

## 基于叶绿体基因组的长梗秦艽、全萼秦艽物种DNA分子标记与鉴定

董博然<sup>1</sup>, 赵志礼<sup>1\*</sup>, 倪梁红<sup>1</sup>, 嘎 务<sup>2</sup>, 刘铜华<sup>2,3</sup>

(1. 上海中医药大学, 上海 201203; 2. 西藏藏医药大学, 西藏 拉萨 850000;  
3. 北京中医药大学, 北京 100029)

**摘要:** 藏药“解吉”为多来源品种之一, 基原涉及龙胆科龙胆属 (*Gentiana*) 秦艽组 (Sect. *Cruciata*) 长梗秦艽 *Gentiana waltonii* Burk. 及全萼秦艽 *Gentiana lhassica* Burk. 等多种植物。在课题组前期民族植物学考察及品种整理基础上, 本文首次分别测定两种基原植物的叶绿体全基因组序列, 进而基于我国青藏高原产秦艽组 10 个物种的叶绿体全基因组数据筛选物种鉴定 DNA 分子标记, 结果如下: ① 长梗秦艽叶绿体全基因组长度为 148 705 bp, 大单拷贝区 (LSC)、小单拷贝区 (SSC) 分别为 81 068 bp 和 17 029 bp, 反向重复区 (IR) 为 25 304 bp; ② 全萼秦艽全基因组长度为 148 652 bp, 大单拷贝区 (LSC)、小单拷贝区 (SSC) 分别为 80 997 bp 和 17 051 bp, 反向重复区 (IR) 为 25 302 bp; ③ 分别注释叶绿体基因 112 个, 其中 78 个编码蛋白基因 (CDS: coding sequence)、30 个 tRNA 基因、4 个 rRNA 基因以及 2 个假基因 (pseudogene: *ψrps16*, *ψinfA*); ④ 构建的特异性引物 PCR 两步鉴定法可在秦艽组内将长梗秦艽、全萼秦艽有效鉴别。本工作可为长梗秦艽及全萼秦艽物种 DNA 分子鉴定、藏药“解吉”品质评价、高山濒危物种种质资源保护及龙胆科龙胆属系统发育分析等工作提供基础科学资料。

**关键词:** 藏药; 长梗秦艽; 全萼秦艽; 叶绿体基因组; 分子标记; PCR 鉴定法

中图分类号: R931 文献标识码: A 文章编号: 0513-4870(2021)09-2584-08

## Molecular markers based upon whole chloroplast genomes and identifying alpine *Gentiana waltonii* and *G. lhassica* (Gentianaceae)

DONG Bo-ran<sup>1</sup>, ZHAO Zhi-li<sup>1\*</sup>, NI Liang-hong<sup>1</sup>, GAAWE Dorje<sup>2</sup>, LIU Tong-hua<sup>2,3</sup>

(1. Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China; 2. Tibetan Traditional Medical College, Lhasa 850000, China; 3. Beijing University of Chinese Medicine, Beijing 100029, China)

**Abstract:** As two original plants of Tibetan herb Jieji, *Gentiana waltonii* Burk. and *Gentiana lhassica* Burk. belong to Section *Cruciata* of *Gentiana*, Gentianaceae. Here, we report on whole chloroplast genome sequences in the alpine species, respectively, and the features of plastomes were investigated. The plastome of *G. waltonii* is 148 705 bp long (148 652 bp in *G. lhassica*) and encodes 112 genes, including 78 protein-coding genes, 30 transfer RNA genes, and 4 ribosomal RNA genes. Two pseudogenes, namely *ψrps16* and *ψinfA*, were found in plastomes. In addition, two novel loci were detected, and a species-specific polymerase chain reaction assay was developed for differentiating *G. waltonii* and *G. lhassica* from 10 alpine species in Section *Cruciata. Gentiana*. Our study provides basic data for identifying Tibetan herbs, alpine species conservation and molecular phylogenetic studies of *Gentiana* and Gentianaceae.

**Key words:** Tibetan herb; *Gentiana waltonii*; *Gentiana lhassica*; plastome; molecular loci; species-specific PCR assay

收稿日期: 2021-04-12; 修回日期: 2021-05-06.

基金项目: 国家自然科学基金面上项目 (82073959); 国家自然科学基金面上项目 (81173654); 藏医药区域协同创新中心项目 (2018XTCX005).

\*通讯作者 Tel: 86-21-51322202, E-mail: zhilzhao@sohu.com

DOI: 10.16438/j.0513-4870.2021-0527

在长期的生活与医疗活动中,藏族人民不断认识身边高山植物、动物及矿物的药用价值,积累了丰富的实践经验,逐步形成了独具特色的藏医药学。藏医药学是祖国民族医药学的重要组成部分,具有丰富的科学与藏族文化内涵。深入开展藏药品种整理,可为藏药材客观、准确的鉴定,活性物质研究及质量标准的制定与完善等工作打下坚实基础,具有重要的科学意义。但由于种种原因,藏药品种整理等工作尚不能适应藏医药学发展的需要,当务之急是开展相关深入的基础性研究工作。

藏药解吉(ཉེ་ཇེ།),常用于治疗“赤巴病”<sup>[1,2]</sup>。品种及植物来源多样,传统上有解吉那保(ཉེ་ཇེ་ནག་པོ།)与解吉嘎保(ཉེ་ཇེ་དཀར་པོ།)两个品种之分;据文献报道、课题组实地考察与品种整理,其来源可涉及龙胆科Gentianaceae龙胆属Gentiana秦艽组(Sect. *Cruciata*)长梗秦艽 *Gentiana waltonii* Burk.、全萼秦艽 *Gentiana lhasica* Burk.等多种植物<sup>[3,4]</sup>。

长梗秦艽 *G. waltonii* Burk.,我国高山特有种,仅在西藏自治区有分布,生于海拔3 000~4 800 m的高山草坡、山坡砾石及林下<sup>[5]</sup>。全萼秦艽 *G. lhasica* Burk.,我国高山特有种,在西藏自治区及青海省有分布,生于海拔4 200~5 050 m的高山草甸<sup>[5]</sup>。秦艽组植物的花色常分为两类,一类常为蓝至蓝紫色,一类则为灰白色或浅黄至黄绿色,这也是将解吉分为解吉那保与解吉嘎保两个品种的主要依据。长梗秦艽与全萼秦艽的花同为蓝紫色,花均具花梗,排列成疏松的花序。尽管植物地上部分形态特征有一定区别,但对于根部药材鉴定而言,传统方法有其局限性。同时,作为高山药用植物,生存环境严酷,加之人为采挖,野生资源状况不容乐观。课题组前期开展了长梗秦艽及全萼秦艽的初步生药学研究<sup>[6]</sup>;但有关上述两种高山特有种的遗传背景研究及深入的种质资源评价等工作,目前尚未见系统报道。

绿色植物的叶绿体拥有自身完整的一套基因组,随着烟草 *Nicotiana tabacum*<sup>[7]</sup>和地钱 *Marchantia polymorpha*<sup>[8]</sup>等叶绿体全基因组序列的发表,植物叶绿体基因结构逐渐被人们所熟知。大多数被子植物的叶绿体基因结构为双链环状<sup>[9,10]</sup>,包括大单拷贝区(LSC)、小单拷贝区(SSC)以及两个反向重复区(IR)。其大小、结构和基因种类一般较为保守<sup>[11,12]</sup>,编码区和非编码区进化速率与模式具有物种鉴别意义<sup>[13]</sup>,在研究植物系统进化、物种遗传背景分析、物种DNA分子标记筛选等方面具有其独特优势<sup>[14,15]</sup>。

为此,在民族植物学考察、标本采集与分类学鉴定等工作基础<sup>[4,16,17]</sup>上,进行长梗秦艽及全萼秦艽叶绿体全基因组序列测定及遗传背景分析,筛选物种DNA分子标记、构建特异性PCR鉴定方法,以期为长梗秦艽及全萼秦艽物种DNA分子鉴定、藏药“解吉”品质评价、高山濒危物种种质资源保护及龙胆科龙胆属系统发育分析等工作提供基础科学资料。

## 材料与方法

**样品采集及分类学鉴定** 叶绿体基因组测序样品,长梗秦艽 *G. waltonii* 采集日期2018-08-08,西藏拉萨,高山草坡,海拔4 364 m,样点地理坐标N 29°46.132', E 91°10.148',凭证标本:董博然等XZ2018CG001(4)。全萼秦艽 *G. lhasica* 采集日期2018-08-10,西藏米拉山,高寒草甸,海拔5 018 m,样点地理坐标N 29°49.596', E 92°20.829',凭证标本:赵志礼,倪梁红Z2018XZ003(5)。同时,采集龙胆属秦艽组其他8个近缘种标本,每个居群至少取3株样品(2016XZ001除外),总计10种22个居群91份。

去英国皇家植物园(Kew)植物标本馆查阅长梗秦艽 *G. waltonii*、全萼秦艽 *G. lhasica* 及其近缘物种的模式标本,以确保物种分类学鉴定的准确性。凭证标本存放于上海中医药大学中药学院药用植物标本室(表1,图1)



**Figure 1** *Gentiana lhasica* and *Gentiana waltonii*. A: Habit of *G. lhasica*; B: Type of *G. lhasica* in Kew! C: Habit of *G. waltonii*; D: Lectotype of *G. waltonii* in Kew!

**Table 1** Samples and voucher specimens

Species	Pop. ID	Voucher specimens	Locality	Indiv. No.
<i>G. waltonii</i>	A	Xiong B et al XZ201302	Mozhugongka, Tibet	5
	B	Xiong B et al XZ201317	Lhasa, Tibet	5
	C	Lu JN et al J2014081803	Mozhugongka, Tibet	7
	E	Lu JN et al J2014081901	Lhasa, Tibet	3
	F	Zhao ZL and Ni NH 2016XZ001	Langkazi, Tibet	1
	I	Dong BR et al XZ2018CG001	Lhasa, Tibet	5
	K	Zhao ZL and Ni NH Z2018XZ001	Mozhugongka, Tibet	5
<i>G. lhasica</i>	L	Xiong B et al XZ201303	Mt. Mila, Tibet	5
	N	Xiong B et al XZ201304	Mt. Mila, Tibet	9
	O	Xiong B et al XZ201319	Lhasa, Tibet	5
	P	Lu JN et al J2014081802	Mt. Mila, Tibet	3
	Q	Lu JN et al J2014081804	Mt. Mila, Tibet	3
	R	Zhao ZL et Ni NH 2017XZ011	Mt. Mila, Tibet	6
	S	Zhao ZL et Ni NH Z2018XZ003	Mt. Mila, Tibet	5
<i>G. straminea</i>	M	Zhao ZL et Ni NH 2016XZ010	Changdu, Tibet	3
<i>G. crassicaulis</i>	J	Zhao ZL et Ni NH 2016XZ011	Changdu, Tibet	3
<i>G. robusta</i>	Z	Xiong B et al XZ201301	Mozhugongka, Tibet	3
<i>G. macrophylla</i>	D	Zhao ZL et Ni NH XZ201308	Linzi, Tibet (cultivated)	3
<i>G. officinalis</i>	H	Xiong B et al QH201306	Henan, Qinghai Province	3
<i>G. siphonantha</i>	G	Xiong B et al QH201302	Zeku, Qinghai Province	3
<i>G. dahurica</i>	X	Ni NH et al GS201304	Hezuo, Gansu Province	3
<i>G. tibetica</i>	T	Zhao ZL et Ni NH 2016XZ004	Yadong, Tibet	3

**总 DNA 提取** 硅胶快速干燥叶片, 液氮中研磨, 改良 CTAB 法<sup>[18]</sup>提取总 DNA。

**叶绿体全基因组序列的获得及物理图谱绘制** 参考被子植物叶绿体全基因组测序通用引物<sup>[19]</sup>, Sanger 法测序<sup>[20]</sup>。其中通用引物无法扩出的片段以及 gaps, 参考相关序列<sup>[16,21,22]</sup>基于 Primer Select (Lasergene 7.0) 软件, 重新设计引物, 其中长梗秦艽 21 对, 全萼秦艽 7 对。扩增出叶绿体基因组各片段序列, 应用 SeqMan 软件进行正反向拼接。针对叶绿体基因组结构的特征, 设计 IR 区与 LSC 区、SSC 区的 4 个接口位置的特异性引物, 经 PCR 扩增、测序后再次进行序列准确性验证。所用引物序列及验证区域见表 2。

**Table 2** Primers for validation of the boundary regions

Primer ID	Sequence (5' to 3')	Boundary region
SSC-a1	TTACTTTGGGGGTTTTAGGATA	SSC/IRa
SSC-a2	GGTAGCCAGATTAGTCAACATTTA	
LSC-b1	ACTTGCCTTGACTTCTGCTTTACT	LSC/IRb
LSC-b2	GGGGGAGGTCCGTTTGATA	
b-SSC1	ATGGCTTGGATTGGTATTA	IRb/SSC
b-SSC2	GAAAAAGGGAGGGTCTATT	
a-LSC1	ACAGGTGGGGAATCTTGAGGTGAA	IRa/LSC
a-LSC2	CCGTGCTAATCTTGGTATGGA	

叶绿体基因组在线注释软件 DOGMA<sup>[23]</sup>对样品叶绿体全基因组进行初步注释; 参考近缘种序列<sup>[16,21,22]</sup>, 结合编码蛋白基因起始密码子、终止密码子, 进行 BLAST<sup>[24]</sup>手工校正; tRNA 基因利用在线软件 tRNAscan-SE 2.0<sup>[25]</sup>

进行验证。将注释完成的基因组序列上传至 NCBI, 应用在线绘图软件 OGDRAW<sup>[26]</sup>, 分别绘制出长梗秦艽及全萼秦艽叶绿体基因组物理图谱。

**基于龙胆属秦艽组叶绿体全基因组结构分析的物种 DNA 分子标记筛选** 基于秦艽组已报道的叶绿体全基因组序列, 应用 MEGA<sup>[27]</sup>软件分析长梗秦艽、全萼秦艽的基因结构差异, 进而序列对比筛选分子标记位点; MrBayes 软件<sup>[28]</sup>, 设置 GTR+I+G 模型, 以夹竹桃科 Apocynaceae 植物长春花 *Catharanthus roseus* 为外类群, 构建龙胆科 21 个分类群的贝叶斯系统进化树。DnaSP 软件计算秦艽组 10 个物种叶绿体全基因组的核苷酸多样性 (Pi), 以寻找潜在的分子标记位点, Pi 值设置: > 0.02。

**基于特异性引物的长梗秦艽与全萼秦艽 DNA 分子鉴定** 对初筛的分子标记进行进一步分析, 筛选原则: ① 差异片段尽可能长; ② 尽可能避免重复序列。基于分析选定的 2 个分子标记, 设计物种特异性鉴定引物, 优化 PCR 扩增程序, 通过琼脂糖凝胶电泳对目的条带长度进行比较。先选取秦艽组 10 个物种 14 个居群, 每个居群单株样品共 14 株, 对 DNA 分子标记进行初步验证; 在此基础上, 扩大样本量至龙胆属秦艽组 10 种 22 个居群 91 份样品, 再次验证分子标记的专属性, 最终建立鉴别方法。

## 结果与分析

### 1 长梗秦艽、全萼秦艽叶绿体基因组序列结构

IR 区与 LSC 区、SSC 区的 4 个接口位置验证实验

结果表明, 拼接所得叶绿体基因组序列相应接口位置均准确无误, 最终获得长梗秦艽、全萼秦艽叶绿体全基因组序列。长梗秦艽序列总长 148 705 bp (MK780032), 包括两个反向重复区域 (IRa、IRb: 25 304 bp)、一个大单拷贝区 (LSC: 81 068 bp) 和一个小单拷贝区 (SSC: 17 029 bp); 全萼秦艽序列总长 148 652 bp (MK790135), 包括两个反向重复区域 (IRa、IRb: 25 302 bp)、一个大单拷贝区 (LSC: 80 997 bp) 和一个小单拷贝区 (SSC: 17 051 bp) (表 3)。

**Table 3** Chloroplast genome structure in *G. waltonii* and *G. lhassica*

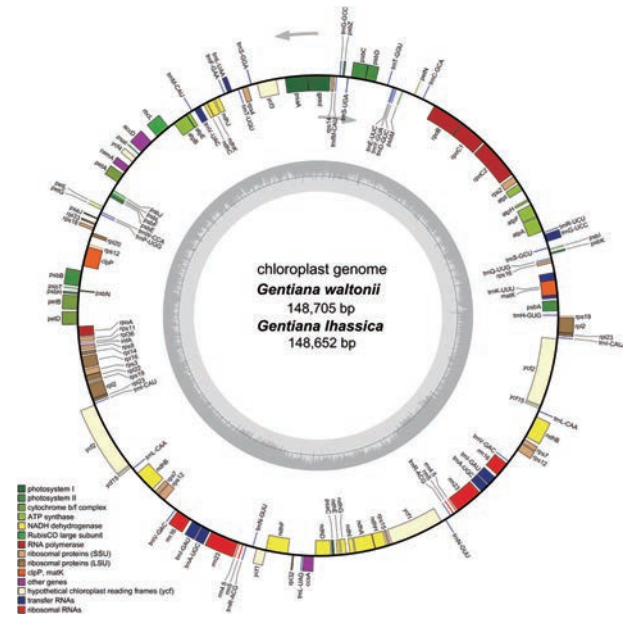
Species	GenBank number	Total /bp	LSC /bp	SSC /bp	IR /bp	G+C /%
<i>G. waltonii</i>	MK780032	148 705	81 068	17 029	25 304	37.73
<i>G. lhassica</i>	MK790135	148 652	80 997	17 051	25 302	37.73

两个物种的基因种类及数量相同, 包括 78 个编码蛋白的基因 (CDS: coding sequence)、30 个 tRNA 基因、4 个 rRNA 基因, 以及 2 个假基因 (pseudogene: *ψrps16*, *ψinfA*) (表 4)。基因组物理图谱见图 2。

## 2 叶绿体基因结构对比及系统发育关系分析

**2.1 序列差异分析** 以长梗秦艽为参照, 利用 MEGA 软件将其余秦艽组 9 种植物的叶绿体基因组序列进行差异位点的人工比对, 差异类型分为 4 种: 插入、缺失、碱基转换、碱基颠换。结果显示, 全萼秦艽与长梗秦艽间序列差异性最小, 二者的差异主要是碱基替换; 而与黄管秦艽 *G. officinalis* 差异性最大 (表 5)。

**2.2 DnaSP 高变区筛选** DnaSP 软件计算秦艽组 10 个物种叶绿体全基因组的 Pi (Pi 值设置: > 0.02)。多态性分布的热点区域常位于大单拷贝区和小单拷贝区, 而 IR 区相对保守。叶绿体基因中的 IGS (*trnH-psbA*)、



**Figure 2** Chloroplast genome maps of *G. waltonii* and *G. lhassica*. Genes inside the circle are transcribed clockwise, whereas those outside the circle counterclockwise. Dashed area in the inner circle indicates the G+C content of the genome

IGS (*trnS-trnG*)、IGS (*ndhC-trnV*)、*ycf1*、*rpl2* 基因, 可以作为秦艽组内潜在的分子标记。

**2.3 分子系统学分析** 基于龙胆科龙胆属 *Gentiana* 秦艽组 (Sect. *Cruciata*) 10 个物种、多枝组 (Sect. *Kudoa*) 8 个分类群及獐牙菜属 *Swertia* 3 个物种的叶绿体基因组序列, 以夹竹桃科植物长春花 *Catharanthus roseus* 为外类群, 构建贝叶斯系统进化树 (图 3)。

进化树中各类群间的拓扑结构与经典分类学观点基本相一致: 首先外类群夹竹桃科长春花被分出, 而另一大支为龙胆科各类群。进而龙胆科獐牙菜属 *Swertia* 和龙胆属 *Gentiana* 分成两支; 其中龙胆属一支除短柄龙

**Table 4** Chloroplast genes in *G. waltonii* and *G. lhassica*. \*: Genes containing introns; (×2): Two gene copies in the IRs; *ψ*: Pseudogenes

No.	Group of genes	Gene names
1	Photosystem I	<i>psaB</i> , <i>psaA</i> , <i>psaI</i> , <i>psaJ</i> , <i>psaC</i>
2	Photosystem II	<i>psbA</i> , <i>psbK</i> , <i>psbI</i> , <i>psbM</i> , <i>psbD</i> , <i>psbC</i> , <i>psbZ</i> , <i>psbJ</i> , <i>psbL</i> , <i>psbF</i> , <i>psbE</i> , <i>psbB</i> , <i>psbT</i> , <i>psbN</i> , <i>psbH</i>
3	Cytochrome b/f complex	<i>petN</i> , <i>petA</i> , <i>petL</i> , <i>petG</i> , <i>petB*</i> , <i>petD*</i>
4	ATP synthase	<i>atpA</i> , <i>atpF*</i> , <i>atpH</i> , <i>atpI</i> , <i>atpE</i> , <i>atpB</i>
5	NADH dehydrogenase	<i>ndhA*</i> , <i>ndhB*</i> (×2), <i>ndhC</i> , <i>ndhD</i> , <i>ndhE</i> , <i>ndhF</i> , <i>ndhG</i> , <i>ndhH</i> , <i>ndhI</i> , <i>ndhJ</i> , <i>ndhK</i>
6	RubisCO large subunit	<i>rbcL</i>
7	RNA polymerase	<i>rpoC2</i> , <i>rpoC1*</i> , <i>rpoB</i> , <i>rpoA</i>
8	Ribosomal proteins (SSU)	<i>ψrps16</i> , <i>rps14</i> , <i>rps4</i> , <i>rps18</i> , <i>rps2</i> , <i>rps11</i> , <i>rps8</i> , <i>rps3</i> , <i>rps19</i> , <i>rps7</i> (×2), <i>rps12*</i> (×2), <i>rps15</i>
9	Ribosomal proteins (LSU)	<i>rpl2*</i> (×2), <i>rpl14</i> , <i>rpl16*</i> , <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> (×2), <i>rpl32</i> , <i>rpl33</i> , <i>rpl36</i>
10	Other genes	<i>clpP*</i> , <i>matK</i> , <i>accD</i> , <i>ccsA</i> , <i>ψinfA</i> , <i>cemA</i>
11	Proteins of unknown function	<i>ycf1</i> , <i>ycf2</i> (×2), <i>ycf3*</i> , <i>ycf4</i> , <i>ycf15</i> (×2)
12	Ribosomal RNAs	<i>rrn4.5</i> (×2), <i>rrn5</i> (×2), <i>rrn16</i> (×2), <i>rrn23</i> (×2)
13	Transfer RNAs	<i>trnH-GUG</i> , <i>trnK-UUU*</i> , <i>trnQ-UUG</i> , <i>trnS-GCU</i> , <i>trnG-UCC*</i> , <i>trnR-UCU</i> , <i>trnC-GCA</i> , <i>trnD-GUC</i> , <i>trnY-GUA</i> , <i>trnE-UUC</i> , <i>trnT-GGU</i> , <i>trnS-UGA</i> , <i>trnG-GCC</i> , <i>trnJ-CAU</i> , <i>trnS-GGA</i> , <i>trnT-UGU</i> , <i>trnL-UAA*</i> , <i>trnF-GAA</i> , <i>trnV-UAC*</i> , <i>trnM-CAU</i> , <i>trnW-CCA</i> , <i>trnP-UGG</i> , <i>trnI-CAU*</i> (×2), <i>trnL-CAA</i> (×2), <i>trnV-GAC</i> (×2), <i>trnI-GAU</i> (×2), <i>trnA-UGC*</i> (×2), <i>trnR-ACG</i> (×2), <i>trnN-GUU</i> (×2), <i>trnL-UAG</i>

**Table 5** SNPs and InDels in plastome sequences within section *Cruciata* spp. SNP Tr: SNP transition; SNP Trv: SNP transversion

No.	Species	Deletions	Insertions	SNP Tr.	SNP Trv.	Total
1	<i>G. waltonii</i>	-	-	-	-	-
2	<i>G. lhassica</i>	15	6	29	26	76
3	<i>G. tibetica</i>	33	26	101	53	213
4	<i>G. straminea</i>	53	54	208	138	453
5	<i>G. robusta</i>	58	46	209	139	452
6	<i>G. crassicaulis</i>	29	29	99	54	211
7	<i>G. officinalis</i>	48	46	236	247	577
8	<i>G. siphonantha</i>	52	50	238	147	487
9	<i>G. dahurica</i>	51	56	249	141	497
10	<i>G. macrophylla</i>	21	82	256	170	529

胆 *G. stipitata* 的归属外, 其他各类群又被分为秦艽组 (Sect. *Cruciata*) 与多枝组 (Sect. *Kudoa*) 两小支。同时, 长梗秦艽和全萼秦艽显示了更为接近的亲缘关系。

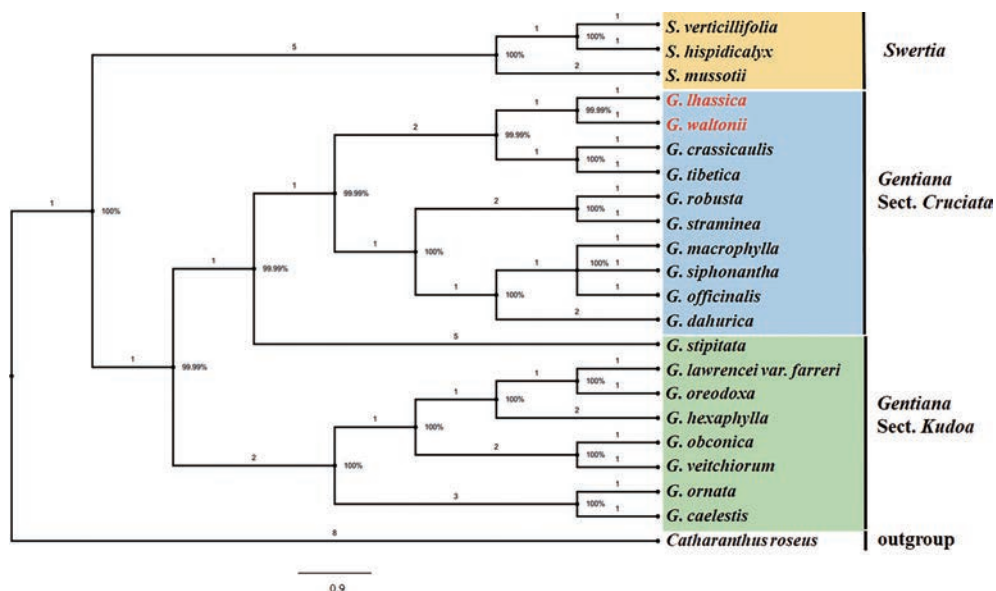
**3 基于叶绿体基因组的长梗秦艽、全萼秦艽物种分子标记及其鉴定**

秦艽组组内 10 个物种叶绿体基因组序列对比分

析显示, 长梗秦艽与全萼秦艽两个物种序列的同源性、相似度极高, 秦艽组内难于直接将这两个物种进行区分。两步法鉴定策略: ① 设计 DNA 特异性引物, 将长梗秦艽、全萼秦艽归为一类, 与其他 8 个近缘种有效区分。② 再设计特异性引物实现两个近缘种之间的鉴定。

“第一步”预筛出 5 个片段 (表 6); “第二步”预筛出 3 个片段 (表 7)。从中再复筛并各确定 1 个片段: IGS (*trnT*-UGU, *trnL*-UAA) 间隔区一个特异性“75 bp deletion/insertion”以及 *ycf1* 基因中的一个特异性“21 bp deletion/insertion”。进而扩大样本量至秦艽组 10 种 91 个个体进行验证实验, 其中, *G. waltonii* 7 居群 31 个个体, *G. lhassica* 7 居群 36 个个体。由此构建如下鉴定方法:

第一步: 引物 (Primer 2-8) 5'-CGACCGTTCAAGT ATGCAAAAGGTT 及 5'-TAGTTTATTGTGAGGTCG CTCCG; PCR 程序为 94 °C, 5 min; 42× (94 °C, 30 s; 60 °C, 40 s; 72 °C, 1~3 min); 72 °C, 7 min。凝胶电泳条件: 2.0% 琼脂糖凝胶, 120 V, 45 min, 65 mA。长梗



**Figure 3** Bayesian phylogenetic analysis based on 21 plastomes within Gentianaceae, with *Catharanthus roseus* as the outgroup. Numbers at nodes are posterior probabilities

**Table 6** Candidate loci for identifying *G. waltonii* and *G. lhassica* within section *Cruciata* spp

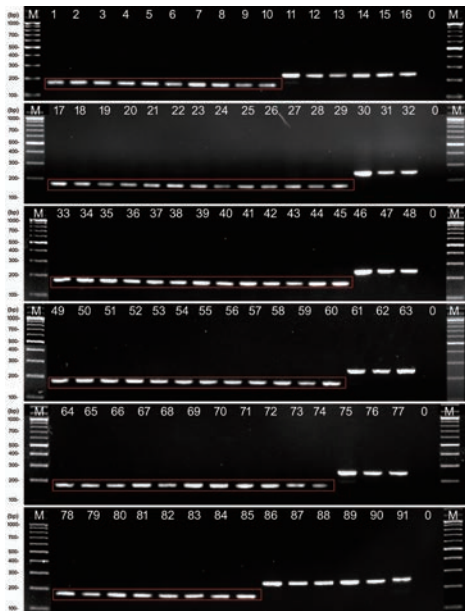
Taxon	IGS ( <i>trnQ</i> -UUG, <i>psbK</i> )	IGS ( <i>trnT</i> -GGU, <i>psbD</i> )	IGS ( <i>trnT</i> -UGU, <i>trnL</i> -UAA)	IGS ( <i>ndhC</i> , <i>trnV</i> -UAC)	IGS ( <i>atpB</i> , <i>rbcL</i> )
<i>G. waltonii</i>	33 bp (Del)	7 bp (In)	75 bp (Del)	6 bp (In)	11 bp (Del)
<i>G. lhassica</i>	+	+	+	+	+
<i>G. tibetica</i>	-	-	-	-	-
<i>G. straminea</i>	-	-	-	-	-
<i>G. robusta</i>	-	-	-	-	-
<i>G. crassicaulis</i>	-	-	-	-	-
<i>G. officinalis</i>	-	-	-	-	-
<i>G. siphonantha</i>	-	-	-	-	-
<i>G. dahurica</i>	-	-	-	-	-
<i>G. macrophylla</i>	-	-	-	-	-

**Table 7** Candidate loci for differentiating between *G. waltonii* and *G. lhasica*

Taxon	<i>atpB-rbcL</i>	<i>psaJ-rpl33</i>	<i>ycf1</i>
<i>G. lhasica</i>	11 bp (In)	15 bp (Del)	21 bp (In)
<i>G. waltonii</i>	-	-	-

秦艽与全萼秦艽目的条带长度为 173 bp, 而秦艽组其他 8 种植物为 248 bp, 差距明显 (75 bp), 图 4。

第二步: 引物 (Primer 1-1) 5'-CGGATTGTTGGTAATACCGTACC 及 5'-CCAAATAGGCGGTGGCAAATTCGC; PCR 程序为 94 °C, 5 min; 42× (94 °C, 30 s; 59.9 °C, 40 s; 72 °C, 1~3 min); 72 °C, 7 min。凝胶电泳条件: 2.7% 琼脂糖凝胶, 145 V, 45 min, 106 mA。长梗秦艽目的条带长度为 172 bp, 全萼秦艽为 193 bp, 差距较为明显 (21 bp), 图 5。

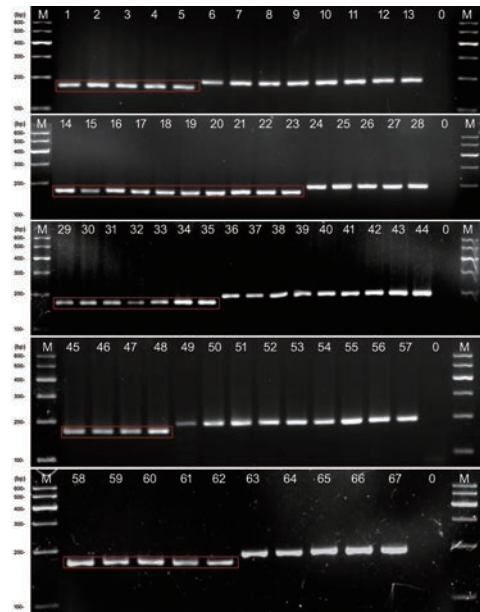


**Figure 4** A group-specific PCR profile of *Gentiana* section *Cruciata* spp. Bands in the box only for *G. waltonii* and *G. lhasica*. M: DNA Ladder Mix; 0: Negative control; *G. waltonii*: 1-5 (A), 17-21 (B), 33-39 (C), 49-51 (E), 64 (F), 65-69 (K), 78-82 (I); *G. lhasica*: 6-10 (L), 22-26 (O), 27-29 (P), 40-45 (R), 52-60 (N), 70-74 (S), 83-85 (Q); 11-13 (J); *G. crassicaulis*; 14-16 (Z); *G. robusta*; 30-32 (M); *G. straminea*; 46-48 (G); *G. siphonantha*; 61-63 (H); *G. officinalis*; 75-77 (T); *G. tibetica*; 86-88 (X); *G. dahurica*; 89-91 (D); *G. macrophylla*; Bracket: Pop. ID

最终视琼脂糖凝胶电泳目的条带的位置, 完成物种鉴别。

## 讨论

植物叶绿体基因组序列含有丰富的植物系统学信息。本文基于龙胆科 21 个分类群的质体基因组相关



**Figure 5** A species-specific PCR profile of *G. waltonii* and *G. lhasica*. Bands in the box only for *G. waltonii*. M: DNA Marker-B; 0: Negative control; *G. waltonii*: 1-5 (A), 14-18 (B), 19-23 (I), 29-35 (C), 45-47 (E), 48 (F), 58-62 (K); *G. lhasica*: 6-10 (L), 11-13 (P), 24-28 (O), 36-44 (N), 49-51 (Q), 52-57 (R), 63-67 (S); Bracket: Pop. ID

资料构建的贝叶斯系统发育树, 拓扑关系与经典分类学观点基本一致。秦艽组内, 长梗秦艽 *G. waltonii* 则与全萼秦艽 *G. lhasica* 互为姐妹种, 后验概率 99.99%, 亲缘关系更为接近; 两者的序列分析进一步显示出高度的同源性。因此, 在构建物种条形码时, 工作较为复杂、困难。本文工作, 筛选出位于 IGS (*trnT-trnL*) 和 *ycf1* 基因中的两个分子标记, 大样本验证, 由此构建长梗秦艽、全萼秦艽的特异性引物 PCR 两步鉴定法, 可在秦艽组 10 个物种间有效鉴定姐妹种: 长梗秦艽与全萼秦艽, 方法稳定, 易于操作, 准确性高。

多态性分布的热点区域分析结果提示: 叶绿体基因中的 IGS (*trnH-psbA*)、IGS (*trnS-trnG*)、IGS (*ndhC-trnV*) 及 *ycf1*、*rpl2* 基因, 可以作为秦艽组组内潜在的分子标记。

全萼秦艽基因组内在 IGS (*rps16-trnQ*) 中出现了两个基因型。I型: 该片段序列完整, II型: 出现一 47 bp 的碱基缺失, 两者比例为 15:21。该基因分型的遗传学意义及是否与植株生长环境和产地等影响因子有关, 值得深入研究。

多枝组 (Sect. *Kudoa*) 短柄龙胆 *G. stipitata* 在系统树中被聚入秦艽组 (Sect. *Cruciata*) 一支中, 与经典分类学处理有所不同。龙胆属物种间自然杂交现象多有报道<sup>[29,30]</sup>, 短柄龙胆是否为一杂交起源物种或两近缘物

种同域分布个体间的杂交类型? 需要更多的证据支持。

综上所述, 叶绿体基因组序列与结构特性在龙胆科及龙胆属系统发育研究及构建物种DNA条形码方面同样具有重要的意义。龙胆属是龙胆科的第一大属, 国产近250种<sup>[5]</sup>, 药用价值高。但目前已报道的质体基因组相关资料有限, 有必要广泛深入的开展相关序列测定及分析工作。

每个物种均有其独特的形成历史及复杂的生物学特性, 因此, 具体的鉴定对象, 应具体分析其遗传背景, 构建个性化的DNA条形码。面对丰富多样的植物类群及中药基原植物, 探究并构建准确、高效的DNA条形码鉴定体系还有很长的路要走。

**作者贡献:** 董博然是本文的第一作者, 负责分子生物学实验及稿件撰写; 倪梁红负责样品采集工作; 嘎务负责民族植物学考察工作; 刘铜华参与研究工作总体设计; 赵志礼为本文的通讯作者, 负责样品分类学鉴定, 研究工作总体设计及稿件修改等工作。

**利益冲突:** 本文的研究内容无任何利益冲突。

## References

- [1] Yutok Yonten Gonpo. The Four Medical Tantras (四部医典) [M]. Shanghai: Shanghai Science & Technology Press, 1987: 42.
- [2] Dimaer Danzeng Pengcuo. Jing Zhu Ben Cao (晶珠本草) [M]. Shanghai: Shanghai Science & Technology Press, 1986: 119.
- [3] Northwest Institute of Plateau Biology, Chinese Academy of Sciences. Tibetan Medicine (藏药志) [M]. Xining: Qinghai People Press, 1991: 9-12.
- [4] Zhao ZL, Dorje Gaawe, Wang ZT. Identification of medicinal plants used as Tibetan Traditional Medicine Jie-Ji [J]. J Ethnopharmacol, 2010, 132: 122-126.
- [5] Ho TN. Gentianaceae in Flora Reipublicae Popularis Sinicae: Vol 62 (中国植物志: 62卷) [M]. Beijing: Science Press, 1988: 14-64.
- [6] Gu CY, Zhao ZL, Wu JR, et al. Pharmacognostical study on *Gentiana lhasica* Burk. and *Gentiana waltonii* Burk. [J]. Chin J Inf Tradit Chin Med (中国中医药信息杂志), 2010, 17: 41-43.
- [7] Shinozaki K, Ohme M, Tanaka M, et al. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression [J]. EMBO J, 1986, 5: 2043-2049.
- [8] Ohyama K, Fukuzawa H, Kohchi T, et al. Chloroplast gene organization deduced from complete sequence of Liverwort *Marchantia polymorpha* chloroplast DNA [J]. Nature, 1986, 322: 572-574.
- [9] Palmer JD, Stein DB. Conservation of chloroplast genome structure among vascular plants [J]. Curr Genet, 1986, 10: 823-833.
- [10] Palmer JD. Plastid chromosomes: structure and evolution. In: Vasil IK, Bogorad L (eds) Cell Culture and Somatic Cell Genetics in Plants, the Molecular Biology of Plastids, 7A[M]. San Diego, Academic Press, 1991: 5-53.
- [11] Clegg MT, Gaut BS, Learn GH, et al. Rates and patterns of chloroplast DNA evolution [J]. Proc Natl Acad Sci U S A, 1994, 91: 6795-6801.
- [12] Daniell H, Lin CS, Yu M, et al. Chloroplast genomes: diversity, evolution, and applications in genetic engineering [J]. Genome Biol, 2016, 17: 134.
- [13] Sabater B. Evolution and function of the chloroplast. current investigations and perspectives [J]. Int J Mol Sci, 2018, 19: 3095.
- [14] Zhang YJ, Li DZ. Advances in phylogenomics based on complete chloroplast genomes [J]. Plant Divers Resour (植物分类与资源学报), 2011, 33: 365-375.
- [15] Ni LH, Zhao ZL, Mi M. Advances in chloroplast genome of medicinal plants [J]. J Chin Med Mater (中药材), 2015, 38: 1990-1994.
- [16] Ni LH, Zhao ZL. A morphometric comparison of three closely related species of *Gentiana* (Gentianaceae), endemic to the region of the Qinghai-Tibet Plateau [J]. Botany, 2018, 96: 209-215.
- [17] Zong FF, Zhao ZL, Ni LH, et al. DNA fingerprinting and genetic diversity in a threatened alpine medicinal plant, *Gentiana crassicaulis* (Gentianaceae) using AFLP [J]. Acta Pharm Sin (药学报), 2020, 55: 1941-1950.
- [18] Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf material [J]. Phytochem Bull, 1987, 19: 11-15.
- [19] Qian J, Song J, Gao H, et al. The complete chloroplast genome sequence of the medicinal plant *Salvia miltiorrhiza* [J]. PLoS One, 2013, 8: e57607.
- [20] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors [J]. Proc Natl Acad Sci U S A, 1977, 74: 5463-5467.
- [21] Ni LH, Zhao ZL, Xu HX, et al. The complete chloroplast genome of *Gentiana straminea* (Gentianaceae), an endemic species to the Sino-Himalayan subregion [J]. Gene, 2015, 577: 281-288.
- [22] Zhou T, Wang J, Jia Y, et al. Comparative chloroplast genome analyses of species in *Gentiana* section *Cruciata* (Gentianaceae) and the development of authentication markers [J]. Int J Mol Sci, 2018, 19: 1962.
- [23] Wyman SK, Jansen RK, Boore JL. Automatic annotation of organellar genomes with DOGMA [J]. Bioinformatics, 2004, 20: 3252-3255.
- [24] Altschul SF, Gish W, Miller W, et al. Basic local alignment search tool (BLAST) [J]. J Mol Biol, 1990, 215: 403-410.
- [25] Goremykin VV, Barbara H, Hirsch-Ernst KI, et al. Analysis of *Acorus calamus* chloroplast genome and its phylogenetic implications [J]. Mol Biol Evol, 2005, 22: 1813-1822.

- [26] Lohse M, Drechsel O, Bock R. Organellar Genome DRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes [J]. *Curr Genet*, 2007, 52: 267-274.
- [27] Tamura K, Peterson D, Peterson N, et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods [J]. *Mol Biol Evol*, 2011, 28: 2731-2739.
- [28] Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models [J]. *Bioinformatics*, 2003, 19: 1572-1574.
- [29] Xiong B, Zhao ZL, Ni LH, et al. DNA-based identification of *Gentiana robusta* and related species [J]. *China J Chin Mater Med* (中国中药杂志), 2015, 40: 4680-4685.
- [30] Li XJ, Wang LY, Yang HL, et al. Confirmation of natural hybrids between *Gentiana straminea* and *G. siphonantha* (Gentianaceae) based on molecular evidence [J]. *Acta Bot Yunnan* (云南植物研究), 2007, 29: 91-97.