

广藿香的基因序列与挥发油化学型的相关性分析

刘玉萍¹, 罗集鹏², 冯毅凡², 郭晓玲², 曹 晖^{3*}

(1. 中国中医研究院中药研究所, 北京 100700; 2. 广东药学院, 广东 广州 510224;
3. 国家中药现代化工程技术研究中心, 广东 珠海 519020)

摘要: 目的 探讨“南药”广藿香 *Pogostemon cablin* (Blanco) Benth. 不同产地间的叶绿体和核基因组的基因型与挥发油化学型的关系, 为广藿香道地性品质评价、规范化种植提供分子依据。方法 用 PCR 直接测序技术对广藿香 6 个产地样本的叶绿体 *matK* 基因和核 18S rRNA 基因核苷酸序列进行测序分析研究。结果 广藿香 6 个样本的 *matK* 基因序列长均为 1 245 bp, 编码 415 个氨基酸成熟酶。18S rRNA 基因序列长为 1 803 ~ 1 805 bp。根据排序比较, 广藿香 6 个样本间的 *matK* 基因序列存在 47 个变异位点, 18S rRNA 基因存在 17 个变异位点, 非加权组平均法构建的系统分支树表明广藿香基因序列分化与其产地、所含挥发油化学变异类型呈良好的相关性。结论 结合挥发油分析数据, 基因测序分析技术可作为广藿香道地性品质评价方法这一以及规范化种植过程关键技术“物种鉴定”的有力工具。

关键词: 广藿香; 道地药材; DNA 测序; 物种鉴定

中图分类号: Q78 文献标识码: A 文章编号: 0513-4870(2002)04-0304-05

目前, 关于中药道地性的研究仅限于形态组织学、化学、药效学等方面。从分子生物学角度视中药的道地性本质为该物种在特定条件下形成的基因型 (genotype), 而其次生代谢产物 (药效成分) 为药材品质的外在表型 (phenotype), 其形态及成分组成的多样性是由基因的遗传和变异引起的。因此, 采用基因序列分析手段对中药物种的地理分布和变迁进行评判与划分, 无疑为从分子水平阐释中药道地性提供崭新的线索, 同时结合药效成分分析, 对道地药材的品质评价将起指导性作用。

随着 DNA 分析技术的发展, 基于分子杂交的 RFLP 及基于 PCR 的 RAPD 和 DNA 测序等技术已应用于药用植物居群、地理品系研究领域^[1,2]。近年来, 我们对 PCR 直接测序技术在中药材鉴别方面的应用进行了许多开拓性研究^[3-6], 并成功开发出高效特异基因芯片简便快速地鉴别半夏与虎掌南星药材^[2]。此外还采用基因技术和化学分析手段对蛇床子叶绿体基因型、地理分布和香豆素化学型相关性 (道地性评价) 进行了探索性研究^[7]。

18S rRNA 基因为编码核糖体 DNA 的小亚基, 系多拷贝基因, 序列长度保守, 约 1 800 bp, 而 *matK* 基

因编码一种 500 余个氨基酸成熟酶, 为单拷贝基因, 位于 *tmK* 基因内含子中, 长约 1 500 bp, 是叶绿体基因组蛋白编码基因中进化速率最快的基因之一, 分辨率高。利用结构十分保守的 18S rRNA 基因和进化速率较快的 *matK* 基因进行序列比较与变异分析, 就能找出一种很好的分子标记用于中药材的鉴别和道地性评判与划分。

广藿香为唇形科植物广藿香 *Pogostemon cablin* (Blanco) Benth. 的干燥地上部分。系广东道地药材之一, 有芳香化浊、开胃止呕、发表解暑的功效, 是多种著名中成药“藿香正气丸”、“藿胆丸”等以及现代中药制剂“抗病毒口服液” (颗粒) 的主要原料。传统的商品药材按产地分为“石牌广藿香” (枝香)、“高要广藿香” (肇香)、“湛江广藿香” (湛香) 和“海南广藿香” (琼香) 4 种, 但因历史渊源与产区的生态环境、种植和采收时间及加工方法的不同, 其商品药材的性状和品质有所差异, 前两者产于广东省广州市郊和肇庆地区 (50 年代末自广州引种), 传统认为是道地药材, 质量优, 供药用; 后两者主要产于广东省湛江地区 (60 年代初自海南岛引种) 以及海南省等地, 质量较次, 不供药用, 仅用作提取广藿香油 (patchouli oil)。我们对不同产地广藿香挥发油成分分析, 发现其挥发油组成有明显差异^[8-10]。据此将其分为 2 个化学型: ① 广藿香酮型 (pogostone-type): 含氧组份以广藿香酮 (pogostone) 为主 (> 30%), 而广藿香醇含量低 (< 16%), 含氧萜烯类和不含氧萜烯类组份含量

收稿日期: 2001-08-27.

基金项目: 广东省自然科学基金和广东省中医药管理局资助课题。

作者简介: 刘玉萍 (1963-), 女, 博士, 副研究员;

曹 晖 (1961-), 男, 博士, 研究员。

* 通讯作者 Tel: (0756) 8135676, Fax: (0756) 8289500,

E-mail: kovhuicao@yahoo.com.cn

相近,栽培于广州市郊的黄村、石牌、棠下及肇庆地区高要;② 广藿香醇型 (Patchouliol-type): 含氧组份以广藿香醇 (patchouliol) 为主 (36.4%~47.8%), 广藿香酮含量较低 (2.3%~13.8%), 而不含氧萜烯类组份含量较高 (46.7%~54.3%), 栽培于湛江地区海康、吴川、遂溪以及海南省万宁等地^[11] (图 1)。

为探讨广藿香基因型与其产地分布、挥发油化学型的相关性, 本文采用 PCR 直接测序技术对 6 个

不同产地的广藿香核 18S rRNA 基因和叶绿体 *matK* 基因核苷酸序列进行了测序分析研究。

材 料 与 方 法

样品 本实验所用广藿香样品均为 1996~1998 年采自广东省和海南省不同产地的干燥叶片, 每个产地取 3~5 个样本。原植物经作者鉴定, 凭证标本存广东药学院中药标本馆 (表 1)。

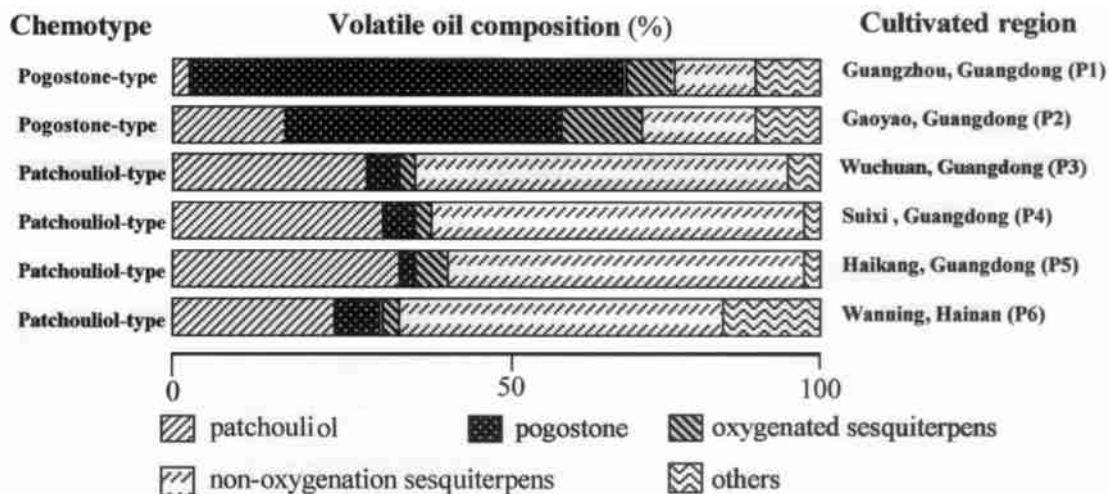


Figure 1 Volatile oil composition in the leaves of *Pogostemon cablin* from samples of different localities

Table 1 Origin of *Pogostemon cablin* samples used

Sample No.	Voucher	Place cultivated	Date collected	Chemotype *
P1	Luo JP 98081001	Guangzhou, Guangdong	Aug. 10, 1998	PO
P2	Luo JP 97112001	Gaoyao, Zhaoqing, Guangdong	Nov. 20, 1997	PO
P3	Li XQ 97111901	Wuchuan, Zhanjiang, Guangdong	Nov. 19, 1997	PA
P4	Li XQ 97111801	Suixi, Zhanjiang, Guangdong	Nov. 18, 1997	PA
P5	Li XQ 97111701	Haikang, Zhangjiang, Guangdong	Nov. 17, 1997	PA
P6	Luo JP 96082001	Wanning, Hainan	Aug. 20, 1996	PA

*PO: Pogostone-type; PA: Patchouliol-type

引物设计 18S rRNA 基因 PCR 扩增的通用引物据 Sogin 设计^[12] 而合成, *matK* 基因 PCR 扩增的通用引物据大井等设计^[13] 而合成。由于 18S rRNA 基因和 *matK* 基因长度分别为 1.8 kb 和 1.2 kb, 利用单个引物进行一次 PCR 测序反应无法测得此区段全部序列, 因此我们还设计合成了 2~3 对测序引物 (表 2)。

方法 总 DNA 按作者摸索的 CTAB 微量提取方法^[14], PCR 扩增反应和测序反应条件同前文^[3,4], 序列数据分析采用 Clustal W 程序进行 DNA 双向排序, 用 Genetyx-SV/R 软件以非加权组平均法 (UPGMA) 进行分支分析。

Table 2 Amplification and sequencing primers used in this study

Name of primer	Sequence (5'→3')
18S rRNA gene	
PCR primer	18S-5F CAA CCT GGT TGA TCC TGC CAG T
	18S-3R CTG ATC CTT CTG CAG GTT CAC CTA C
Sequencing primer	18SF1 GAT CCT GCC AGT CAT ATG C
	18SF2 GTT GGC CTT CGG GAT CGC AG
	18SF3 TTT GTC TGG TTA ATT CCG TT
	18SR1 TTA TTG TCA CTA CCT CCC CG
	18SR2 CCT GGT CGG CAT CGT TTA TG
	18SR3 GAC CTA CGG AAA CCT TGT TAC G
<i>matK</i> gene	
PCR primer	<i>matK</i> -AF CTA TAT CCA CTT ATC TTT CAG GAG T
	<i>matK</i> -8R AAA GTT CTA GCA CAA GAA AGT CGA
Sequencing primer	<i>matK</i> F1 GCA CTT GCG CAT GAT CAT GGT
	<i>matK</i> F2 TGG TTG TTT AAG GAC CCT TT
	<i>matK</i> R1 TCC AAA TAC GAT ACT CGT GGA G
	<i>matK</i> R2 AAA GTC GAA GTA TAT ACT TTA

结 果

1 DNA 提取和 PCR 扩增

在本研究中, 6 个不同产地广藿香用作者改进的 CTAB 微量法从干燥药材叶片提取到较高质量的

总基因组 DNA, 其分子长度均在 23.1 kb 左右(图 2A)。18S rRNA 基因和 *matK* 基因的 PCR 扩增能分别得到预期 1.8 kb 和 1.2 kb 的双链产物(图 2B, 2C)。

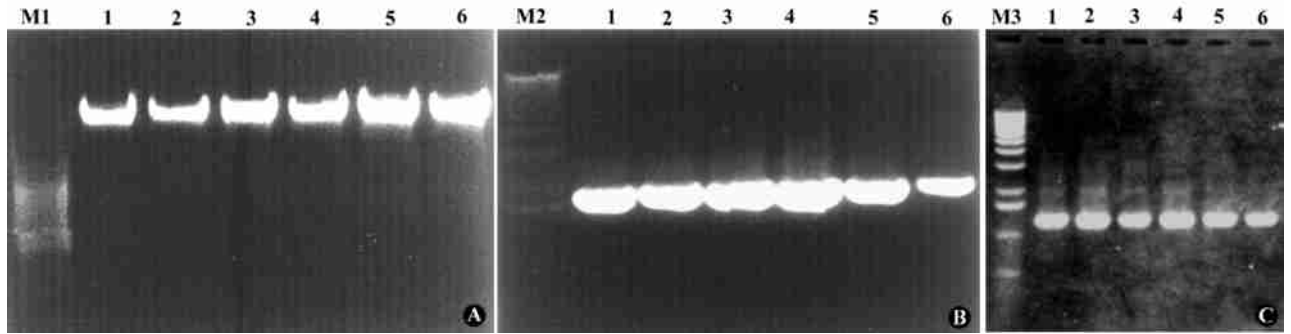


Figure 2 Agarose electrophoresis pattern of total DNA isolated (A), 18S rRNA gene amplification (B) and *matK* gene amplification (C) of *Pogostemon cablin* from six samples from different localities
Lane 1~6, P1~P6; M1, λHind III marker; M2~3, 1 kb DNA ladder marker

2 DNA 序列测定

不同产地广藿香 *matK* 基因的核苷酸序列长度相同, 均为 1 245 bp, 编码 415 个氨基酸成熟酶。根据排列比较, 6 个广藿香样本间的 *matK* 基因核苷酸序列存在 47 个变异位点(图 3)。而 18S rRNA 基因的核苷酸序列长度不等, 其中广州黄村的“石牌广藿

香”(P1)为 1 805 bp, 肇庆地区高要的“高要广藿香”(P2)与湛江地区海康的“湛江广藿香”(P5)、海南万宁的“海南广藿香”(P6)为 1 804 bp, 湛江地区吴川(P3)、遂溪(P4)的“湛江广藿香”为 1 803 bp。它们的 18S rRNA 核苷酸序列间也有 17 个变异位点(图 4)。

```

8 9 1 1 1 1 1 1 1 2 2 2 3 5 5 5 5 6 6 6 6 6 6 6 7 7 7 7 7 8 8 9 9 9 1 1 1 1 1 1 1 1 1 1 1 1 1
9 6 2 2 3 3 3 4 5 9 0 3 4 6 5 6 7 8 9 0 1 1 3 4 7 9 0 0 1 1 2 6 7 2 5 5 0 0 1 1 1 1 1 1 1 1 2 2
4 8 0 4 9 3 2 8 0 6 0 8 4 9 1 0 5 0 2 9 3 1 4 6 1 6 1 8 9 7 0 5 0 4 7 7 3 4 4 4 6 6 9 0 0
                                0 2 8 0 2 4 1 3 9 0 1
P1  G T C C C C T T F C C C A G T C T G G A G A G T T F G G C C C G G G G C T A C C C G G C T T G T T
P2  G G A G T C A G T G G G G T G T F C C A C C T T C C C C C C T G G G T G C C T G T T C T C C G
P3  G G A G T A G G T G T G G T C G C A A C G G T T C G C A G G G T G C T A C T T T T G G T T T T
P4  T G A A T G G T G T T T G T C G T C A T G G T G T G G T G T G G G G G G G G C T G C C C C G T G
P5  T G A A T G G T T T T G G C G A G A C G G G T C A G A A C A T G C C G A G C C T G C C C T C T
P6  T G A G T T G T G T G A T T C T C A A C A A G T T T A C G C T T G G G A G C C G T C C T G C G
    
```

Figure 3 Comparison of variable sites in *matK* sequence from six samples of *Pogostemon cablin* in different localities (P1~P6)
Upper number indicates nucleotide position from upstream of *matK* sequence

	89	200	357	432	471	501	680	846	1022	1060	1130	1421	1477	1479	1512	1517		
P1	C	G	G	C	G	C	C	C	C	T	A	T	C	A	C	A	C	= 1805 bp
P2	—	*	A	A	C	*	*	*	*	*	*	C	*	*	*	*	*	= 1804 bp
P3	*	A	*	*	*	*	A	—	G	—	G	*	*	*	*	*	C	= 1803 bp
P4	*	A	*	*	*	*	A	—	*	—	T	*	*	*	*	*	C	= 1803 bp
P5	*	*	*	*	*	T	A	—	*	*	*	*	*	*	*	A	G	= 1804 bp
P6	*	*	*	*	*	T	A	—	*	*	*	*	A	A	*	A	G	= 1804 bp

Figure 4 Comparison of variable sites in 18S rRNA sequence from six samples of *Pogostemon cablin* in different localities (P1~P6)
Upper number indicates nucleotide position from upstream of 18S rRNA sequence, asterisk (*) indicates identity to base composition of P1 sequence; hyphen (—) indicates alignment gap

3 基因序列位点多态性分析

本实验选取的广藿香有 2 个化学型, 表 3 显示 6 个产地广藿香(2 个化学型)间 *matK* 基因位点多态性(site polymorphism), 结果表明其序列的同源性为 94%~99%(相对遗传距离为 1%~6%)。化学型内和化学型间的碱基变异位点数(绝对遗传距离)相近, 如广藿香酮型(P1~P2)与广藿香醇型(P3~P6)间为 24~32 个, 而 4 个广藿香醇型样本间的碱基变异位点是 23~31 个。但是 2 个化学型间碱基变异类型明显不同, 如广藿香酮型和广藿香醇型间 *matK* 基因序列的碱基颠换数目 6~14 个; 广藿香醇型内颠换数目 7~9 个, 反映了 *matK* 基因序列在化学型间的分化离散程度大于化学型内。图 4 亦显示 4 个广藿香醇型内 18S rRNA 基因位点多态性以海南省万宁(P6)样本的碱基变异位点数多于广东省湛江地区(P3~P5)样本, 说明广藿香的 18S rRNA 和 *matK* 序列均存在种内多态性(intraspecific polymorphism)。

Table 3 Variable site number (below diagonal) and transition/transversion numbers (above diagonal) of *matK* gene sequence of *Pogostemon cablin* from different localities

Sample No.	Chemotype*					
	PO		PA			
	P1	P2	P3	P4	P5	P6
P1		21/9	19/8	22/10	17/14	23/7
P2	30		18/6	19/11	24/8	16/9
P3	27	24		22/9	19/9	20/9
P4	32	30	31		16/7	14/9
P5	31	32	28	23		22/7
P6	30	25	29	23	29	

* PO: Pogostone-type; PA: Patchouliol-type

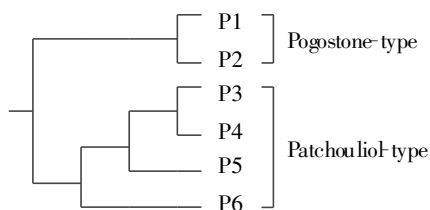


Figure 5 Cluster tree of both *matK* and 18S rRNA gene sequences of *Pogostemon cablin* of six samples from different localities constructed by UPGMA method

讨 论

广藿香植物作为一个自然分类群属于亚洲热带种, 原产东南亚菲律宾、马来西亚等国, 叶缘锐锯齿状^[5]。然自宋代从南洋传入我国岭南地区(今广

东)后, 经过长期人工栽培, 叶形态发生变化, 叶缘成钝锯齿状^[6]。在我国各地引种的广藿香形态上虽有一些变异, 但干燥药材差别不明显, 采用传统的形态组织学方法难以区分种内的分化类型, 且各地广藿香含油量有所差异。我们曾报道^[1], 广藿香挥发油化学组成与产地分布、种植和采收季节有关。

本研究测得的不同产地广藿香间的 *matK* 基因和 18S rRNA 基因 DNA 序列存在较大的差异(图 3, 4)。根据 *matK* 和 18S rRNA 基因序列共同构建的 UPGMA 系统分支树(图 5), 6 个产地广藿香明显可分成两个组, 即属广藿香酮型的广东省广州“石牌广藿香”(P1)与邻近的肇庆地区“高要广藿香”(P2)聚为一组, 而远离属广藿香醇型的广东省湛江地区“湛江广藿香”与海南省万宁“海南广藿香”组。在后一组中, 吴川(P3)和遂溪(P4)产广藿香先聚为一支, 再与海康(P5)产广藿香相聚, 最后与万宁(P6)相聚。说明广藿香植物的 *matK* 和 18S rRNA 序列分化(基因型)与地理分布由北向西南(道地性)及挥发油成分变异(化学型)呈良好的相关性。由此, 我们认为广藿香存在产地及挥发油成分不同的分化类型, 是适应特定环境的结果, 本质上由遗传变异所决定。由于本研究测得的 *matK* 属于叶绿体基因组 DNA, 一般通过母性单亲遗传, 而 18S rRNA 为核基因组 DNA, 系双亲遗传的基因, 由 rDNA 基因构建的分支关系能真正反映种内进化方向。*matK* 和 18S rRNA 基因序列和有关挥发油化学组成数据, 不仅可为“南药”广藿香的道地性评价提供分子依据, 而且基因测序分析技术可以成为广藿香规范化种植过程关键技术“物种鉴定”的强有力工具。

REFERENCES:

- [1] Cao H, Liu YP, Komatsu K, et al. DNA molecular profiling: A new approach to quality control of Chinese drugs [J]. *Chin J Integrated Tradit Chin West Med* (中国中西医结合杂志), 2000, 6(1): 71-75.
- [2] Liu YP, Cao H, Komatsu K, et al. Quality control for Chinese herbal drugs using DNA probe technology [J]. *Acta Pharm Sin* (药学报), 2001, 36(6): 475-480.
- [3] Cao H, Liu YP, Zhang SL, et al. Nuclear ribosomal RNA small subunit (18S rRNA) nucleotide sequencing and characterization of Sialonggu (whole bone of *Myosorex baileyi* Thomas) [J]. *China J Chin Mater Med* (中国中药杂志), 2001, 26(2): 90-94.
- [4] 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): 398-402.

- [5] 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.
- [6] 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 Tradit Herb Drugs* (中草药), 2001, **32**(11): 1026-1030.
- [7] Cao H, Cai JN, Liu YP, et al. Correlativity analysis between geographical distribution and nucleotide sequence of chloroplast *matK* gene of *Cnidium monnieri* fruit in China [J]. *Chin Pharm J* (中国药学杂志), 2001, **36**(6): 373-376.
- [8] Luo JP, Feng YF, Guo XL. Analysis of volatile oil of *Pogostemon cablin* [J]. *Chin Tradit Herb Drugs* (中草药), 2001, **32**(4): 299-302.
- [9] Luo JP, Feng YF, Guo XL, et al. Analysis of volatile oil of *Pogostemon cablin* from Gaoyao county [J]. *J Chin Med Mater* (中药材), 1999, **22**(1): 25-27.
- [10] Feng YF, Guo XL, Luo JP, et al. GC-MS analysis of volatile oil of *Pogostemon cablin* from Leizhou [J]. *J Chin Med Mater* (中药材), 1999, **22**(5): 241-242.
- [11] Luo JP, Feng YF, Guo XL. Influence in volatile oil constituents of *Pogostemon cablin* in different cultivated area and harvesting time [C]. *Proceedings of 2000 Annual Meeting of Chinese Pharmaceutical Association* (2000年中国药学会学术年会论文集). Part 2. Ningbo, 2000. 592-593.
- [12] Sogin ML. Amplification of ribosomal RNA genes for molecular evolution studies [A]. Innis M, Gelfand D, Sninsky J. *PCR Protocols: A Guide to Methods and Application* [M]. San Diego: Academic Press 1990. 307-314.
- [13] Ooi K, Endo Y, Yokoyama J, et al. Useful primer designs to amplify DNA fragments of the plastid gene *matK* from angiosperm plants [J]. *J Jpn Bot*, 1995, **70**(3): 328-331.
- [14] Cao H, But PPH, Shaw PC. Methodological studies on DNA extraction and purification from commercial drug materials [J]. *J Chin Pharm Sci*, 1998, **7**(3): 130-137.
- [15] Quisumbing E. *Medicinal Plants of the Philippines* [M]. Manila: Bureau of Printing, 1951. 829-830.
- [16] Huang YC. *Flora Reipublicae Populans Sinicae* (中国植物志) [M]. Vol 66. Beijing: Science Press 1977. 370-371.

DNA PROFILING OF *POGOSTEMON CABLIN* CHEMOTYPES DIFFERING IN ESSENTIAL OIL COMPOSITION

LIU Yu-ping¹, LUO Ji-peng², FENG Yi-fan², GUO Xiao-ling², CAO Hui³

(1. Institute of Chinese Materia Medica, China Academy of Traditional Chinese Medicine, Beijing 100700, China; 2. Guangdong College of Pharmacy, Guangzhou 510224, China; 3. National Engineering Research Center for Modernization of Traditional Chinese Medicine, Zhuhai 519020, China)

ABSTRACT: **AIM** To provide molecular evidence for quality evaluation and GAP production of *Pogostemon cablin* (Blanco) Benth. cultivated in different regions in Guangdong and Hainan provinces, China, by comparing two sequences (1.2 kb of plastid *matK* gene and 1.8 kb of nuclear 18S rRNA gene) and two chemotypes (Pogostone-type and Patchouliol-type in essential oil composition). **METHODS** PCR direct sequencing was applied to determine the *matK* and 18S rRNA sequences for six samples of *Pogostemon cablin* from different localities. **RESULTS** The *matK* sequences of six samples of *Pogostemon cablin* from different regions of cultivation are 1 245 bp in length, which coding 415 amino acids of protein (maturase), and 18S rRNA sequences are 1 803 ~ 1 805 bp in size. Based on multiple sequence alignment, there are 47 variable sites in the *matK* sequence of these six samples, 17 in the 18S rRNA sequence. The cluster tree reconstructed by UPGMA method shows that the sequence divergence both in *matK* and 18S rRNA genes among six samples of *Pogostemon cablin* was well correlative with their regions of cultivation and intraspecific chemotypes of essential oil composition. **CONCLUSION** Combining with chemical and biogeographical data, DNA sequencing can become a powerful tool in the key technique-species identification of quality evaluation and GAP production of *Pogostemon cablin*.

KEY WORDS: *Pogostemon cablin* (Blanco) Benth.; DNA sequencing; species identification