Identification of Two Chemotypes of *Pogostemon cablin* (Blanco) Benth. through DNA Barcodes

Yang He, Feng Wan, Liang Xiong, Dong-Mei Li, and Cheng Peng*

State Key Laboratory Breeding Base of Systematic Research, Development and Utilization of Chinese Medicine Resources, Chengdu University of Traditional Chinese Medicine, No. 1166 Liutai Avenue, Chengdu 610075, P. R. China. Fax: +86 (28) 61801919.
E-mail: tcmpengcheng@gmail.com

* Author for correspondence and reprint requests

Z. Naturforsch. **69c**, 253 – 258 (2014) / DOI: 10.5560/ZNC.2013-0180
Received November 17, 2013 / February 19, 2014 / published online June 5, 2014

Pogostemon cablin (Blanco) Benth. is an important medicinal plant in Traditional Chinese Medicine (TCM). Because of differences in the chemical composition, this species has been classified into two major chemotypes, *i. e.* the patchouliol-type and the pogostone-type; however, no quick and effective method is presently available for the precise identification of these two chemotypes. DNA barcoding, using a standardized DNA fragment, is a promising molecular diagnostic method for species identification. We have established a reliable and quick method for the identification of the *P. cablin* chemotypes. Of five potential barcodes [*rbcL*, *psbA-trnH*, *rpoB*, ITS (internal transcibed spacer), and *ndhJ*], tested among 103 samples, ITS was the best candidate, as comparative studies between patchouliol-type and pogostone-type *P. cablin* revealed that ITS had more variable regions among these five barcodes. We suggest that ITS can serve as the most suitable barcode for differentiating between the chemotypes of *P. cablin*.

Key words: Pogostemon cablin, Patchouliol-Type, Pogostone-Type, ITS Barcoding

Introduction

Pogostemon cablin (Blanco) Benth. (Lamiaceae) is an important traditional Chinese medicinal plant (Guan *et al.*, 1992; Paul *et al.*, 2012), also called Guanghuoxiang in Chinese (Chen *et al.*, 2013). Presently, *P. cablin* is widely spread in southern China, including Guangdong, Hainan, and Guangxi (Hu *et al.*, 2006). In Traditional Chinese Medicine (TCM), this plant is used for removing dampness, relieving summer-heat or exterior syndrome, stopping vomiting, and stimulating appetite (Xu *et al.*, 2010; Chinese Pharmacopoeia Commission, 2010). The terpenoids patchouli alcohol (patchouliol) and pogostone are important secondary metabolites in *P. cablin* (Lu *et al.*, 2011).

Due to quantitative variations in the contents of patchouliol and pogostone, the effectiveness of *P. cablin* preparations in patients is variable. *P. cablin* can be divided into two chemotypes, *i. e.* the patchouliol-type and the pogostone-type, which cannot be readily distinguished on the basis of morphological criteria, colour, texture or odour, especially when the herbs are in a dried and/or sliced state. Up to date, the GS-MS method has been advocated to identify the chemotypes of *P. cablin*. Unfortunately, there have been shortages in the availability of the standard samples (patchouliol and pogostone), so the GS-MS method does have limitations in practice. Therefore, it is imperative to find new methods for the identification of these two chemotypes of *P. cablin*.

DNA barcording is a new biological tool to achieve accurate, rapid, and automatable identification of species and/or cultivars, without morphological knowledge, by using a short universal genome sequence (Hebert *et al.*, 2003). DNA barcodes can provide a very efficient way for the quick discovery of new species and the accurate discrimination between species (Kress and Erickson, 2007; Lahaye *et al.*, 2008; Chen *et al.*, 2010). Different environmental conditions might change the chemical composition of patchouli essential oil (PEO), not only in quality but also in quantity. Thus, we propose that DNA bar-

coding can be used for identifying cultivars of *P. cablin* with different chemotypes. Here, we selected five candidate DNA barcodes including *rbcL*, *psbA-trnH*, *rpoB*, ITS (internal transcibed spacer), and *ndhJ* for this purpose. Our aims were (1) to examine the reliability, suitability, and reproducibility of these five genomic regions as barcodes, and (2) to evaluate the congruence between the GC-MS method and the DNA barcoding for accurate authentication and identification of the patchouliol- and/or the pogostone-type cultivars of *P. cablin*.

Materials and Methods

Plant specimens and genomic DNA extraction

Fresh leaves of samples were kept in silica gel upon collection from Guangdong Province, Hainan Province, and Guangxi Zhuang Autonomous Region, China. Voucher specimens were deposited at the College of Pharmacy, Chengdu University of Traditional Chinese Medicine, Chengdu, China. Overall 10 cultivation areas covering all main areas currently used in Southern China were examined for sequence variations within the regions of the candidate barcodes. Details of the origins of the herbarium vouchers are provided in Fig. 1 and Table I.

The specimens were frozen in liquid nitrogen and ground into a fine powder. Total genomic DNA was extracted from the powder using the cetyltrimethylammonium bromide (CTAB) method as modified from Doyle and Doyle (1987). The concentration of genomic DNA was measured in an Eppendorf biophotometer (Hamburg, Germany).

PCR amplification and sequencing

PCR (polymerase chain reaction) amplification was carried out in a Gene Amp PCR system 9700 (PE Applied Biosystems, Foster City, CA, USA) with 20 μL reaction mixture containing 5 μL 5×PCR buffer (including 2.5 mM $MgCl_2$), 2 μL dNTP Mix (2.5 mM each), 1 μL 10 μM each primer, 2–5 μg template DNA, and 0.25 μL Pfu DNA polymerase. Information on primers and thermocycling conditions for the five markers used in this study are listed in Table II. Purified PCR products were sent to Invitrogen Biotechnology (Shanghai, China) for sequencing of both strands.

Sequence alignment and analysis

Sequences were assembled using SeqMan of the DNASTAR software package (Burland, 2000); the sequences were aligned using CLUSTAL X (Thompson *et al.*, 1997) and adjusted manually if necessary. The complete sequences of the candidate DNA barcodes from each sample were stored as a separated text file. The newly acquired DNA sequences have been deposited in the NCBI Database (http://www.ncbi.nlm.nih.gov).

Fig. 1. Map showing the 10 collection sites in the main cultivation areas of *P. cablin* in Southern China. Some sites are superimposed: (1) Tanshui village, Yangchun City; (2) Sanjia village, Yangchun City; (3) Qindou village, Leizhou City; (4) Yingli village, Leizhou City; (5) Shipai village, Guangzhou City; (6) Liantang village, Gaoyao City; (7) Guding village, Gaozhou City; (8) Santang village, Nanning City; (9) Changfeng village, Wanning City; (10) Liji village, Wanning City.

Table I. List of samples used in this study.

No.	Place of collection	Date of collection	Voucher number	Number of collections in different fields
1	Tanshui village, Yangchun City, Guangdong Province	September 2012	CDCM 20120920	11
2	Sanjia village, Yangchun City, Guangdong Province	September 2012	CDCM 20120921	7
3	Qindou village, Leizhou City, Guangdong Province	September 2012	CDCM 20120923	8
4	Yingli village, Leizhou City, Guangdong Province	September 2012	CDCM 20120924	10
5	Shipai village, Guangzhou City, Guangdong Province	October 2012	CDCM 20121014	10
6	Liantang village, Gaoyao City, Guangdong Province	October 2012	CDCM 20121016	11
7	Guding village, Gaozhou City, Guangdong Province	October 2010	CDCM 20121018	12
8	Santang village, Nanning City, Guangxi Province	August 2012	CDCM 20120818	12
9	Changfeng village, Wanning City, Hainan Province	August 2012	CDCM 20120825	11
10	Liji village, Wanning City, Hainan Province	August 2012	CDCM 20120826	10

Table II. Primers and reaction conditions used in PCR.

Gene	Primer sequence $5' - 3'$	Reaction condition
ITS	Forward: GAAGTCGTAACAAGGTTTCCGTAGG Reverse: TCCTCCGCTTATTGATATGC	98 °C, 5 min 98 °C, 10 s; 56 °C, 15 s; 72 °C, 45 s; 35 cycles 72 °C, 7 min
rpoB	Forward: AAGTGCATTGTTGGAACTGG Reverse: GATCCCAGCATCACAATTCC	98 °C, 5 min 98 °C, 10 s; 53 °C, 15 s; 72 °C, 40 s; 35 cycles 72 °C, 7 min
rbcL	Forward: ATGTCACCACAAACAGAGACTA Reverse: TCGCATGTACCTGCAGTAGC	98 °C, 5 min 98 °C, 10 s; 55 °C, 15 s; 72 °C, 1 min; 35 cycles 72 °C, 7 min
psbA-trnH	Forward: CGCGCATGGTGGATTCACAATCC Reverse: GTTATGCATGAACGTAATGCTC	98 °C, 5 min 98 °C, 10 s; 55 °C, 15 s; 72 °C, 1 min; 34 cycles 72 °C, 7 min
ndhJ	Forward: CATAGATCTTTGGGCTTYGA Reverse: TCAATGAGCATCTTGTATTTC	98 °C, 5 min 98 °C, 10 s; 53 °C, 15 s; 72 °C, 40 s; 40 cycles 72 °C, 7 min

Sources: Takaiwa *et al.* (1985), White *et al.* (1990), and Song *et al.* (2009).

Identification of chemotype by DNA barcodes

The chemotype of each sample was identified by the GC-MS method (Zhang *et al.*, 2012). Five candidate DNA barcodes (*rbcL*, *psbA-trnH*, *rpoB*, ITS, and *ndhJ*) of standard samples of these two chemotypes of *P. cablin* were sequenced and aligned. Based on the sequence alignment, the variable regions of each candidate DNA barcode were calculated.

ITS regions from the samples collected in the various cultivation areas were amplified and sequenced. Each amplified sequence was aligned, and homologies with those of the pogostone-type (GenBank accession No. KJ413188) and patchouliol-type (GenBank accession No. KF771661) were analysed. Based on the respective extent of the homology with sequences from either of the two chemotypes, assignment to either of the two chemotypes was done and compared with the respective result of the GC-MS analysis.

Results and Discussion

Amplification efficiency and sequence recovery

High amplification efficiency and sequence recovery are the basic requirements for the suitability of a DNA barcode (Meyer and Paulay, 2005; CBOL Plant Working Group, 2009). In this study, each single primer pair for the five potential barcodes performed without fail, as all samples of *P. cablin* were successfully amplified (Fig. 2) and sequenced, indicating the high applicability of all five DNA barcodes in identifying *P. cablin*. However, as far as these five barcodes are concerned, amplification efficiency was low

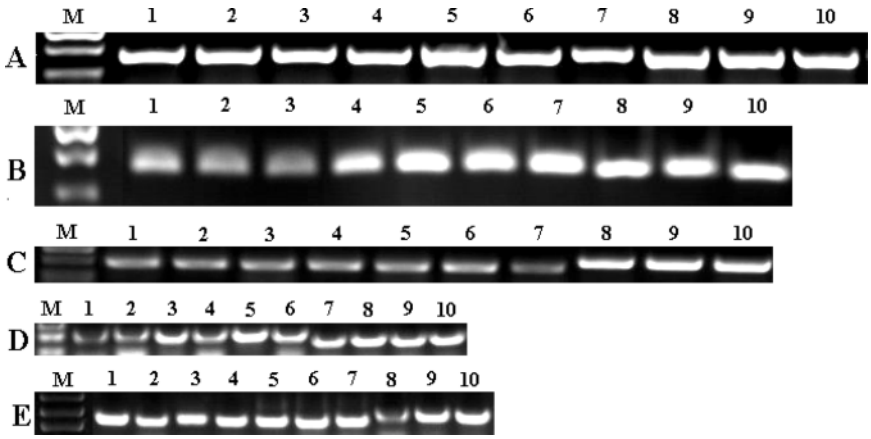

Fig. 2. Amplification of DNA barcoding candidate genes of *P. cablin*. (A) *rbcL*: lane M, marker; lanes 1 – 10, different samples; (B) ITS: lane M, marker; lanes 1 – 10, different samples; (C) *psbA-trnH*: lane M, marker; lanes 1 – 10, different samples; (D) *rpoB*: lane M, marker; lanes 1 – 10, different samples; (E) *ndhJ*: lane M, marker; lanes 1 – 10, different samples. (Origins of samples are given in Fig. 1 and Table I.)

Table III. Comparison of the variability of the five barcode candidates in *P. cablin* samples.

Feature	ITS	*rpoB*	*rbcL*	*psbA-trnH*	*ndhJ*
Primer universality	Yes	Yes	Yes	Yes	Yes
PCR success (%)	100	100	78	100	100
Sequencing success (%)	100	100	100	100	100
Aligned sequence length [bp]	618	481	695	286	389
Variable sites	22	6	2	3	0

for the *rbcL* sequence (78 %), while it was 100 % for the other four DNA candidate barcodes (Table III). All PCR products of these five DNA sequences could be successfully sequenced.

Validation of DNA barcoding markers

In recent years, DNA barcodes have attracted increasing attention and have been employed to identify different biological materials (Hebert *et al.*, 2004; Kress *et al.*, 2009; Chen *et al.*, 2010) including those from medicinal plants (Techen *et al.*, 2014), but identification of different chemotypes in *P. cablin* by DNA barcoding has not yet been reported. We selected five candidate DNA barcodes (*rbcL*, *psbA-trnH*, *rpoB*, ITS, and *ndhJ*) to potentially assign each of the 103 samples from 10 different cultivation areas to either of the two chemotypes. Large dataset screening had previously revealed that *ndhJ* sequences of seed plants are identical, while the *rbcL*, *psbA-trnH*, and *rpoB* sequences exhibited slight differences, and only the ITS sequences showed clear differences (Kress *et al.*, 2005; China Plant BOL Group, 2011) (Table III). These results indicate that ITS is more discriminating than any other candidate barcode (Ren *et al.*, 2010; Liu *et al.*, 2011).

Applicability of barcode to accurate chemotype discrimination

At present, except the GS-MS method, no reliable method is available to assign a sample to either of the two chemotypes of *P. cablin* (Zhang *et al.*, 2012). For identification of the chemotype by GC-MS, preparation of the essential oil from a sample of about 100 g plant material is required. For the barcoding method, a sample of only about 0.3 g is sufficient for DNA isolation. So, compared to the GC-MS method, the DNA barcoding method can be recommended as the primary (if only small amounts of samples and/or no standard samples of patchouliol and pogostone are available) or subsidiary method to identify the chemotype of a *P. cablin* sample.

We suggest that DNA barcoding can serve as an efficacious method to identify the chemotype of samples of *P. cablin* collected in Southern China, including Guangdong, Hainan, and Guangxi. After alignment of the sequences of the five candidate DNA bar-

codes (*rbcL*, *psbA-trnH*, *rpoB*, ITS, and *ndhJ*), we confirmed that ITS has more variable sites (22) than *rbcL* (2), *psbA-trnH* (3), *rpoB* (6), and *ndhJ* (0) (Table III). Thus, ITS was chosen as the most suitable DNA barcode for the effective identification of *P. cablin*. The results showed that the samples from product areas Nos. 5 and 6 (Guangzhou and Gaoyao) are of the pogostone-type, with concentrations of patchouliol between 15 % and 33 % and pogostone of 13 % to 36 % in the essential oil. Samples from the other production areas (Nos. 1, 2, 3, 4, 7, 8, 9, 10 from Yangchun, Leizhou, Gaozhou, Nanning, and Wanning) are of the patchouliol-type, with patchouliol concentrations between 22 % and 43 %, and those of pogostone in the range of 0 % to 8 %. Comparing the results obtained with the DNA barcoding method and the GC-MS method, we calculated that the identification accuracy can reach 93.2 %, as a total of 96 of the 103 samples was correctly identified, while only a single sample from each product area (Nos. 1, 4, 5, 7, 8, 9, 10; see also Table I) could not be accurately identified. Taken together, we advocate that barcoding using ITS can serve as a practical tool to verify and identify the chemotype of a *P. cablin* sample.

Acknowledgement

This work was supported by grants from the China Postdoctoral Science Foundation funded project (No. 2013M531944) and the Applied Basic Research Program of Sichuan Province (No. 2012JYZ004).

Burland T. G. (2000), DNASTAR's Lasergene sequence analysis software. Methods Mol. Biol. **132**, 71 – 91.

CBOL Plant Working Group (2009), A DNA barcode for land plants. Proc. Natl. Acad. Sci. USA **106**, 12794 – 12797.

Chen S. L., Yao H., Han J. P., Liu C., Song J. Y., Shi L. C., Zhu Y. J., Ma X. Y., Gao T., Pang X. H., Luo K., Li Y., Li X. W., Jia X. C., Lin Y. L., and Leon C. (2010), Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One **5**, e8613.

Chen M., Zhang J., Lai Y., Wang S., Li P., Xiao J., Fu C., Hu H., and Wang Y. (2013), Analysis of *Pogostemon cablin* from pharmaceutical research to market performances. Expert Opin. Invest. Drugs **22**, 245 – 257.

China Plant BOL Group (2011), Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. Proc. Natl. Acad. Sci. USA **108**, 19641 – 19646.

Chinese Pharmacopoeia Commission (2010), The Pharmacopoeia of the People's Republic of China, Vol. 1. Chemical Industry Press, Beijing, China.

Doyle J. J. and Doyle J. L. (1987), A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. **19**, 11 – 15.

Guan L., Quan L., and Cong P. (1992), Study on chemical constituents of volatile oil from *Pogostemon cablin* (Blanco) Benth. Nat. Prod. Res. Dev. **2**, 5.

Hebert P. D., Cywinska A., Ball S. L., and de Waard J. R. (2003), Biological identifications through DNA barcodes. Proc. Biol. Sci. **270**, 313 – 321.

Hebert P. D., Penton E. H., Burns J. M., Janzen D. H., and Hallwachs W. (2004), Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc. Natl. Acad. Sci. USA **101**, 14812 – 14817.

Hu L. F., Li S. P., Cao H., Liu J. J., Gao J. L., Yang F. Q., and Wang Y. T. (2006), GC-MS fingerprint of *Pogostemon cablin* in China. J. Pharm. Biomed. **42**, 200 – 206.

Kress W. J. and Erickson D. L. (2007), A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. PLoS One **2**, e508.

Kress W. J., Wurdack K. J., Zimmer E. A., Weigt L. A., and Janzen D. H. (2005), Use of DNA barcodes to identify flowering plants. Proc. Natl. Acad. Sci. USA **102**, 8369 – 8374.

Kress W. J., Erickson D. L., Jones F. A., Swenson N. G., Perez R., Sanjur O., and 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.

Lahaye R., Van der Bank M., Bogarin D., Warner J., Pupulin F., Gigot G., Maurin O., Duthoit S., Barraclough T. G., and Savolainen V. (2008), DNA barcoding the floras of biodiversity hotspots. Proc. Natl. Acad. Sci. USA **105**, 2923 – 2928.

Liu J., Möller M., Gao L. M., Zhang D. Q., and Li D. Z. (2011), DNA barcoding for the discrimination of Eurasian yews (*Taxus* L., Taxaceae) and the discovery of cryptic species. Mol. Ecol. Resour. **11**, 89 – 100.

Lu T. C., Liao J. C., Huang T. H., Lin Y. C., Chiu Y. J., and Peng W. H. (2011), Analgesic and anti-inflammatory activities of the methanol extract from *Pogostemon cablin*. Evid.-Based Compl. Alt. Med., 1 – 9.

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

Paul A., Bakshi S., Sahoo D. P., Kalita M. C., and Sahoo L. (2012), *Agrobacterium*-mediated genetic transformation of *Pogostemon cablin* (Blanco) Benth. using leaf explants: bactericidal effect of leaf extracts and counteracting strategies. Appl. Biochem. Biotechnol. **166**, 1871 – 1895.

Ren B. Q., Xiang X. G., and Chen Z. D. (2010), Species identification of *Alnus* (Betulaceae) using nrDNA and cpDNA genetic markers. Mol. Ecol. Resour. **10**, 594 – 605.

Song J., Yao H., Li Y., Li X., Lin Y., Liu C., Han J., Xie C., and Chen S. (2009), Authentication of the family Polygonaceae in Chinese pharmacopoeia by DNA barcoding technique. J. Ethnopharmacol. **124**, 434 – 439.

Takaiwa F., Oono K., and Sugiura M. (1985), Nucleotide sequence of the 17S – 25S spacer region from rice rDNA. Plant Mol. Biol. **4**, 355 – 364.

Techen N., Parveen I., Pan Z., and Khan I. A. (2014), DNA barcoding of medicinal plant material for identification. Curr. Opin. Biotechnol. **25**, 103 – 110.

Thompson J. D., Gibson T. J., Plewniak F., Jeanmougin F., and Higgins D. G. (1997), The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24**, 4876 – 4882.

White T. J., Bruns T., Lee S., and Taylor J. (1990), Amplication and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: PCR Protocols: A Guide to Methods and Applications (Innis M. A., Gelfand D. H., Sninsky J. J., and White T. J., eds.). Academic Press, San Diego, USA, pp. 315 – 322.

Xu X., Tang Z., and Liang Y. (2010), Comparative analysis of plant essential oils by GC-MS coupled with integrated chemometric resolution methods. Anal. Methods UK **2**, 359 – 367.

Zhang H., Chen X., Wei J., Yang X., Yu C., and Zou Z. (2012), Determination of patchouli alcohol and pogostone in Hainan *Pogostemon cablin*. China Pharm. **21**, 28.