Journal of Medicinal Plants Research Vol. 5(28), pp. 6413-6417, 30 November, 2011
Available online at http://www.academicjournals.org/JMPR
ISSN 1996-0875 ©2011 Academic Journals
DOI: 10.5897/JMPR11.545

*Full Length Research Paper*

# Identification of the genus *Epimedium* with DNA barcodes

**Yu Jiang[1], Chun-bang Ding[1]*, Li Zhang[1], Ruiwu Yang[1], Yonghong Zhou[2] and Li Tang[1]**

[1]College of Biology and Science, Sichuan Agricultural University, Yaan 625014, China.
[2]Key Laboratory of Crop Genetic Resources and Improvement Ministry of Education, Sichuan Agricultural University, Yaan 625014, China.

Accepted 24 October, 2011

**The traditional identification for *Epimedium* is inadequate due to the morphological variation and complicated procedure. DNA barcoding is a new technique for species identification by a standard DNA region. However, there has been no consensus emerged for plant barcoding due to the low rates of evolution in plastid DNA. In this paper, we compared the performance of four candidate DNA regions (*psb*A-*trn*H, ITS, *rbc*L, *mat*K) in ten species of *Epimedium.* To select a suitable barcode, ranking criteria included: the success efficiency of polymerase chain reaction (PCR) amplification and sequencing; genetic divergences; the barcoding gap. The *psb*A-*trn*H represented easy amplification, high level of interspecific divergence and well-separated barcoding gap. The other three loci (ITS, *rbc*L, *mat*K) was inadequate for application. We concluded that the *psb*A-*trn*H region was the better potentially DNA barcode for the genus *Epimedium*.**

**Key words:** *Epimedium*, DNA barcode, ITS, *rbc*L, *mat*K, *psb*A-*trn*H.

## INTRODUCTION

The genus *Epimedium* (Beberaceae) comprises 54 species and 43 in central-southeastern China (Stearn, 2002; Sun et al., 2005). Certain species of *Epimedium* have been used as a tonic in China for over 2000 years. The official species listed in the Chinese Pharmacopoeia (2010) are *Epimedium brevicornum* Maxim.*, Epimedium pubescens* Maxim.*, Epimedium koreanum* Nakai and *Epimedium sagittatum* Maxim. Owing to large-scale exploitation, it is essential to protect genetic diversity of *Epimedium*. An easy and accurate method is required for identifying *Epimedium* species. Recently, Corolla characteristics, Biochemical markers and molecular marker, such as RAPD, have been used as classification tools for *Epimedium* species (Nakai et al., 1996; Sun et al., 2005). However, the tranditional identification is inadequate due to the morphological variation and complicated procedure. A new molecular identification-DNA barcoding- has been proposed by workers (Hebert et al., 2003).

used for species identification, biodiversity assessment DNA barcoding is a novel classification that can be and ecological studies by a short gene sequence fragment (Hebert et al., 2003; Barraclough and Savolainen, 2008). DNA barcode have been used to make accurate identification in cases where specimens are insufficient or morphological characteristics are absent (Kress et al., 2005). At present, the mitochondrial cytochrome oxidase subunit I (coxI) is chosed as the standard barcode for animal identification with almost 100% accuracy (Chase et al., 2007). For plants, coxI has not been successfully used for specie identification due to slow evolutionary rate. An appropriate DNA region is necessary for plant. The vast numbers of genetic loci have been proposed for plant DNA barcoding, but no consensus has emerged (Gonzalez et al., 2009). Several regions (*mat*K, *rpo*C1, *rpo*B, *acc*D, trnL$^{UAA}$, *ndh*J, *psb*A-*trn*H, *trn*L, *atp*F-*atp*H, ITS, *rbc*L) have been suggested as potential barcodes (Kress and Erickson, 2007; CBOL, 2009; Taberlet et al., 2007; Chase et al., 2007).

Here, we selected four DNA regions, two plastid regions (*mat*K, *rbc*L), one non-coding regions (*psb*A-*trn*H) and one nuclear region ITS, to explore the best candidate barcode for the genus *Epimedium*.

*Corresponding author. E-mail: dcb@sicau.edu.cn. Tel: 86-835-2886170. Fax: 86-835-2886170.

**Table 1.** PCR primer and reaction conditions used in this study.

| Locus | Primer | Sequence 5'-3' | Reaction condition | Cycles |
|---|---|---|---|---|
| ITS | ITS-4 | TCCTCCGCTTATTGATATGC | 97˚C/1 min | 35 |
| | ITS-5 | GGA AGTAAA AGT CGT AAC AAG G | 97˚C /45 s | |
| | | | 52˚C /45 s | |
| | | | 72/2 min | |
| | | | 72/10 min | |
| *mat*K | 390 | CGATCTATTCATTCAATATTTC | 97˚C /1 min | 35 |
| | 1326 | TCTAGCACACGAAAGTCGAAGT | 97˚C /45 s | |
| | | | 50˚C /45 s | |
| | | | 72/2 min | |
| | | | 72/10 min | |
| *rbc*L | 1 | ATGTCACCACAAACAGAAAC | 97˚C /1 min | 35 |
| | 724 | TCGCATGTACCTGCAGTAGC | 97˚C /45 s | |
| | | | 55˚C /45 s | |
| | | | 72/2 min | |
| | | | 72/2 min | |
| *psb*A-*trn*H | psbA3'f | GTTATGCATGAACGTAATGCTC | 97˚C /1 min | 35 |
| | trnH | CGCGCATGGTGGATTCACAATCC | 97˚C /45 s | |
| | | | 55˚C /45 s | |
| | | | 72/2 min | |
| | | | 72/2 min | |

## MATERIALS AND METHODS

37 specimens belonging to 10 species of *Epimedium* collected from different part of Sichuan Province and three of them (*E. brevicornum* Maxim., *E. pubescens* Maxim., *E. sagittatum* Maxim) are officially recorded in Chinese pharmacopoeia (2010). The voucher specimens were preserved in the Herbarium of Sichuan Agricultural University (SAU).

Total DNA was extracted from silica gel-dried leaves following the CTAB method of Doyle and Doyle (1987). In this study, we collected four regions (*mat*K, *rbc*L, *psb*A-*trn*H, ITS) that have been proposed as potential DNA barcoding by the Consortium of Barcode of Life (CBOL Plant Working Group, 2009).

Polymerase chain reaction (PCR) was performed for the selected loci following the protocols of Kress et al. (2005). The primer pairs used in the present study are: *mat*K-390F, *mat*K-1326R (Cuénoud et al., 2002); *rbc*L-1F, *rbc*L-724R (Kress et al., 2005); *psb*A3'f, *trn*H (Kress and Erickson, 2007); ITS-5, ITS-4 (Sun et al., 2005). Primers and PCR reaction condition are showed in Table 1. The PCR products were sequenced with the amplification primers.

DNA sequences were aligned using DNAman software and then manually modified. The inter- and intraspecific variation of the samples was calculated following the method of Meyer and Paulay (2005). Two parameters were applied to evaluate the interspecific divergence (Meyer and Paulay, 2005): (i) all interspecific distances between species that had at least two representatives (ii) the minimum interspecific distance, that is, the average of smallest interspecific distance between species with at least two species. Three parameters were used to compare intraspecific divergence (Barraclough and Savolainen, 2008; Meyer and Paulay, 2005): (i) mean all intraspecific distance within species with at least two individuals; (ii) "mean theta", Where theta was the average pairwise distance calculated within species to eliminate biases associated with difference sampling among species; (iii) average coalescent depth, that is, the maximum distance within each species. The comparation of inter/intra-specific variation for each region was performed by Wilcoxon signed rank tests with the software MEGA 4.0 (Tamura et al., 2007). The gaps were evaluated by comparing the distribution of inter/intra-specific divergence.

## RESULTS

The data indicated that *mat*K exhibited 100% PCR success with two primers and the other regions could not be amplified in one or more samples (Figure 1). The PCR success for four barcodes were 100% (*mat*K), 90% (*psb*A-*trn*H), 83% (*rbc*L), 80% (ITS). Among all samples, success rates for sequenced were highest for ITS (95.8%) and *psb*A- *trn*H (88.9%), followed by *rbc*L (84%) and *mat*K (76.7%).

Over the four tested DNA regions, *psb*A-*trn*H had showed considerable variation (35 variable characters of 533 total base), ITS (18 of 687), *mat*K (11 of 799), *rbc*L (4 of 706) was the least. For regions, *mat*K provided the longest amplification fragment and *psb*A-*trn*H was the shortest. The other two sequences had no significant length differences between species.

Wilcoxon signed rank tests on this four candidate DNA barcodes showed that *psb*A-*trn*H exhibited higher inter-

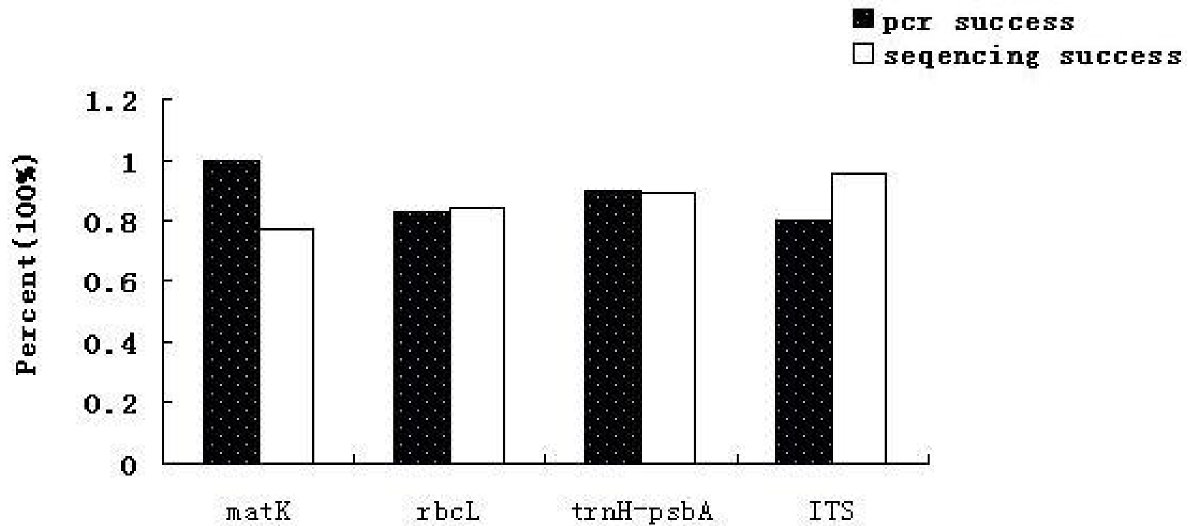


**Figure 1.** Efficiency of amplification and sequencing for four loci (*mat*K, *rbc*L, *psb*A-*trn*H, ITS) tested.

**Table 2.** Wilcoxon signed rank test for inter-specific divergence among loci.

| W- | W+ | Relative ranks, n, P value | Result |
|---|---|---|---|
| *mat*K | *rbc*L | W+=15, W-=0, n=5, p<=0.001 | *mat*K >*rbc*L |
| *mat*K | *psb*A-*trn*H | W+=0, W-=15, n=5, p<=0.001 | *psb*A-*trn*H > *mat*K |
| *mat*K | ITS | W+=15, W-=0, n=5, p<=0.001 | *mat*K > ITS |
| *psb*A-*trn*H | *rbc*L | W+=0, W-=21, n=6, p<=0 | *psb*A-*trn*H > *rbc*L |
| ITS | *rbc*L | W+=20, W-=1, n=6, p<=0 | ITS > *rbc*L |
| ITS | *psb*A-*trn*H | W+=45, W-=0, n=9, p<=0 | *psb*A-*trn*H > ITS |

specific divergence than other three loci, followed by *mat*K (Table 2). However, *psb*A-*trn*H indicated that average coalescent depth was a litter larger than the minimum interspecific distance (Table 3). Intraspecific variability was crucial for the assessment of a suitable barcode. As shown in Table 3, all four DNA regions exhibited low variation at intraspecific level and the lowest was *rbc*L.

An ideal DNA barcoding must exhibit a “gap” between inter-/ intraspecific divergence. To assess whether such a "gap" occurred, we evaluated the distribution of inter/intra specific divergence in classes of 0.002 distance units (Figure 2). However, the result did not show any large gap as coxI in animal, although with *psb*A-*trn*H the distribution inter/intra specific divergence was relatively well separated. The data revealed obvious overlapping when the divergence distributions were calculated using *mat*K, *rbc*L and ITS.

## DISCUSSION

A favorable DNA barcoding must contain short sequence length to faciliate amplification and sequence, enough genetic divergence for identification and conserved regions for developing universal primers (Kress et al., 2005; Sass et al., 2007). The most promising single loci for plant DNA barcode is the non-coding *psb*A-*trn*H which has been accepted as a candidate barcode in recent years (Kress and Erickson, 2007). Recently, authors demonstrated that *psb*A-*trn*H had a high level of amplification success in angiosperms, gymnosperms, ferns and liverworts (Kress et al., 2005). The *psb*A-*trn*H marker performed well as a DNA barcode for *Epimedium* in this study. The results show that *psb*A-*trn*H has been universally amplifiable with an average length of 520 bp in *Epimedium* and the inter-specific divergence is higher than the other three loci. In gap assessment test, *psb*A-*trn*H exhibits distinct gaps between intra- and interspecific divergence. Because of its complex molecular evolution (Barraclough and Savolainen, 2008), the locus *psb*A-*trn*H needs further research to analyse whether the non-coding region could continue do the best performance in the expanded sampling.

ITS which was proposed by Kress et al. (2005) as a potential barcode for plants, gives limited amplification

**Table 3.** Analysis of intra- and interspecific divergence for four candidate barcodes.

| | ***mat*K** | ***rbc*L** | ***psb*A-*trn*H** | **ITS** |
|---|---|---|---|---|
| All interspecific distance | 0.0033±0.0017 | 0.0022±0.0011 | 0.0129±0.0056 | 0.0032±0.0023 |
| All intraspecific distances | 0.0011±0.0009 | 0.0003±0.0003 | 0.0085±0.0019 | 0.0014±0.0014 |
| The minimum interspecific distance | 0.0027±0.0019 | 0.0016±0.0010 | 0.0078±0.0031 | 0.0023±0.0019 |
| Theta | 0.0010±0.0009 | 0.0003±0.0004 | 0.0079±0.0019 | 0.0017±0.0017 |
| Coalecent depth | 0.0021±0.0019 | 0.0006±0.0007 | 0.0112±0.0022 | 0.0023±0.0021 |

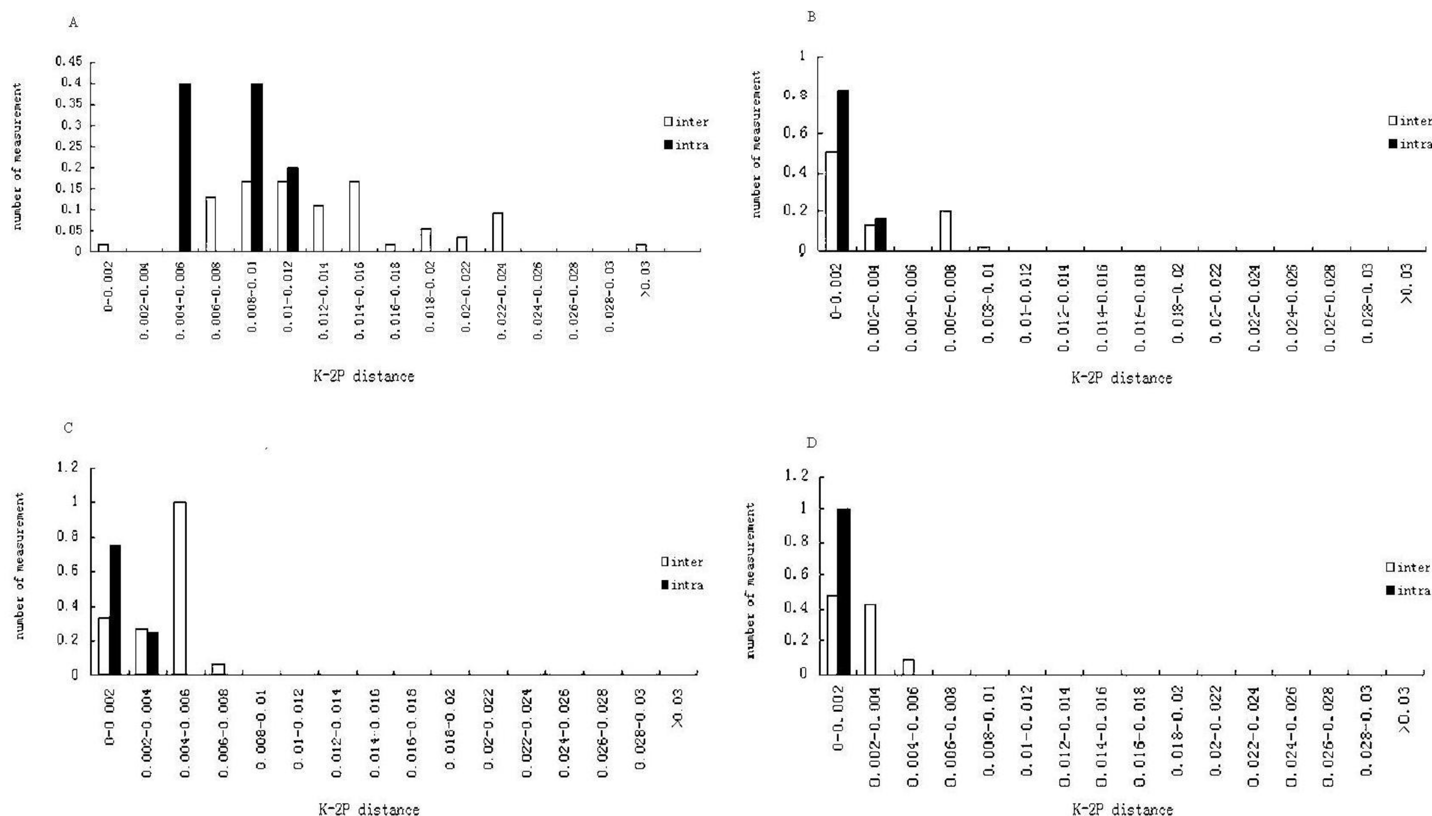


**Figure 2.** Relative distribution between intra- and interspecific divergence; Barcoding gaps were assessed in classes of 0.002 distance units. (A) *psb*A-*trn*H. (B) ITS. (C) *mat*K. (D) *rbc*L.

success observed in this study. The results indicate that ITS is not a good barcode for the genus *Epimedium*. Yet, the nuclear ITS region could be used as identification tool for particular species such as the *Sapotaceae* (Gonzalez et al., 2009).

*mat*K and *rbc*L have been recommended by CBOL as a candidate region in their lastest study (CBOL, 2009). Kress and Erickson (2007) indicated that the primary problem of *mat*K for its use was the poor amplification. Of our samples, *mat*K is amplified by the primer 390F and 1326R from Cuénoud et al. (2002) with a 100% success. This region provides variable value in selected groups of plants but a low sequencing success using two primer pairs. Fazekas et al. (2008) indicated that a sequenceing success of 88% for this region with 10 primer pairs. Ford et al. (2009) reported that a combination of standard and nested multiplexed-tandem PCR (MT-PCR) resulted in a 85% success for *mat*K. Such sophisicated procedures are not suitable for DNA barcoding. In our study, the *mat*K shows poor performance because of low sequencing success. We agreed with Kondo et al. (2007) that *mat*K is not appropriate for species identification. The *rbc*L marker proved to be easily in amplification and sequence with universal primers (Newmaster et al., 2006). A number of *rbc*L data are deposited in GenBank. *rbc*L had the worst performance with insufficient variable, low interspecific variation and entirely no gap. We suggested that this locus is not suitable for the genus *Epimedium*.

Based on these results, we demonstrated that *psb*A-

*trn*H can be treated as a suitable plant DNA barcode for the *Epimedium* genus. In summary, DNA barcodes can be very effective in species identification even in ecological application (Gonzalez et al., 2009). This new molecular technique would benefit plant diversity surveys especially closely related species and species-rich genera lacking morphological characters. The result of this effort makes a contribution for application of DNA barcodes in plant.

## REFERENCES

Barraclough TG, Savolainen V (2008). DNA barcoding the floras of biodiversity hotspots. Proc. Natl. Acad. Sci. U.S.A., 105: 2923-2928.

CBOL Plant Working Group (2009). DNA barcode for land plant. Proc. Natl. Acad. Sci. U.S.A., 106: 12794-12797.

Chase MW, Cowan RS, Hollingsworth PM (2007). A proposal for a standardised protocol to barcode all land plants. TAXON, 56: 295-299.

Cuénoud P, Savolainen V, Chatrou LW, Powell M, Grayer RJ, Chase M (2002). Molecular Phylegenetic of Caryophyllales based on Nulear 18S r DNA and Plastid *rbcL*,*atpB*, and *matK* DNA Sequences. Am. J. Bot., 89: 132-144.

Gonzalez MA, Baraloto CB, Engel J, Mori SA, Pétronelli, Riéra B, Roger A, Thébaud C, Chavel J (2009). Identification of Amazonian Trees with DNA Barcodes. PLOS ONE, 4: e7483.

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

Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, Percy DM, Hajibabaei M, Barrett SCH (2008). Multiple Multilocus DNA Barcodes from the Plastid Genome Discriminate Plant Species Equally Well. PLOS ONE, 3: e2802.

Ford CS, Ayres KL, Toomey N, Haider N, van Alphen SJ (2009). Selection of candidate coding DNA barcoding regions for use on land plants. Bot. J. Linn. Soc., 159: 1–11.

Hebert PD, Cywinska A, Ball SL (2003). Biological identifications through DNA barcodes. Proc. R. Soc. Biol. Sci. Ser. B, 270: 313–321.

Kondo K, Shiba M, Yamaji H, Morota T, Zhengmin C, Huixia P, Shoyama Y (2007). Species Identification of Licorice Using nrDNA and cpDNA Genetic Markers. Biol. Pharm. Bull., 30: 1497–1502.

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005). Use of DNA barcodes to identify flowering plants. Proc. Natl. Acad. Sci. U.S.A., 102: 8369-8374.

Kress WJ, Erickson DL (2007). A Two-Locus Global DNA Barcode for Land Plants: The Coding *rbcL* Gene Complements the Non-Coding *trnH-psbA* Spacer Region. PLOS ONE, 6: e508.

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

Nakai R, Shoyama Y, Shiraishi S (1996). Genetic Characterization of Epimedium Species Using Random Amplified Polymorphic DNA (RAPD) and PCR-Restriction Fragment Length Polymorphism (RFLP) Diagnosis. Biol. Pharm. Bull., 19: 67-70.

Newmaster SG, Fazekas AJ, Ragupathy S (2006). DNA barcoding in lant plant: evalution of *rbcL* in a multigene tiered approach. Can. J. Bot., 84: 335-341.

Pharmacopoeia Committee of China (2010). Pharmacopoeia of People's Republic of China. Beijing: Chemical Industry Press.

Sass C, Little DP, Stevenson DW, Specht CD (2007). DNA Barcoding in the Cycadales: Testing the Potential of Proposed Barcoding Markers for Species Identification of Cycads. PLOS ONE, 11: e1154.

Stearn WT (2002). The Genus *Epimedium* and Other Herbaceous Berberidaceae.Timber Press, Inc., Portland, OR, USA.

Sun Y, Fung KP, Leung PC, Shaw PC (2005). A phylogenetic analysis of *Epimedium* (Berberidaceae) based on nuclear ribosomal DNA sequences. Mol. Phylogenetics Evol., 35: 287-291.

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G, Christian B, Willerslev E (2007). Christian Brochmann and Eske Willerslev. Power and limitations of the chloroplast *trn*L (UAA) intron for plant DNA barcoding. Nucleic Acids Res., 35: e14.

Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. Mol. Biol. Evol., 24: 1596–1599.