGACGATGCTAC		GGCATCGAGCCCGTT		GCTGGAGAAGAAAAT		CAATATATTGCTTAT	
2. *Afgekia mahidolae*	ACGGATGGGCTTACC		AGTCTTGATCGTTAT		AAAGGACGATGCTAC		GGCATCGAGCCCGTT		GCTGGAGAAGAAAAT		CAATTTATTGCTTAT	
3. *Millettia japonica*	ACCGATGGGCTTACC		AGTCTTGATCGTTAT		AAAGGACGATGCTAC		CACATCGAGCCCGTT		GCTGGAGAAGAAAAT		CAATTTATTGCTTAT	
4. *Tephrosia grandiflora*	ACCGATGGGCTTACC		AGTCTTGATCGTTAC		AAAGGACGATGCTAC		CACATCGAACCCGTT		GCTGGAGAAGAAAGT		CAATTTATTGCTTAT	
5. *Tectona grandis*	ACCGATGGACTTACC		AGCCTTGATCGTTAC		AAAGGGCGATGCTAC		CACATCGAGCCCGTT		CCTGGAGAAAAAGAT		CAATATATCTGTTAT	
	271	285	286	300	301	315	316	330	331	345	346	360
1. *Afgekia sericea*	GTAGCTTATCCCTTA		GACCTTTTTGAAGAA		GGTTCTGTTACTAAC		ATGTTCACCTCCATT		GTAGGTAATGTATTT		GGGTTCAAGGCCTTG	
2. *Afgekia mahidolae*	GTAGCTTATCCCTTA		GACCTTTTTGAAGAA		GGTTCTGTTACTAAC		ATGTTCACCTCCATT		GTAGGTAATGTATTT		GGGTTCAAGGCCTTG	
3. *Millettia japonica*	GTAGCTTATCCCTTA		GACCTTTTTGAAGAA		GGTTCTGTTACTAAC		ATGTTTACCTCCATT		GTAGGTAATGTATTT		GGGTTCAAGGCCTTG	
4. *Tephrosia grandiflora*	GTAGCTTATCCCTTA		GACCTTTTTGAAGAA		GGTTCTGTTACTAAT		ATGTTTACTTCCATT		GTAGGTAATGTATTT		GGGTTCAAAGCCCTG	
5. *Tectona grandis*	GTAGCTTACCCTTTA		GACCTTTTTGAAGAA		GGTTCTGTTACTAAC		ATGTTTACTTCCATT		GTAGGAAATGTATTT		GGATTCAAAGCCCTA	
	361	375	376	390	391	405	406	420	421	435	436	450
1. *Afgekia sericea*	CGCGCTCTACGTCTG		GAGGATTTGAGAATT		CCTGTTTCTTATATT		AAAACTTTCCAGGGT		CCGCCTCACGGAATC		CAATCTGAGAGAGAT	
2. *Afgekia mahidolae*	CGCGCCCTACGTCTG		GAGGATTTGAGAATT		CCCGTTTCTTATATT		AAAACTTTCCAGGGT		CCGCCTCACGGAATC		CAATCTGAGAGAGAT	
3. *Millettia japonica*	CGCGCTCTACGTCTG		GAGGATTTGCGAATC		CCTATTTCTTATGTT		AAAACTTTCCAAGGT		CCGCCTCACGGAATC		CAAGTTGAGAGAGAT	
4. *Tephrosia grandiflora*	CGTGCTCTACGTTTG		GAGGATTTGCGAATC		CCCAATTCTTATATT		AAAACTTTCCAAGGT		CCACCTCATGGTATC		CAAGTTGAAAGAGAT	
5. *Tectona grandis*	CGTGCTCTACGTCTG		GAAGATCTGCGAATC		CCTACTGCTTATATT		AAAACTTTCCAAGGC		CCGCCTCATGGGATC		CAAGTTGAGAGAGAT	
	451	465	466	480	481	495	496	510	511	525	526	540
1. *Afgekia sericea*	AAATTGAATAAGTAT		GGTCGTCCCCTATTG		GGATGTACTATTAAA		CCTAAATTGGGATTA		TCCGCTAAGAATTAC		CGGAGAGCAGTTTAT	
2. *Afgekia mahidolae*	AAATTGAATAAGTAT		GGGCGTCCCCTATTG		GGATGTACTATTAAA		CCAAAATTGGGATTA		TCCGCTAAGAATTAC		GGTAGAGCAGTTTAT	
3. *Millettia japonica*	AAATTGAACAAGTAT		GGCCGTCCCCTATTG		GGATGTACTATTAAA		CCTAAATTGGGGTTA		TCCGCTAAGAATTAC		GGTAGAGCAGTTTAT	
4. *Tephrosia grandiflora*	AAATTGAACAAGTAT		GGACGTCCCTTATTG		GGATGTACTATTAAA		CCTAAATTGGGGTTA		TCCGCTAAGAATTAT		GGTAGAGCGGTTTAT	
5. *Tectona grandis*	AAATTGAACAAGTAC		GGTCGTCCTCTGTTG		GGATGTACTATTAAA		CCGAAATTGGGGTTA		TCTGCTAAAAACTAT		GGTAGAGCAGTTTAT	
	541	555	556	570	571	585	586	600	601	615	616	630
1. *Afgekia sericea*	GAATGTCTCCGCGGA		GGACTTGATTTTACC		AAAGATGATGAAAAT		GTGAACTCCCCACCA		TTTATGCCTTTGAGA		GACCC-GTTCTTATT	
2. *Afgekia mahidolae*	GAATGTCTCCGCGGG		GGACTTGTTTTTTCC		AAAGATGATGAAAAT		GTGAACTCCCCACCA		TTTATGCGTTGGAGA		GACCCGTTTTTTTTT	
3. *Millettia japonica*	GAATGTCTCCGCGGG		GGACTTGATTTTACC		AAAGATGATGAAAAT		GTGAACTCCCAACCA		TTTATGCGTTGGAGA		GACCG-TTTCTGCTT	
4. *Tephrosia grandiflora*	GAATGTCTACGTGGG		GGACTTGATTTTACC		AAAGATGATGAAAAT		GTGAATTCCCAACCA		TTTATGCGATGGAGA		GACCG-TTTCTTATT	
5. *Tectona grandis*	GAATGTCTTCGCGGT		GGACTTGATTTTACC		AAAGATGATGAGAAC		GTGAACTCCCAGCCA		TTTATGCGTTGGAGG		GATCG-CTTCTTATT	
	631	645	646	660	661	675	676	690	691	705	706	720
1. *Afgekia sericea*	TTGTGCCCGAGGCAT		TTATTAAGCACAGGC		CGAAACAGGTGAAAT		CAAAGGGCTTATTTT		GAATGCCTCTGCGGG		GCCCTGCCAAGAAAT	
2. *Afgekia mahidolae*	TTGTGCCGAAGGGAA		TTTTAAAGCACAGGC		-GAAACAGGGGAAAT		AAAAGGGCCTTTTTT		GAATGCCTCTGCGGG		TCCCTGCGAAGAAAT	
3. *Millettia japonica*	TTGTGCCGAAGCACT		TTATAAAGCACAGGC		CGAAACAGGTGAAAT		CAAAGGGCATTATTT		GAATGCTACTGCGGG		TACATGCGAAGAAAT	
4. *Tephrosia grandiflora*	TTGTGCAGAAGCTAT		TTTTAAATCACAGGC		TGAAACGGGTGAAAT		CAAAGGGCATTATTT		AAACGCTACTGCAGG		TACATGCGAAGAAAT	
5. *Tectona grandis*	TTGTGCCGAAGCAAT		TTATAAAGCACAGGC		TGAAACAGGTGAAAT		CAAAGGGCATTACTT		GAATGCTACTGCGGG		TACATGCGAAGAAAT	

Fig 5. (continued) Aligned *rbcL* gene sequences of the two species of the genus *Afgekia* and other species. Gaps are indicated by a "-".

	721 735	736 750	751 765	766 780	781 795	796 810
1. *Afgekia sericea*	GATGAAAAGAGCTTT	TTTTTCCCAAAAAAT	AGGCGTTCCTTTGGT	AATGCATGACTCCTT	AACAGGGGGGTCCCC	GGCAAAAACTAACTT
2. *Afgekia mahidolae*	GATGAAAAGAGCTTT	TTTTTCCCCGAAATT	TGGCGTTCCTTTTTT	TATGCCTGACTCCTT	AACAGGGGGGTTCAC	TGCAAAAAACAACTT
3. *Millettia japonica*	GCTAAAAAGAGCTGT	ATTTGCCAGAGAATT	AGGAGTTCCTATCGT	AATGCATGACTACTT	AACAGGGGGGTTCAC	TGCAAATACTACCTT
4. *Tephrosia grandiflora*	GATAAAAAGAGCTGT	ATTTGCCAGAGAATT	GGGCGTTCCTATCGT	AATGCATGATTACTT	AACAGGGGGATTCAC	TGCAAATACTAGCTT
5. *Tectona grandis*	GATGAAAAGGGCTGT	ATTTGCTAGAGAATT	GGGAGTTCCTATCGT	AATGCATGACTACTT	AACAGGAGGATTCAC	TGCAAATACTAGCTT

	811 825	826 840
1. *Afgekia sericea*	AGCTCACTATTGCCG	GGATAATGGG
2. *Afgekia mahidolae*	ACCCCCCTATTGCCG	GGATAATGGG
3. *Millettia japonica*	GGCTCACTATTGCCG	GGATAATGGT
4. *Tephrosia grandiflora*	AGCTCATTATTGCCA	GAATAATGGT
5. *Tectona grandis*	GGCTCATTATTGCCG	AGATAATGGC

Fig 5. (continued) Aligned *rbcL* gene sequences of the two species of the genus *Afgekia* and other species. Gaps are indicated by a "-".

	1 15	16 30	31 45	46 60	61 75	76 90
1. *Afgekia sericea*	VGFKAGVKDYKLTYY	TPDYETKDTDILAAF	RVTPQPGVPPEEAGA	AVAAESSTGTWTTVW	TDGLTSLDRYKGRCY	GIEPVAGEENQYIAY
2. *Afgekia mahidolae*	VGFKAGVKDYKLTYY	TPDSETKDTDILAAF	RVTPQPGVPPEEAGA	AVAAESSTGTWTTVW	TDGLTSLDRYKGRCY	GIEPVAGEENQFIAY
3. *Millettia japonica*	VGFQAGVKDYKLNYY	TPDYETKDTDILAAF	RVSPQPGVPPEEAGA	AVAAESSTGTWTTVW	TDGLTSLDRYKGRCY	HIEPVAGEENQFIAY
4. *Tephrosia grandiflora*	VGFKAGVKDYKLTYY	TPDYETKDTDILAAF	RVTPQPGVPPEEAGA	AVAAESSTGTWTTVW	TDGLTSLDRYKGRCY	HIEPVAGEESQFIAY
5. *Tectona grandis*	VGFKAGVKEYKLTYY	TPEYETKDTDILAAF	RVTPQPGVPPEEAGA	AVAAESSTGTWTTVW	TDGLTSLDRYKGRCY	HIEPVPGEKDQYICY

	91 105	106 120	121 135	136 150	151 165	166 180
1. *Afgekia sericea*	VAYPLDLFEEGSVTN	MFTSIVGNVFGFKAL	RALRLEDLRIPVSYI	KTFQGPPHGIQSERD	KLNKYGRPLLGCTIK	PKLGLSAKNYRRAVY
2. *Afgekia mahidolae*	VAYPLDLFEEGSVTN	MFTSIVGNVFGFKAL	RALRLEDLRIPVSYI	KTFQGPPHGIQSERD	KLNKYGRPLLGCTIK	PKLGLSAKNYGRAVY
3. *Millettia japonica*	VAYPLDLFEEGSVTN	MFTSIVGNVFGFKAL	RALRLEDLRIPISYV	KTFQGPPHGIQVERD	KLNKYGRPLLGCTIK	PKLGLSAKNYGRAVY
4. *Tephrosia grandiflora*	VAYPLDLFEEGSVTN	MFTSIVGNVFGFKAL	RALRLEDLRIPNSYI	KTFQGPPHGIQVERD	KLNKYGRPLLGCTIK	PKLGLSAKNYGRAVY
5. *Tectona grandis*	VAYPLDLFEEGSVTN	MFTSIVGNVFGFKAL	RALRLEDLRIPTAYI	KTFQGPPHGIQVERD	KLNKYGRPLLGCTIK	PKLGLSAKNYGRAVY

	181 195	196 210	211 225	226 240	241 255	256 270	271 278
1. *Afgekia sericea*	ECLRGGLDFTKDDEN	VNSQPFMRLRDRFCF	CAEAIYKAQAETGEI	KGHILNASAGPTQEM	MKRAFFSQKIGVPLV	MHDSLTGGFPAKTNL	AHYCRDNG
2. *Afgekia mahidolae*	ECLRGGLDFTKDDEN	VNSQPFMRWRDRFCF	CAEANFKAQAETGEI	KGHFLNASAGPTEEM	MKRAFFSQKIGVPFF	MPDSLTGGFTAKNNL	PPYCRDNG
3. *Millettia japonica*	ECLRGGLDFTKDDEN	VNSQPFMRWRDRFCF	CAEALYKAQAETGEI	KGHYLNATAGTCEEM	LKRAVFARELGVPIV	MHDYLTGGFTANTTL	AHYCRDNG
4. *Tephrosia grandiflora*	ECLRGGLDFTKDDEN	VNSQPFMRWRDRFLF	CAEAIFKSQAETGEI	KGHYLNATAGTCEEM	IKRAVFARELGVPIV	MHDYLTGGFTANTSL	AHYCQNNG
5. *Tectona grandis*	ECLRGGLDFTKDDEN	VNSQPFMRWRDRFLF	CAEAIYKAQAETGEI	KGHYLNATAGTCEEM	MKRAVFARELGVPIV	MHDYLTGGFTANTSL	AHYCRDNG

Fig 6. Comparison of deduced amino acid sequences of *rbcL* proteins of the two species of the genus *Afgekia* and three other species.

among examined plants of each taxon. This suggested that *A. sericea* and *A. mahidolae* have not recently diverged or they have mutated very rapidly. Sequence variation of these spacers between the two taxa occurred not only by base substitution but also by a few insertion and deletion events. The base substitution appeared to be congruent based on the aligned data. Since nucleotide sequence divergence was high, the ITS sequences divergence of the two taxa may reflect the evolutionary forces (such as mutation and natural selection) that generated divergence in the process of plant speciation from their common ancestor.

The entire *rbcL* sequence of plant in this subfamily is 1420 bp in length. The length of nucleotide sequences of *rbcL* gene reported here were 834 bp, with a divergence of 5.6% (all changes) between the two species. In addition, amino acid sequence divergence is about 5.1%. Kaess and Wink[24] reported that all plants of Papilionoideae studied have no changes (deletions and/or insertions, or inversions of nucleotides) on *rbcL* sequence. The results from this study contradict that previous study. However, divergence of *rbcL* sequence in chloroplast DNA has been reported within plants in the genus *Fagopyrum*.[23]

The additional molecular data obtained from this study will provide a foundation for future studies on genome evolution within the genus *Afgekia*.

Acknowledgements

The authors would like to thank two anonymous readers for the valuable comments on the draft manuscript. This work was partly supported by the TRF/BIOTEC Special Program for Biodiversity Research and Training grant BRT 142008.

References

1. Burtt BL and Chermsirivathana C (1971) A second species of *Afgekia* (Leguminosae) BL Burtt and C Chermsirivathana. *Reprint from the Royal Botanic Garden Edinburgh* **30**, 131-3.
2. Geesink R (1984) *Scala millettiearum* : A survey of the genera of *Millettieae* (Leguminosae-Papilionoideae) with methodological considerations. EJ Brill Leiden University Press, Leiden.
3. Prathepha P and Baimai V (1999a) Cytological and morphological characteristics of two closely related species of the genus *Afgekia* Craib (Leguminosae) from Thailand. *Mahidol Journal* **6,** 23-6.
4. Rabinowitz D (1981) Seven forms of rarity, In: *The biological aspects of rare plant conservation* (Edited by Sunge H), pp 205-17. John Wiley & Sons, Chichester, UK.
5. Prathepha P and Baimai V (1999b) Genetic differentiation in Thai populations of the rare species *Afgekia serica* Craib (Leguminosae) revealed by RAPD-PCR assays. *Genetica* **105**, 193-202.
6. Prathepha P (1994) Chromosome number of the genus *Afgekia* Craib (Leguminosae) *Cytologia* **59**, 437-8.
7. Clegg MT (1993) Chloroplast gene sequences and the study of plant evolution. *Proc Natl Acad Sci USA* **90**, 363-7.
8. Haymer DS (1994) Random amplified polymorphic DNAs and microsatellites: What are they, and can they tell us anything we don't already know? *Ann Entomol Soc Am* **87**, 717-22.
9. Williams JGK Kubelik AR Livak KJ Rafalski JA and Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers *Nucleic Acids Res* **18**, 6531-5.
10. Welsh J and McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res* **18**, 7213-8.
11. Wilkie SE Isaak PG and Slater RJ (1993) Random amplified polymorphic DNA (RAPD) markers for genetic analysis in Allium. *Theor App Genet* **56**, 497-504.
12. Campos LP Raelson JV and Grant WF (1994) Genome relationship among Lotus species based on random amplified polymorphic DNA (RAPD) *Theor Appl Genet* **88**, 417-22.
13. Gehrig HH Rosicke H and Kluge M (1997) Detection of DNA polymorphisms in the genus Kalanchoe by RAPD PCR fingerprint and its relationships to infrageneric taxonomic position and ecophysiological photosynthetic behaviour of the species. *Plant Science* **125**, 41-51.
14. Noli E Salvi S and Tuberosa R (1997) Comparative analysis of genetic relationships in barley based on RFLP and RAPD markers. *Genome* **40**, 607-16.
15. Charmet G Ravel C and Balfourier F (1997) Phylogenetic analysis in the Festuca-Lolium complex using molecular markers and ITS rDNA. *Theor Appl Genet* **94**, 1038-46.
16. Baldwin BG (1992) Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants : an example from the compositae. *Mol Phylo Evol* **1**, 3-16.
17. Baldwin BG Sanderson MJ Porter JM Wojciechowski MF Campbell CS and Donoghue MJ (1995) The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. *Ann Miss Bot Gard* **82**, 247-77.
18. Bayer RJ Soltis DE and Soltis PS (1996) Phylogenetic inferences in Antennaria (Asteraceae : Gnaphalieae : Cassiniinae) based on sequences from nuclear ribosomal DNA internal transcribed spacers (ITS). *Am J Bot* **83**, 516-27
19. Appels R and Honeycutt RL (1986) rDNA: evolution over a billion years. In: *DNA systematics, vol. II Plants* (Edited by S.K. Dutta SK), pp. 81-135. CRC Press, Boca Raton, Florida.
20. Palmer JD Jansen RK Michaels HJ Chase MW and Manhart, JR (1988) Chloroplast DNA variation and plant phylogeny. *Ann Mo Bot Gard* **75**, 1180-206.
21. Clegg MT Learn GH and Golenberg EM (1991) Molecular evolution of chloroplast DNA. In: *Evolution at the molecular level* (Edited by Selander RK, Clark AG and Whittam TS), pp. 135-49. Sinauer, Sunderland, MA.
22. Palmer JD (1987) Chloroplast DNA evolution and biosystematic uses of chloroplast DNA variation. *Am Nat* **130**, s6-s26.
23. Yasui Y and Ohnishi O (1996) Comparative study of rbcl gene in Fagopyrum and related taxa. *Genes Genet Syst* **71**, 219-24.
24. Kaess E and Wink M (1997) Phylogenetic relationships in the *Papilionoideae* (Family Leguminosae) based on nucleotide sequences of cpDNA (rbcL) and ncDNA (ITS1 and 2). *Mol Phylogen Evol* **8**, 65-88.
25. Doyle JJ and Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bull* **19**: 11-5.
26. Nei M and Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA* **76**, 5269-73.
27. Mummenhoff K Pranzke A and Koch M (1997) Molecular phylogenetics of Thlaspi S.L. (Brassicaceae) based on chloroplast DNA restriction site variation and sequences of the internal transcribed spacers of nuclear ribosomal DNA. *Can J Bot* **75**, 469-82.
28. Smith R F Wiese BA Wojzynski MK Davison DB and Worley KC (1996) BCM Search Launcher: An Integrated Interface to Molecular Biology Data Base Search and Analysis Services Available on the World Wide Web. *Genome Res* **6**, 454-62.
29. Downie SR Ramanath S Katz-Downie DS and Llanas E (1998) Molecular systematics of Apiaceae subfamily Apioideae: phylogenetic analysis of nuclear ribosomal DNA internal transcribed spacer and plastid rpoC1 intron sequences. *Am J Bot* **85**, 563-91.
30. Moller M and Cronk QCB (1997) Origin and relationships of *Saintpaulia*(Gesneriaceae) based on ribosomal DNA internal transcribed spacer (ITS) sequences. *Am J Bot* **84**, 956-65.