Molecular identification of six medicinal *Curcuma* plants produced in Sichuan: Evidence from plastid *trn*K gene sequences

CAO Hui[1 *] , KOMATSU Katsuko[2]

(1. National Engineering Research Center for Modernization of Traditional Chinese Medicine, Zhuhai 519020, China; 2. Research Center for Ethnomedicine, Institute of Natural Medicine, Toyama Medical and Pharmaceutical University, Toyama 930 - 0194, Japan)

Abstract: **Aim** To establish a rapid and simple molecular identification method for six medicinals: *Curcuma*: *C. longa*, *C. phaeocaulis*, *C. sichuanensis*, *C. chuanyujin*, *C. chuanhuangjiang*, and *C. chuanezhu* in Sichuan Province. **Methods** A molecular approach (*trn*K nucleotide sequencing) was used in this study. **Results** The sequenced entire chloroplast *trn*K gene region spanned 2 699 - 2 705 bp. The *mat*K gene (an intron embodied in *trn*K gene) sequence and the intron spacer region of the *trn*K gene have great diversity within these six medicinal *Curcuma* species. There were six single bases substitutions between *trn*K coding region and *mat*K region, the 9-bp deletion and 4-bp or 14-bp insertion repeat at some sites of *mat*K region in each taxon. **Conclusion** These relatively variable sequences were potentially informative in the identification for these six *Curcuma* species at the DNA level.

Key words: *Curcuma*; nucleotide sequencing; *trn*K gene; molecular identification

CLC number: Q81 **Document code**: A **Article ID**: 0513 - 4870(2003)11 - 0871 - 05

6种川产姜黄属药用植物叶绿体 *trn*K 基因序列变异分析及其分子鉴定

曹 晖[1 *] , 小松かつ子[2]

(1. 国家中药现代化工程技术研究中心, 广东 珠海 519020; 2. 富山医科药科大学和汉药研究所, 日本 富山 930 - 0194)

摘要: 目的 建立6种川产姜黄属(*Curcuma*)药用植物快速简单的分子鉴定方法。方法 采用叶绿体赖氨酸tRNA基因(*trn*K)测序与序列变异分析方法。结果 6种姜黄属药用植物(包括姜黄 *C. longa*、莪术 *C. phaeocaulis*、川郁金 *C. sichuanensis*、川郁金 *C. chuanyujin*、川黄姜 *C. chuanhuangjiang*、川莪术 *C. chuanezhu*)完整 *trn*K 基因长度在2699 ~ 2705 bp。序列可变区包括 *mat*K 基因编码区和 *trn*K 外显子与 *mat*K 内含子之间区域,共有6个单核苷酸多态性(SNPs)位点、1个9-bp的缺失重复序列和2个4-bp、14-bp插入重复序列。结论 *trn*K 基因序列可变位点可以作为6种川产姜黄属药用植物快速简单的分子鉴定标记,并为它们之间种的归并提供了分子依据。

关键词: 姜黄属; 核苷酸测序; *trn*K 基因; 分子鉴定

Species identification of Chinese drugs is a prerequisite for standardization of themselves. Traditional ways rely on the inspection of morphological, histological as well as phytochemical markers. DNA is the basic component of a living organism, whereas phenotypic and chemical characters are controlled by the arrangement and expression of genes. With recent advances in molecular biology, comparative DNA sequences have become a widespread tool for species identification of medicinal materials. In our laboratories, many medicinal plants such as *Cnidium*, *Dioscorea*, *Ligusticum*, *Panax*, *Pinellia* and *Pogostemon* were successfully identified using 18S rRNA gene and *mat*K gene sequences by molecular means[1 - 7].

About 20 species of *Curcuma* are estimated to be found in China, Some of them are of medicinal and economic importance. Huangsi Yujin is derived from the tuber of *Curcuma longa*, Ezhu from the rhizome of *C.*

Received date: 2002-06-05.

* Corresponding author Tel:86 - 756 - 8135676; Fax:86 - 756 - 8289500; E-mail: kovhuicao@yahoo.com.cn

phaeocaulis. These two Chinese drugs are officially recorded in the Chinese Pharmacopoeia[8]. However, there are some confusion of botanical origins on Yujin and Ezhu in commercial market in Sichuan[9]. For instance, the former was often substituted by Chuanyujin/Huangbaisi Yujin (tubers of *C. sichuanensis* or *C. chuanyujin*), the latter by Chuanezhu/Wenzhu (rhizomes of *C. chuanezhu* or *C. chuanhuangjiang*) (Table 1), it makes species identification of a particular *Curcuma* drug difficult.

Table 1　The overlapping and confusion in botanical origin and commercial names of Yujin and Ezhu retailed in Sichuan market

Origin	Part used	Commercial drug name
Curcuma longa	Rhizome	Jianghuang (姜黄)
	Tuber	Huangsi Yujin (黄丝郁金)
C. sichuanensis	Rhizome	Chuanezhu (川莪术)
	Tuber	Chuan Yujin(川郁金)
C. chuanyujin	Rhizome	Chuanezhu(川莪术)
	Tuber	Huangbaisi Yujin (黄白丝郁金)
C. chuanhuangjiang	Rhizome	Wenzhu (文术)
	Tuber	Baisi Yujin(白丝郁金)
C. phaeocaulis	Rhizome	Ezhu(莪术)
	Tuber	Lüsi Yujin (绿丝郁金)
C. chuanezhu	Rhizome	Wenzhu/Chuanezhu(文术/川莪术)
	Tuber	Lüsi Yujin(绿丝郁金)

Detection of some medicinal *Curcuma* plants have currently been investigated using random amplified polymorphic DNA (RAPD) technique[10-11]. Although RAPD method does not require prior genetic background of the genome and can examine DNA polymorphism, the reliability of RAPD-based genomic fingerprint results can be greatly influenced by experimental conditions. These studies also indicate that interspecific differentiation can not be resolved completely. DNA sequencing method focuses on a defined locus on the genome and may provide definitive means for species identification of medicinal *Curcuma* plants[12-13]. In the present study, sequence analysis was performed using a chloroplast transfer RNA for lysine (*trn*K) among six medicinal *Curcuma* species in Sichuan for establishing a rapid and simple molecular identification of *Curcuma* drugs at DNA level.

Materials and methods

Plant materials　Six species of medicinal *Curcuma* were collected from different localities in Sichuan province, *Curcuma aromatica* from Botanical Garden of Institute of Medicinal Plants, The Chinese Academy of Medical Sciences, Beijing and one species of *Hedychium* from the Medicinal Plant Research Center of Toyama Prefecture, Japan were chosen as ingroup and outgroup taxon, respectively (Table 2). Four to five accessions were tested in each taxon from one source. All voucher specimens were identified by authors and deposited in the Museum of Materia Medica, Institute of Natural Medicine, Toyama Medical and Pharmaceutical University, Japan.

Table 2　Plant materials investigated in this study

Taxon	Source *	Voucher	GenBank Accession**
Curcuma longa L.	Shuangliu, Sichuan	Cao 9946	AB047738
C. sichuanensis X.X.Chen	Shuangliu, Sichuan	Cao 9948	AB047739
C. phaeocaulis Val.	Chongqin, Sichuan	Cao 9956	AB047735
C. chuanezhu Z.Y.Zhu	Emei, Sichuan	Cao 9951	AB047736
C. chuanyujin C.K.Hsieh et H.Zhang	Chongqin, Sichuan	Cao 9954	AB047733
C. chuanhuangjiang Z.Y.Zhu	Emei, Sichuan	Cao 9952	AB047732
C. aromatica Salisb.	BGIMP, Beijing	Cao 9801	AB047731
Hedychium spicatum Smith	MPRC, Toyama, Japan	Cao 0002	AB047754

* BGIMP: Botanical Garden of Institute of Medicinal Plants, The Chinese Academy of Medical Sciences. MPRC: Medicinal Plant Research Center of Toyama Prefecture. ** All the nucleotide sequence data of *trn*K gene in above table will appear in the GenBank, DDBJ and EMBL nucleotide sequence databases with accession numbers

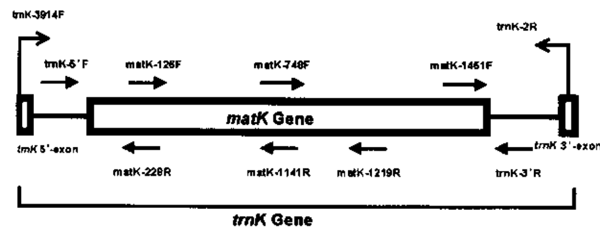
Genomic DNA preparation　Total DNA was prepared from fresh rhizomes or tubers, the tissues were frozen with liquid nitrogen into a fine powder. The DNA extraction method was a modification described in the operating manual of DNeasy plant mini-extraction kit (Giagen, Hilden, Germany).

The molecular size was determined by 1.0% in agarose gel electrophoresis comparing with λ*Hind* III marker (Life Technologies Inc., Rockville, MD). Photograph of the DNA banding pattern was made using an AE-6911FX printgraph system (Atto Co., Tokyo, Japan). The amount was detected using an UltroSpec 4000 UV-Visible Spectrophotometer (Pharmacia Biotech (Biochrom) Ltd., Cambridge, UK).

DNA amplification　Double-stranded DNAs of *trn*K regions were amplified using PCR with primers *trn*K-3914F and *trn*K-2R. The amplification reaction was performed in a volume of 50 μL containing 1x *Taq* buffer [10 $mmol\cdot L^{-1}$ Tris-HCl (9.0), 50 $mmol\cdot L^{-1}$ KCl, 1.5 $mmol\cdot L^{-1}$ $MgCl_2$ and 0.1% Triton X-100], 0.2 $mmol\cdot L^{-1}$ of each dNTPs, 1.5 U *Taq* polymerase (Promega, Madison, WI), 0.25 $mmol\cdot L^{-1}$ of each primer and 10 - 100 ng template DNA in a PTC-100

Programmable Thermal Controller (MJ Research, Inc., Watertown, MA), then subjected to the following thermal cycles: one hot start cycle for 3 min at 94 ℃, 35 cycles for 1 min at 94 ℃, 1 min at 52 ℃ and 2.5 min at 72 ℃, and a final extension cycle for 10 min at 72 ℃. All specimens were analyzed at least twice to confirm the results. The amplified product was resolved by 1.0% agarose gel electrophoresis as described above, 1 kb DNA ladder (Life Technologies Inc., Rockville, MD) was used as a size marker.

DNA Sequencing The PCR product was purified by QIAquick™ PCR purification kit (Qiagen, Hilden, Germany). The purified products were sequenced using a Thermo Sequenase™ fluorescent labeled primer cycle sequencing kit with 7-deaza-dGTP (Amersham Pharmacia Biotech UK Limited, Burkinghamshire, UK) and a set of fluorescent-labeled internal primers according to the manufacturer's recommendation (Figure 1). Labeled sequencing fragments were separated in 4% Long Ranger™ (acrylamide) gel and analyzed on a 4000L Automated DNA Sequencer (Li-Cor Inc., Lincoln, NE) according to the manual supplied.



The PCR amplification primers were *trn*K-3914F (5'-TGGGTTGCTAACTCAATGG-3') and *trn*K-2R (5'-AACTAGTCGGATGGAGTAG-3'). The base composition (5'-3') of sequencing primers is *trn*K-5'F: AGCAGCATGTCGTAATACGT; *mat*K-126F: GATAGATCTCTGGCAACAACAG; *mat*K-748F: AAGGATCCTTTCATGCATTAT; *mat*K-1451F: TATCCGTATCAATGACTTGG; *mat*K-228R: CAATGACTGCAAAACCTTCAG; *mat*K-1141R: ACAATCCGCTAAATCGGTCC; *mat*K-1219R: GTCGAAGTATATACTTTATTCG; *trn*K-3'R: TCCTTGTTATAATAGGTAAC

Figure 1 Structure and relative position of the PCR amplification and sequencing primers used for *trn*K gene. Boxed areas represent coding region. Arrows indicate the direction of strand synthesis

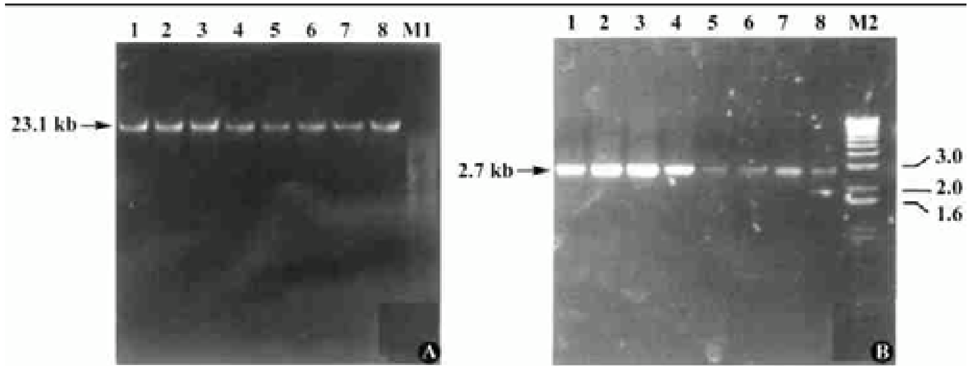
Data analysis The sequences were read and computed using Version 4.0 BaseImagIR™ program (Li-Cor Inc., Lincoln, NE), aligned using Version 1.3.0 AutoAssembler™ program (Applied Biosystems, Foster, CA) with minor manual adjustments. The boundaries of both *trn*K exon and *mat*K of all taxa were determined by comparison with published *trn*K data for rice (GenBank-X15901) and tobacco (GenBank-Z00044). The sequence divergence and pairwise genetic distance were calculated using the Distance Matrix option in Version 4βa PAUP program (Sinauer Associates, Sunderland, MA). The phylogenetic tree was generated by UPGMA method.

Results and discussion

The quality of total genomic DNA isolated from seven *Curcuma* species and one *Hedychium* species using CTAB miniprep method-based kit was shown in Figure 2A, which resulted in a clear band, ca. 23.1 kilobases (kb) in length. The quantitative measurement showed that yields of genomic DNA ranged from 132.4 to 178.9 ng/mg. The polymerase chain reaction amplification resulted in a single band corresponding to approximately 2.7 kb using *trn*K-2914F and *trn*K-2R primers (Figure 2B).

The sequence length of *trn*K region spanned 2699-2705 base-pair (bp). *Curcuma longa* and *C. sichuanensis* are 2699 bp, *C. phaeocaulis* and *C. chuanezhu* 2704 bp, and *C. chuanyujin* and *C. chuanhuangjiang* 2705 bp. The sequence divergence and absolute genetic distance of the *trn*K data among seven *Curcuma* species are shown in Table 3. The sequence divergence ranged from 0.000% - 0.224%, including base substitution numbers 0 - 6, transition numbers 0 - 5 and transversion numbers 0 - 1, there were three patterns in *trn*K sequence within seven *Curcuma* species. Six single base substitutions from the upstream at nucleotide position 146, 147, 645 and the downstream at 2511, 2575, 2602, one 1-bp gap at 512, one 9-bp deletion at 714 - 722, as well as two 4-bp, 14 bp insertion repeats were observed in 5'-intron spacer region between *trn*K exon and *mat*K gene (Figure 3).

The UPGMA dendrogram constructed by *trn*K sequences using Kimura two-parameter method showed that the *Curcuma chuanyujin* and *C. chuanhuangjiang* with hairy both surfaces on leaf were clustered firstly, followed by the order of five other species. Of the later cluster, *C. longa* and *C. sichuanensis* without hairy and with central spike become a group, closely related to *C. phaeocaulis* and *C. chuanezhu* with purple cloud along midrib on leaf blade and lateral spike than the former cluster in phylogenetic relationship (Figure 4).



Lane 1 : *Curcuma longa* ; Lane 2 : *C. sichuanensis* ; Lane 3 : *C. chuanhuangjiang* ; Lane 4 : *C. chuanyujin* ; Lane 5 : *C. phaeocaulis* ; Lane 6 : *C. chuanezhu* ; Lane 7 : *C. aromatica* ; Lane 8 : *Hedychium spicatum*. Lane M1 : λ*Hind* III , M2 : I kb DNA ladder

Figure 2 Agarose gel electrophoresis patterns of total genomic DNA isolated (A) and PCR amplification of plastid *trn*K gene from *Curcuma* and *Hedychium* (B)

	146-7	512	645	713 764		2511	2575	2602	
Curucma aromatica	AT	T	A	CC--------- TACA---- GGTTTTTTATTATTATTGTTTTATTATTATTG	T	A	C	T	(2705 bp)
C. chuanhuanhjiag	AT	T	A	CC--------- TACA---- GGTTTTTTATTATTATTGTTTTATTATTATTG	T	A	C	T	(2705 bp)
C. chuanyujin	AT	T	A	CC--------- TACA---- GGTTTTTTATTATTATTGTTTTATTATTATTG	T	A	C	T	(2705 bp)
C. longa	GC	-	G	CCTCTTTCCTTTACA---- GGTTTTTTATTATTATTG--------------	T	G	C	G	(2699 bp)
C. sichuanensis	GC	-	G	CCTCTTTCCTTTACA---- GGTTTTTTATTATTATTG--------------	T	G	C	G	(2699 bp)
C. phaeocaulis	GC	T	G	CCTCTTTCCTTTACATACAGGTTTTTTATTATTATTG--------------	T	G	T	G	(2704 bp)
C. chuanezhu	GC	T	G	CCTCTTTCCTTTACATACAGGTTTTTTATTATTATTG--------------	T	G	T	G	(2704 bp)

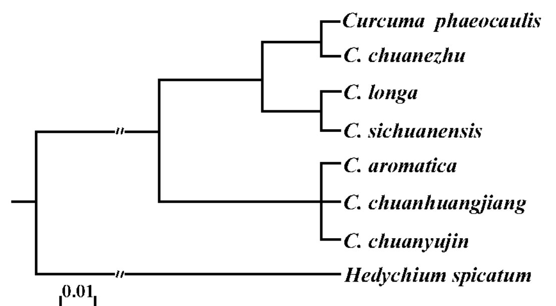
Hyphens (-) denote alignment gaps ; boxed hyphen indicated 9-bp deletion indel ; boxed letter showed 4-bp and 14-bp insertion repeat indels . Numbers in above the sequence indicate the aligned nucleotide position

Figure 3 Comparison of *trn*K gene sequences in seven *Curcuma* taxa

Table 3 Pairwise genetic distances among seven *Curcuma* species based on 2680 nucleotides of the *trn*K gene sequences. Upper-right diagonal: percentages of sequence divergence; lower-left diagonal: numbers of nucleotide substitutions (transition/transversion)

	1	2	3	4	5	6	7
1 *Curcuma longa*	-	0.000	0.037	0.037	0.037	0.224	0.224
2 *C. sichuanensis*	0(0/0)	-	0.037	0.037	0.037	0.224	0.224
3 *C. aromatica*	5(4/1)	5(4/1)	-	0.000	0.000	0.186	0.186
4 *C. chuanyujin*	5(4/1)	5(4/1)	0(0/0)	-	0.000	0.186	0.186
5 *C. chuanhuangjiang*	5(4/1)	5(4/1)	0(0/0)	0(0/0)	-	0.186	0.186
6 *C. phaeocaulis*	6(5/1)	6(5/1)	1(1/0)	1(1/0)	1(1/0)	-	0.000
7 *C. chuanezhu*	6(5/1)	6(5/1)	1(1/0)	1(1/0)	1(1/0)	0(0/0)	-

This investigation did not support some recently published *Curcuma* taxa as new species[14,15] , because the molecular data are not congruent with morphological characters for these seven species of *Curcuma* . *C. chuanhuangjiang* and *C. chuanyujin* have identical sequences in *trn*K gene , seem to be the same species as *C. aromatica* . Similarly , *C. chuanezhu* seem to be combined under *C. phaeocaulis* . Liu *et al* reported that *C. chuanyujin* could be placed under *C. kwangsiensis* based on the morphology[16] , but our result also do not agree their combination treatment .



The tree was outgroup-rooted using the *trn*K sequence of *Hedychium spicatum*. The distance corresponding to 0.01 sequence divergence is indicated by the bar

Figure 4 Dendrogram of phylogenetic relationship among seven *Curcuma* taxa based on *trn*K sequences using UPGMA method

Although Chen *et al* reported that *C. sichuanensis* and *C. wenyujin* seem to be combined according to

RAPD and chemical analysis[10], but our previous study showed that there are more great differences in *trn*K sequences between these two species[13]. The present study showed that *C. sichuanensis* might be combined under *C. longa*.

Conclusion

Nucleotide sequencing of plastid *trn*K gene could provide a novel information for origin identification of *Curcuma* species due to its higher mutation rate. Indeed, each group of *Curcuma* species was found to have a unique sequence pattern in the *trn*K gene region, so that they could be easily distinguished at the DNA level. On the other hand, these relatively variable sites in the non-coding region of *trn*K intron were potentially informative in design of species-specific primer used as a DNA probe to identify such *Curcuma* drugs as Yujin and Ezhu from Sichuan market.

Acknowledgements

We are grateful to Mr. Z. Y. Zhu of Sichuan Vocational School of Chinese Materia Medica in Emei, Mr. H. Shuai of Sichuan University in Chengdu, Dr. X. J. Ma of Institute of Medicinal Plants, Chinese Academy of Medical Sciences in Beijing, and Mr. M. Murakami of Medicinal Plant Research Center of Toyama Prefecture for help with material collection, Prof. N. Liu of South China Institute of Botany, Chinese Academy of Sciences for his valuable information. This work was supported in part by a Grant-in Aid for Scientific Research (B) No. 11695086 in 1999 - 2001 from Japan Society for the Promotion of Science, and by the Yamasaki Foundation for Promotion of Spice Research.

References:

[1] Cao H, Cai JN, Liu YP, *et al*. Correlativity analysis between geographical distribution and nucleotide sequence of chloroplast *mat*K gene of *Cnidium monnieri* fruit in China [J]. *Chin Pharm J* (中国药学杂志), 2001, **36**(6):373 - 376.

[2] Liu YP, He BZ, Cao H. Application of gene technology in quality control of Chinese drugs (II)- Identification of Chinese yam (*Dioscorea polystachya* rhizome) using DNA sequencing [J]. *Chin Trad Herbal Drugs* (中草药), 2001, **32**(11): 1026 - 1030.

[3] Fushimi H, Komatsu K, Isobe M, *et al*. Application of PCR-RFLP and MASA analyses on 18S ribosomal RNA gene sequence for the identification of three ginseng drugs [J]. *Biol Pharm Bull*, 1997, **20**(7):765 - 769.

[4] Liu YP, Cao H, Han GR, *et al*. *mat*K and ITS nucleotide sequencing of crude drug Chuanxiong and phylogenetic relationship between their species from China and Japan [J]. *Acta Pharm Sin* (药学学报), 2002, **37**(1):63 - 68.

[5] Cao H, Liu YP, Fushimi H, *et al*. Identification of Notoginseng (*Panax notoginseng*) and its adulterants using DNA sequencing [J]. *J Chin Med Mater* (中药材), 2001, **24**(6):398 - 401.

[6] Liu YP, Cao H, Wang XT. Application of gene technology in quality control of Chinese drugs (I)- Identification of *Pinellia ternata* species from Yuncheng, Shandong using DNA sequencing [J]. *Chin J Pharm Anal* (药物分析杂志), 2001, **21**(6):423 - 427.

[7] Luo JP, Cao H, Liu YP. DNA sequencing and molecular identification of Patchouli and its substitute Wrinkled Gianthyssop [J]. *Acta Pharm Sin* (药学学报), 2002, **37**(9):739 - 742.

[8] Guo JX. *Pharmacopoeia of the People's Republic of China* (1997 English Ed.) [S]. Beijing: Chemical Industry Press, 1997. Vol 1. 145 - 146, 190.

[9] Zhu ZY. Zingiberaceae. *Flora Sichuanica* (四川植物志) [Z]. Tomus 10. Chengdu: Sichuan Minority Press, 1992. 604 - 610.

[10] Chen YH, Bai SM, Cheng KD, *et al*. RAPD analysis on *Curcuma wenyujin* and *C. sichuanensis* [J]. *China J Chin Mater Med* (中国中药杂志), 1999, **24**(3):131 - 133.

[11] Xiao XH, Liu FQ, Shi CH, *et al*. RAPD polymorphism and authentication of medicinal plants from Turmeric (*Curcuma* L.) in China [J]. *Chin Trad Herbal Drugs* (中草药), 2000, **31**(3):209 - 212.

[12] Cao H, Sasaki Y, Fushimi H, *et al*. Molecular analysis of medicinally-used Chinese and Japanese *Curcuma* based on 18S rRNA gene and *trn*K gene sequences [J]. *Biol Pharm Bull*, 2001, **24**(12):1389 - 1394.

[13] Sasaki Y, Fushimi H, Cao H, *et al*. Sequence analysis of Chinese and Japanese *Curcuma* drugs on the 18S rRNA gene and *trn*K gene and the application of amplification refractory mutation system analysis for their authentication [J]. *Biol Pharm Bull*, 2002, **25**(12):1593 - 1599.

[14] Zhang H., Jiao WX, Fang XP. A new species of *Curcuma* (Zingiberaceae) from Sichuan [J]. *J West China Univ Med Sci* (华西医科大学学报), 1990, **21**(2):179 - 180.

[15] Chen XX. A new species of *Curcuma* Roxb [J]. *Guihaia* (广西植物), 1984, **4**(2):133 - 134.

[16] Liu N, Wu TL. Notes on *Curcuma* in China [J]. *J Trop Subtrop Bot* (热带亚热带植物学报), 1999, **7**(2):146 - 150.