

基于叶绿体全基因组序列的双峰法鉴定粗茎秦艽及其近缘物种

高娜娜¹, 倪梁红^{1*}, 赵志礼^{1*}, 嘎 务²

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

摘要: 粗茎秦艽 *Gentiana crassicaulis* Duthie ex Burk. 为中药“秦艽”的基原植物之一, 而同为龙胆属秦艽组的西藏秦艽 *Gentiana tibetica* King ex Hook. f.、粗壮秦艽 *Gentiana robusta* King ex Hook. f. 为其近缘的非国家药典品种。因存在较大的种内形态变异, 粗茎秦艽与西藏秦艽、粗壮秦艽间表现出较高的形态相似性, 且存在一定的同域分布, 为粗茎秦艽的鉴定带来了一定的难度。为有效鉴定中药秦艽国家药典品种粗茎秦艽与其近缘的非药典品种, 本文基于前期测定的粗茎秦艽、粗壮秦艽叶绿体基因组序列, 首次测定西藏秦艽叶绿体基因组序列, 筛选分子标记并构建这三个物种的分子鉴定方法。结果如下: ① 获得西藏秦艽叶绿体基因组序列, 长度为 148 765 bp, 大单拷贝区 (LSC)、小单拷贝区 (SSC) 分别为 81 163 bp 和 17 070 bp, 反向重复区 (IR) 为 25 266 bp; ② 粗茎秦艽、西藏秦艽、粗壮秦艽的叶绿体基因组结构一致; 粗茎秦艽与西藏秦艽的重复序列、SSR 分布情况一致, 两个物种间叶绿体基因组变异位点数仅为 9; ③ 筛选同时对这三个物种有效鉴定的单核苷酸多态性位点 (SNP), 并验证稳定性; ④ 建立 SNP 双峰检测方法, 对各物种及混合样品进行有效鉴定。本工作可为粗茎秦艽及其近缘物种的 DNA 分子鉴定, 以及相关藏药品种的整理等工作提供基础资料。

关键词: 粗茎秦艽; 西藏秦艽; 粗壮秦艽; 叶绿体基因组; SNP 标记; 双峰检测方法

中图分类号: R931 文献标识码: A 文章编号: 0513-4870(2022)08-2520-08

SNP double peak methods through chloroplast genome profiling for identifying *Gentiana crassicaulis* and two closely related species

GAO Na-na¹, NI Liang-hong^{1*}, ZHAO Zhi-li^{1*}, Gaawe Dorje²

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

Abstract: *Gentiana crassicaulis* Duthie ex Burk. is one of the plant sources of Gentianae Macrophyllae Radix (QinJiao). *Gentiana tibetica* King ex Hook. f. and *Gentiana robusta* King ex Hook. f. are relative species of *G. crassicaulis*. Due to the large intraspecific morphological variation, *G. crassicaulis* showed high morphological similarity with *G. tibetica* and *G. robusta*. And the distribution area of the three species overlaps to some extent, which makes it difficult to identify them. On the basis of morphological identification, the method of molecular identification of the three species was constructed in this study based on chloroplast genomes. The chloroplast genome of *Gentiana tibetica* is 148 765bp long, with LSC, SSC and IR 81 163 bp, 17 070 bp and 25 266 bp, respectively. The structure of the three is consistent. The chloroplast genome sequences of *G. tibetica* and *G. crassicaulis* are highly similar, and the number of variable sites is 9 (149 267 bp in total). Diagnostic SNP that could effectively identify the three species was screened and verified, and a dual-peak SNP detection method was established for the effective identification of each species and mixed samples. Our study provides basic data for the molecular identification of *G. crassicaulis* and its related species, and the arrangement of related Tibetan medicine.

收稿日期: 2022-04-26; 修回日期: 2022-05-15.

基金项目: 国家自然科学基金面上项目 (82073959, 81173654).

*通讯作者 Tel: 86-21-51322202, E-mail: nlhtcm@126.com; zhilzhao@sohu.com

DOI: 10.16438/j.0513-4870.2022-0496

Key words: *Gentiana tibetica*; *Gentiana crassicaulis*; *Gentiana robusta*; chloroplast genome; SNP marker; double peak detection

粗茎秦艽 *Gentiana crassicaulis* Duthie ex Burk. 为《中华人民共和国药典》(2020年版) 收录的中药“秦艽”基原植物之一, 来源于龙胆科 Gentianaceae 龙胆属 *Gentiana* 秦艽组 Section *Cruciata*, 根入药, 可祛风湿, 清湿热, 止痹痛, 退虚热^[1,2]。秦艽组植物在青藏高原呈现出丰富的物种多样性^[3], 并由于杂交和多倍化等^[4,5], 秦艽组植物存在明显的种内形态变异, 以及近缘物种间表现出较高的形态相似性。在市场及应用中, 除了药典收录的秦艽4个基原品种外, 存在较多的混伪品。国内学者针对秦艽组植物及秦艽药材的鉴定开展了一系列的研究工作。罗焜等^[6]应用 ITS2 条形码对秦艽药材及其多种混伪品进行鉴定研究; 熊波等^[7]应用 ITS 和 *atpB-rbcL* 等多片段组合对西藏产4种秦艽组植物进行鉴定; 倪梁红等^[8]应用 ITS 与叶绿体多片段组合构建甘肃产6种秦艽组植物的鉴定策略; 陆佳妮等^[9]评价了线粒体 *nad1/b-c* 及 *nad5/d-e* 对秦艽组物种的鉴定意义; Liu 等^[10]测定了包括秦艽组8种在内的龙胆属6组30种植物的 ITS 及 *rbcL* 等序列, 构建的 DNA 条形码可用于秦艽4种基原植物的鉴定。但是, 因存在致同进化不完全, 秦艽组部分物种的 ITS 序列存在杂合, 需要结合克隆测序及基因分型; 同时因遗传背景相近, 常用条形码片段在部分近缘物种, 如粗茎秦艽与西藏秦艽, 麻花秦艽与粗壮秦艽等的鉴定中提供的信息位点有限。因此, 有待于寻找更为便捷、稳定的分子标记。本文以粗茎秦艽为切入点, 探讨秦艽药典品种与近缘的非药典品种的深入、高效的鉴定方法。

粗壮秦艽 *Gentiana robusta* King ex Hook. f.、西藏秦艽 *Gentiana tibetica* King ex Hook. f., 为与粗茎秦艽形态相近的非药典品种, 这三个物种间还存在一定的同域分布, 这些都为物种鉴定带来了一定的难度和复杂性。同时, 据文献^[11]报道, 粗茎秦艽、西藏秦艽等均为藏药“解吉”的来源, 本课题组在西藏产解吉药材市场品调研中注意到, 因形态相近, 在解吉药材的采集、流通及应用中, 不同地区所涉及来源存在多样性, 影响相关藏药品种的进一步标准化研究。

前期, 本课题组在形态分类的基础上^[12], 对粗茎秦艽、西藏秦艽和粗壮秦艽进行了植物形态学测量, 将大量形态学数据与数理统计方法结合, 细化各物种的繁殖器官(花及花序)各部位的形态特征, 获得将这三个近缘物种有效区分的形态学参数, 为其形态鉴定提供了新思路^[13]。但因秦艽的药用部位为根, 形态鉴定方

法对其药材的鉴定具有局限性, 进一步构建相关物种的分子鉴定方法很有必要。

叶绿体普遍存在于绿色植物中, 拥有自身完整的一套基因组^[14]。叶绿体基因组含有大量的遗传信息, 现已成为药用植物物种鉴定及系统进化研究的有力工具^[15,16]。相较于常用 DNA 条形码片段, 叶绿体基因组可提供更为丰富的信息位点, 同时, 被子植物的叶绿体 DNA 一般为母系遗传, 不存在多拷贝杂合现象^[15], 便于序列分析。目前, 秦艽组多个物种已完成了叶绿体基因组测序^[17-21]。基于叶绿体基因组序列可对长梗秦艽、全萼秦艽、达乌里秦艽、管花秦艽等物种进行有效分子鉴定^[19,20]; 同时, 基于叶绿体基因组序列的系统进化分析进一步阐明了龙胆属内秦艽组与多枝组的系统关系^[21]。

本实验新完成西藏秦艽的叶绿体基因组测序, 结合前期已获得的粗茎秦艽、粗壮秦艽叶绿体基因组序列^[8], 通过序列比较, 筛选出可同时对这三个物种进行有效鉴定的 SNP 位点。基于各物种分布区广泛采样的基础上, 进行 SNP 鉴别位点的稳定性验证。同时, 采用 SNP 双(多)峰检测的方法^[22], 对混合样品进行分析。本研究可为这粗茎秦艽及近缘物种的有效分子鉴定、遗传背景分析等工作提供依据。

材料与方法

样品采集与鉴定 采集粗茎秦艽14个居群36个单株, 西藏秦艽12个居群33个单株, 粗壮秦艽13个居群35个单株(表1), 样点覆盖各物种主要分布区。采集时将新鲜叶片用硅胶快速干燥, 得分子生物学实验材料, 备用。凭证标本经赵志礼教授鉴定, 保存于上海中医药大学药用植物标本室。

西藏秦艽叶绿体基因组测序 用改良的 CTAB 法^[23]提取西藏秦艽(标本号: 2015XZ002-A)的总 DNA。参考本课题组方法^[19], 获得西藏秦艽的叶绿体基因组序列, 进行叶绿体基因组的注释及结构分析。

叶绿体基因组比较 对粗茎秦艽、西藏秦艽、粗壮秦艽的叶绿体基因组的结构和 DNA 序列进行比较^[19,24]。分析物种间的变异位点及分布情况, 并筛选可同时三个物种进行有效区分的 SNP 鉴别位点。

SNP 鉴别位点验证 对 SNP 鉴别位点所在片段(*ccsA-ndhD*)设计专属引物(上游引物: 5'-GACCCAA AAGAAACGTGGGC-3', 下游引物: 5'-CACTGTCCG

Table 1 Sources of plant materials

Species	Voucher specimens	Number	Locality	Longitude and latitude	
<i>G. crassicaulis</i>	QH201405	3	Banma, Qinghai	N 32°46.608'; E 100°49.695'	
	GS201301	2	Maqu, Gansu	N 34°03.815'; E 102°04.138'	
	GS201601	3	Maqu, Gansu	N 34°03.815'; E 102°04.138'	
	GZ201601	3	Weining, Guizhou	N 27°06.396'; E 104°07.377'	
	GZ201602	2	Weining, Guizhou	N 27°06.393'; E 104°07.362'	
	SC201601	3	Kangding, Sichuan	N 29°54.618'; E 101°59.538'	
	SC201602	2	Daofu, Sichuan	N 30°29.862'; E 101°29.727'	
	YN201601	3	Lijiang, Yunnan	N 27°00.938'; E 100°09.601'	
	YN201603	2	Lijiang, Yunnan	N 27°25.266'; E 100°25.518'	
	2016XZ008	3	Basu, Tibet	N 30°07.148'; E 97°16.886'	
	2016XZ009	2	Basu, Tibet	N 30°08.997'; E 97°17.618'	
	2016XZ011	2	Changdu, Tibet	N 31°09.173'; E 97°14.025'	
	XZ201214	3	Dingqing, Tibet	N 31°21.458'; E 95°52.533'	
	2012XZ014	3	Dingqing, Tibet	N 31°21.458'; E 95°52.533'	
	<i>G. tibetica</i>	2012XZ004	3	Linshi, Tibet	N 29°09.303'; E 94°12.861'
		2012XZ006	2	Linshi, Tibet	N 29°37.450'; E 94°21.184'
		2012XZ007	3	Linshi, Tibet	N 29°46.809'; E 94°24.584'
		XZ201306	3	Linshi, Tibet	N 29°36.913'; E 94°24.426'
		XZ201307	3	Linshi, Tibet	N 29°09.303'; E 94°12.861'
		XZ201314	3	Linshi, Tibet	N 29°43.313'; E 94°43.743'
XZ201316		3	Linshi, Tibet	N 29°35.739'; E 94°20.763'	
2015XZ002		3	Longzi, Tibet	N 28°17.900'; E 92°45.394'	
2015XZ003		3	Cuona, Tibet	N 27°52.891'; E 91°47.500'	
2016XZ004		3	Yadong, Tibet	N 27°29.006'; E 88°53.822'	
2016XZ005		3	Yadong, Tibet	N 27°31.745'; E 88°58.493'	
2016XZ012-2		1	Gongbujiangda, Tibet	N 30°00.689'; E 92°59.311'	
<i>G. robusta</i>		XZ201201	3	Lhasa, Tibet	N 29°42.952'; E 91°10.394'
		XZ201202	3	Gongbujiangda, Tibet	N 29°52.151'; E 92°34.714'
		XZ201209	3	Gongbujiangda, Tibet	N 29°52.160'; E 92°40.557'
	XZ201211	3	Lhasa, Tibet	N 29°42.538'; E 91°10.256'	
	XZ201301	3	Mozhugongka, Tibet	N 29°41.305'; E 92°06.737'	
	XZ201318	3	Lhasa, Tibet	N 29°42.952'; E 91°10.394'	
	XZ201401	3	Mozhugongka, Tibet	N 29°41.305'; E 92°06.737'	
	XZ201404	1	Mozhugongka, Tibet	N 29°41.305'; E 92°06.737'	
	XZ201405	1	Mozhugongka, Tibet	N 29°41.305'; E 92°06.737'	
	2015XZ005	3	Cuona, Tibet	N 28°00.732'; E 91°57.736'	
	2016XZ002	3	Langkazi, Tibet, Tibet	N 29°02.849'; E 90°23.746'	
	J2014081801	3	Mozhugongka, Tibet	N 29°41.305'; E 92°06.737'	
J2014081902	3	Lhasa, Tibet	N 29°42.952'; E 91°10.394'		

TTGACAGGGTA-3'; 扩增产物长度: 450 bp)。对粗茎秦艽、西藏秦艽、粗壮秦艽所有样本进行 PCR 扩增, 程序为: 94 °C, 5 min; 42× (94 °C, 30 s; 59.9 °C, 40 s; 72 °C, 1 min); 72 °C, 7 min。扩增产物经琼脂糖凝胶电泳检测后送上海生工公司应用 Sanger 法测序, 根据测序结果分析该位点在各物种内的稳定性。

SNP 鉴别位点的混合样品检测 分别将三种秦艽的总 DNA 两两混合及三者混合, 对 *ccsA-ndhD* 片段进行扩增、测序, 观察测序峰图中 SNP 鉴别位点的峰形特征, 分析该位点对混合样品的鉴别能力。并分析样品的 DNA 混合比例与 SNP 峰图的相关性, 应用紫外分光光度法测定粗茎秦艽、西藏秦艽的总 DNA 浓度, 并配制相同浓度; 然后分别选取两者比例为 20:1、10:1、5:1、1:1、1:5、1:10、1:20 的总 DNA 混合溶液, 进行

ccsA-ndhD 片段的扩增及测序, 观察 SNP 峰的变化。

结果与分析

1 西藏秦艽叶绿体基因组结构

西藏秦艽的叶绿体基因组总长 148 765 bp, 呈环状, 共分为 4 个组成部分: 一个大单拷贝区 (LSC, 81 163 bp), 一个小单拷贝区 (SSC, 17 070 bp), 以及两个反向互补重复区 (IRa, IRb, 各 25 266 bp) (图 1)。共含有基因 112 个, 包括蛋白编码基因 78 个、tRNA 基因 30 个、rRNA 基因 4 个, 其中有 18 个基因在 IR 区重复。此外, 存在 *ψrps16* 和 *ψinfA* 两个假基因, *rps16* 缺失第二外显子, *infA* 序列中插入终止密码子。基因组的 GC 含量为 37.71%, 其中 IR 区 (43.37%) 高于 LSC 区 (35.47%) 和 SSC 区 (31.62%), 这与 IR 区的 rRNA 基因

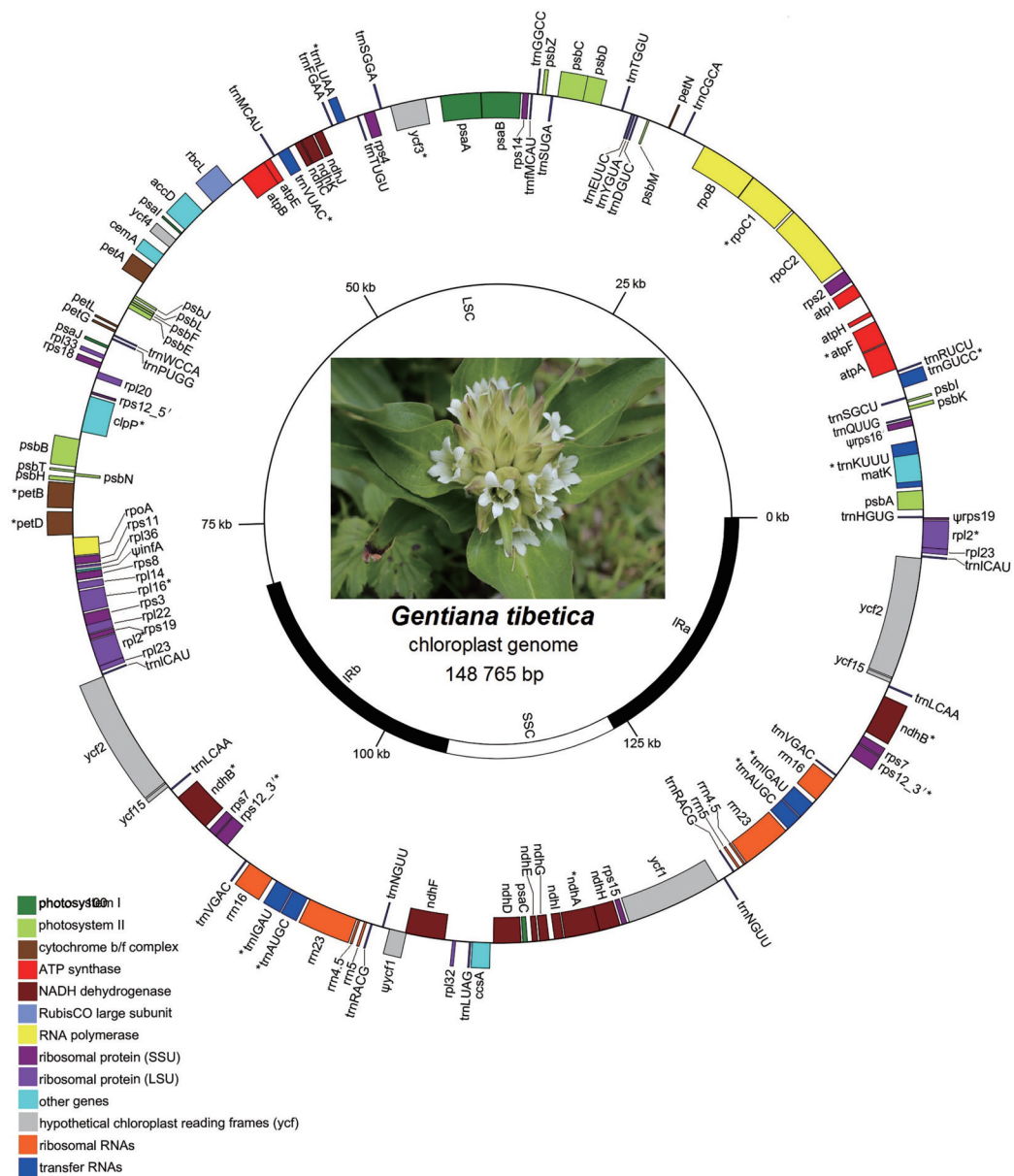


Figure 1 Chloroplast genome map of *Gentiana tibetica*. Genes drawn inside of the map are transcribed in the clockwise direction and genes outside of the map are transcribed in the counter-clockwise direction. Genes belonging to different functional groups are expressed with different colors. Pseudogenes are marked with "ψ" and genes containing introns are marked with "**"

具有较高的GC含量(55.3%)有关^[25]。西藏秦艽叶绿体基因组的结构及基因组成与粗茎秦艽、粗壮秦艽一致^[18], IR区的长度具有较明显差异(表2)。

应用Reputer程序对西藏秦艽叶绿体基因组中的正向重复序列(F, forward repeat sequence)、反向重复序列(R, reverse repeat sequence)、正向互补序列(C, complement repeat)和回文序列(P, palindrome sequence)等4种重复序列进行分析。结果显示,有32对重复序列,包括15对F重复、16对P重复和1对C重复,未检测到R重复。大部分的重复序列位于非编码区,大概有1/4重复序列位于编码区,分别位于*ycf1*、

Table 2 Chloroplast genome features comparison results of *G. tibetica*, *G. crassicaulis* and *G. robusta*

Feature	<i>G. tibetica</i>	<i>G. crassicaulis</i>	<i>G. robusta</i>
GenBank ID	KU975374	KJ676538	KT159969
Total length/bp	148 765	148 776	148 911
GC content/%	37.71	37.72	37.70
LSC length/bp	81 163	81 164	81 164
SSC length/bp	17 070	17 070	17 081
IR length/bp	25 266	25 271	25 333
Total genes	112	112	112
Duplicate genes in IR section	18	18	18
Genes encoding protein	78	78	78
rRNA genes	4	4	4
tRNA genes	30	30	30
Pseudogenes	2	2	2

ycf2、*psaA*、*psaB*、*trnG*-UCC 和 *trnG*-GCC。另有 2 对重复序列跨编码区与基因间隔区排列。应用相同方法检测,粗茎秦艽重复序列的分析结果与西藏秦艽完全一致;粗茎秦艽重复序列的数目、分布情况与其他两个种存在较明显的差异(图2)。

对西藏秦艽叶绿体基因组中的 SSR 进行分析,结果显示,共有 41 个 SSR,其中 27 个单核苷酸 SSR、4 个二核苷酸 SSR、2 个三核苷酸 SSR、7 个四核苷酸 SSR 和 1 个六核苷酸 SSR。单核苷酸 SSR 占比例最大(65.9%),并且均为 A/T 型,这与文献^[26]报道被子植物叶绿体基因组中 SSR 以 A/T 型占主导是相一致的。所有的二核苷酸 SSR 类型都是 AT/TA。其中 31 个 SSR (75.6%) 位于非编码区;10 个 SSR 位于编码基因,包括 *rpoC2*、*rpoC1*、*atpB*、*ycf1* 和 *ndhF*;有 1 个单核苷酸 SSR 跨编码区与基因间隔区排列。应用相同方法检测,粗茎秦艽与西藏秦艽的 SSR 类型及分布完全一致;粗茎秦艽 SSR 的数目、分布情况与其他两个种存在较明显的差异(图3)。

2 叶绿体基因组变异位点分析

将粗茎秦艽、西藏秦艽和粗壮秦艽的叶绿体基因

组 DNA 序列进行全面比较分析。粗壮秦艽与西藏秦艽之间有 415 个变异位点,包括 327 个单核苷酸多态性位点(SNP)和 88 个插入缺失(InDel);粗壮秦艽与粗茎秦艽之间有 407 个变异位点,包括 322 个 SNP 和 85 个 InDel;西藏秦艽与粗茎秦艽序列相似度极高,两者之间仅有 9 个变异位点,包括 6 个 SNP 和 3 个 InDel(图4,表3)。

三个物种间同时存在差异的变异位点仅有 1 个,为位于 *ccsA-ndhD* 片段的第 159 位点,其中粗茎秦艽为 G,西藏秦艽为 T,粗壮秦艽为 C(图5)。该位点可作为同时区分这三个物种的潜在分子标记。

3 SNP 鉴别位点验证

对粗茎秦艽 36 个单株、西藏秦艽 33 个单株、粗壮秦艽 35 个单株的 *ccsA-ndhD* 第 159 位点测序结果表明,该位点在三个物种内部均高度稳定,无种内变异。

结合 GenBank 已知叶绿体基因组序列,对秦艽组 10 种植物的 *ccsA-ndhD* 该位点进行分析。结果表明,西藏秦艽为 T;麻花苳、粗壮秦艽、大叶秦艽、达乌里秦艽、黄管秦艽、管花秦艽为 C;全萼秦艽、长梗秦艽和粗茎秦艽为 G。该位点对西藏秦艽的鉴定有较好的专属性。

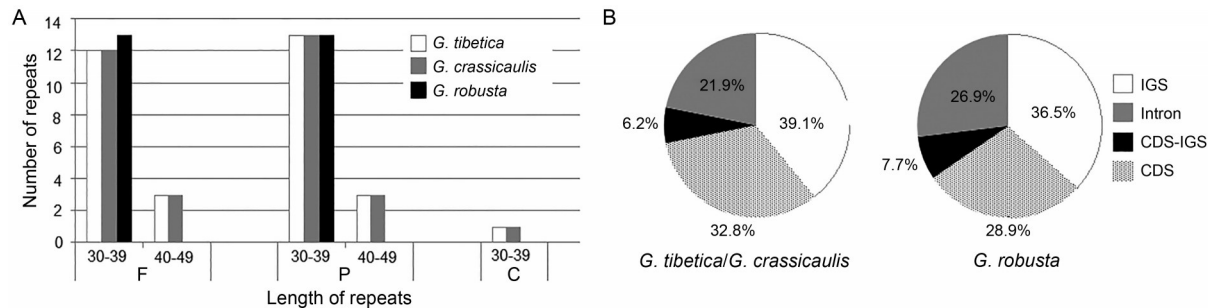


Figure 2 Repeat sequences in *G. tibetica*, *G. crassicaulis* and *G. robusta* chloroplast genomes. A: Frequency of the three repeat types by length; B: Repeat distribution among four different regions: IGS (intergenic spacer), CDS (coding sequence), CDS-IGS (part in CDS and part in IGS) and intron. F: Forward repeat sequence; P: Palindrome sequence; C: Complement repeat

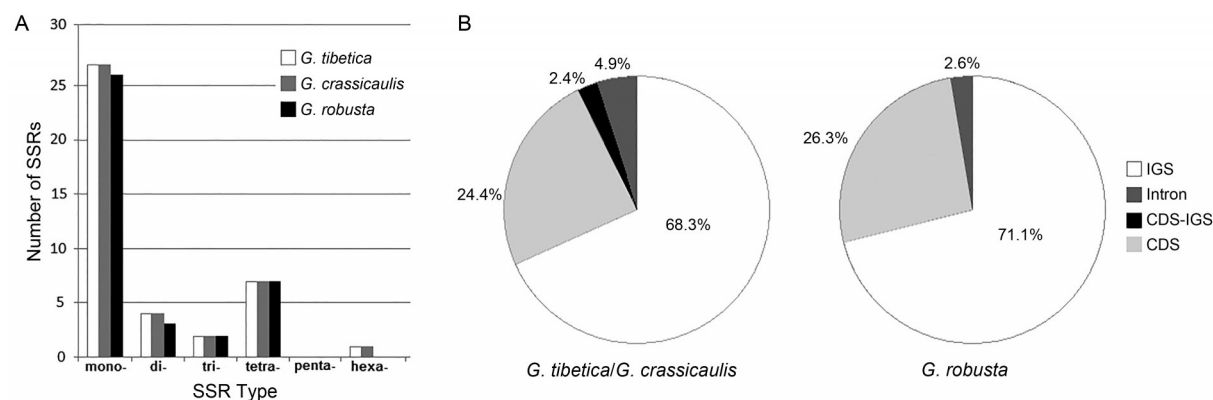


Figure 3 Simple sequence repeats in *G. tibetica*, *G. crassicaulis* and *G. robusta* chloroplast genomes. A: Frequency of different repeat types of SSRs; B: SSR distribution among four different regions: IGS (intergenic spacer), CDS (coding sequence), CDS-IGS (part in CDS and part in IGS) and intron

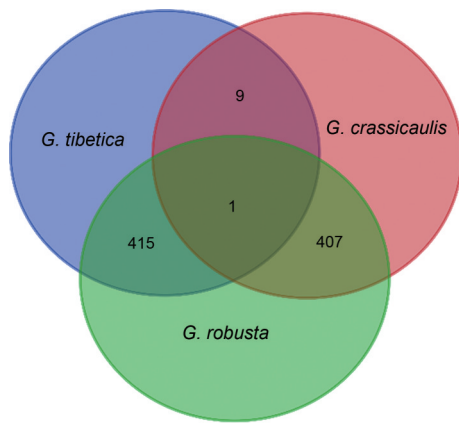


Figure 4 Venn diagram of variable sites among cp genomes of *G. tibetica*, *G. crassicaulis* and *G. robusta* (the number represents the variation sites between the corresponding species)

Table 3 SNPs and InDels among cp genomes of *G. tibetica*, *G. crassicaulis* and *G. robusta*. A: *G. crassicaulis*; B: *G. tibetica*; C: *G. Robusta*. * Different with each other

	LSC		SSC		IR		Total	
	SNP	InDel	SNP	InDel	SNP	InDel	SNP	InDel
A vs B	5	2	1	0	0	1	6	3
C vs B	216	73	96	6	15	9	327	88
A vs C	211	71	96	6	15	8	322	85
A vs B vs C*	0	0	1	0	0	0	1	0

4 混合样品鉴定

分别将三种秦艽组物种的总 DNA 两两混合后进行 *ccsA-ndhD* 片段的扩增和测序。测序图谱中, 该 SNP 位点呈现相应两个碱基的重叠峰, 即粗壮秦艽和粗茎秦艽的混合样为 C/G, 西藏秦艽和粗茎秦艽的混合样为 T/G, 粗壮秦艽和西藏秦艽的混合样为 C/T。将 3 个物种的总 DNA 混合后, 进行 *ccsA-ndhD* 片段的扩增和测序, 该 SNP 位点在测序峰图中呈现相应 3 个碱基的重叠峰 (C/G/T)。结果表明, 该位点可有效区分这三个物种的混合样品 (图 6)。

选取粗茎秦艽和西藏秦艽, 进一步探讨样品混合比例对检测结果的影响。结果表明, 混合比例与峰图中峰面积的比例呈现一定的相关性, 随着西藏秦艽在

混合样品中的比例逐渐加大, 该位点处西藏秦艽位点 T 在峰图中的比例总体呈增大趋势 (图 7)。但不同碱基的响应强度存在差异, 粗茎秦艽该位点处 G 的响应高于西藏秦艽 T。当粗茎秦艽与西藏秦艽的比例为 20:1 时, 测序峰图中仅可见粗茎秦艽的 G 位点, 而当粗茎秦艽与西藏秦艽的比例为 1:20 时, 两个物种的位点 G/T 均非常明显; 当两者等比例时, 粗茎秦艽的 G 位点峰值明显高于西藏秦艽 T。

讨论

秦艽组物种间形态相近, 且存在种下变异幅度较大、同域分布的近缘种间存在自然杂交等现象^[4,5], 导致一些近缘物种的鉴定较难把握。粗茎秦艽为中药秦艽的来源之一, 西藏秦艽和粗壮秦艽, 为其形态相近的近缘物种, 为确保中药秦艽基原植物来源的准确性, 在植物形态测量鉴定方法基础上^[13], 本研究基于粗茎秦艽、西藏秦艽和粗壮秦艽的叶绿体基因组序列, 筛选鉴别位点, 进一步建立相关近缘物种的分子鉴定方法。同时考虑到可能存在的种内基因型, 在各物种分布区广泛取样, 进行鉴别位点的稳定性验证, 筛选出 *ccsA-ndhD* 片段内可同时鉴定这三个物种的 SNP 位点。秦艽组 10 种植物的序列比较结果表明, 该位点对西藏秦艽的鉴定有较好的专属性。*ccsA-ndhD* 为非常规的 DNA 条形码片段, 应用较少, 但在多个分类群的叶绿体基因组比较分析中, 均发现其具有较高的变异性^[27,28], 可结合相关分类群的遗传背景特征进一步开发相应的分子标记。

SNP 在分子鉴定中有着广泛的应用, 尤其是对近缘物种的鉴定^[29,30]。本文进一步采用 SNP 双 (多) 峰检测方法, 通过测序图谱的直接观察, 可方便、直观的对各物种及混合样品进行区分鉴定, 同时在一定范围内可反映混合比例。但该方法对多批次样品的检验效率较低, 同时在定量方面存在局限。随着高通量测序技术的发展, 应用宏条形码技术对混合药材、中成药的鉴定研究日益成熟^[31-33], 可有效弥补以上缺陷。基于本



Figure 5 The alignment of *ccsA-ndhD* of *G. crassicaulis*, *G. tibetica* and *G. robusta*

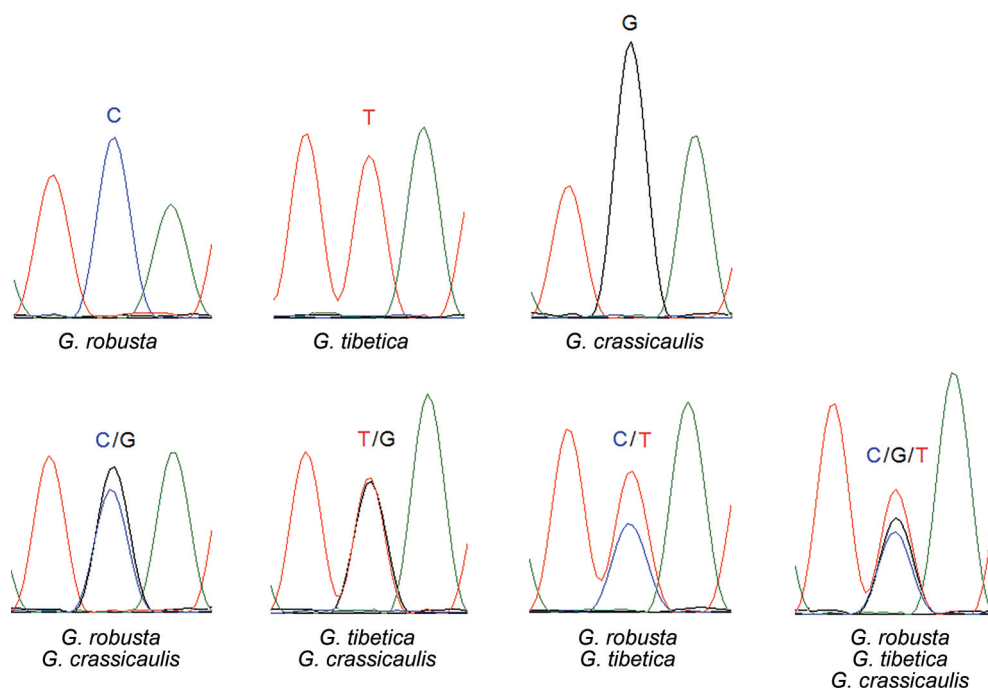


Figure 6 SNP peaks of *G. tibetica*, *G. crassicaulis* and *G. robusta*

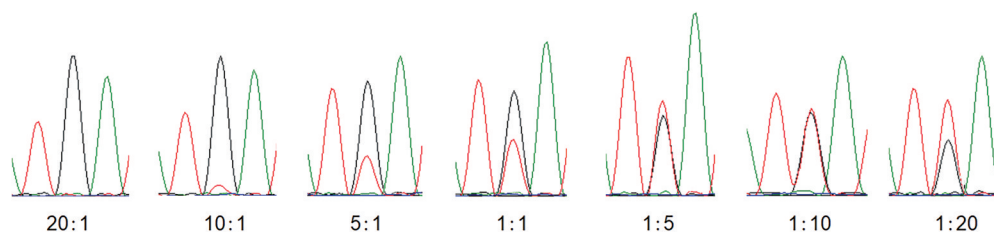


Figure 7 SNP peaks of *G. tibetica* and *G. crassicaulis* with different proportions (*G. crassicaulis* : *G. tibetica*)

实验结果, 下一步可开展粗茎秦艽及其近缘物种的宏条形码鉴定方法。

粗茎秦艽和西藏秦艽的叶绿体基因组序列高度相近, 仅有9处差异位点 (9/149 267), 提示两者可能存在相近的母系来源。系统发育分析也表明, 西藏秦艽可能为一杂交物种, 其叶绿体基因组序列与粗茎秦艽相比未呈现出外类群特征, 而是处于粗茎秦艽种内变异范围, 且在系统树中两者完全聚为一支 (另文发表)。两者在化学成分、药效活性等方面的差异, 也有待于进一步研究。同时藏药“解吉”的来源涉及秦艽组粗茎秦艽、西藏秦艽及粗壮秦艽等多种植物, 通过开展深入的系统性研究工作, 可为相关藏药材品种的鉴定、质量标准完善等工作提供科学依据。

作者贡献: 高娜娜是本文的第一作者, 负责分子生物学实验及稿件撰写; 嘎务负责民族植物学考察及样品采集工作; 倪梁红和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] Chinese Pharmacopocia Commission. Pharmacopocia of the People's Republic of China: Vol I (中华人民共和国药典: 一部) [S]. Beijing: China Medical Science Press, 2020: 282.
- [2] He TN, James SP. Genus *Gentiana*. In: Jussieu AL (ed) Flora of China, Vol 16. Gentianaceae [M]. Science Press, Beijing and Missouri Botanical Garden Press, St. Louis, 1995: 15-98.
- [3] Zhang XL, Wang YJ, Ge XJ, et al. Molecular phylogeny and biogeography of *Gentiana* sect. *Cruciata* (Gentianaceae) based on four chloroplast DNA datasets [J]. *Taxon*, 2009, 58: 862-870.
- [4] Zhang XL, Ge XJ, Liu JQ, et al. Morphological, karyological and molecular delimitation of two gentians: *Gentiana crassicaulis* versus *G. tibetica* (Gentianaceae) [J]. *Acta Phytotaxon Sin* (植物分类学报), 2006, 44: 627-640.
- [5] 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.
- [6] Luo K, Ma P, Yao H, et al. Identification of Gentianae Macro-

- phyllae Radix using the ITS2 barcodes [J]. Acta Pharm Sin (药学报), 2012, 47: 1710-1717.
- [7] 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.
- [8] Ni LH, Zhao ZL, Xiong B, et al. A strategy for identifying six species of Sect. *Cruciata* (*Gentiana*) in Gansu using DNA barcode sequences [J]. Acta Pharm Sin (药学报), 2016, 51: 821-827.
- [9] Lu JN, Zhao ZL, Ni LH, et al. The identification of Sect. *Cruciata* (*Gentiana*) species using mtDNA *nad1/b-c* and *nad5/d-e* fragments [J]. Acta Pharm Sin (药学报), 2019, 54: 166-172.
- [10] Liu J, Yan HF, Ge XJ. The use of DNA barcoding on recently diverged species in the genus *Gentiana* (Gentianaceae) in China [J]. PLoS One, 2016, 11: e0153008.
- [11] Yang YC. Tibetan Medicine (藏药志) [M]. Xining: Qinghai People Press, 1991: 9-12.
- [12] Zhao ZL, Dorje G, Wang ZT. Identification of medicinal plants used as Tibetan traditional medicine Jie-Ji [J]. J Ethnopharmacol, 2010, 132: 122-126.
- [13] 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.
- [14] McFadden GI. Primary and secondary endosymbiosis and the origin of plastids [J]. J Phycol, 2001, 37: 951-959.
- [15] Pal M. Plastid transformation in higher plants [J]. Annu Rev Plant Biol, 2004, 55: 289-313.
- [16] Ni LH, Zhao ZL, Mi M. Research progress on chloroplast genome of medicinal plants [J]. J Chin Med Mater (中药材), 2015, 38: 1991-1995.
- [17] 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, 2016, 577: 281-288.
- [18] Ni LH, Zhao ZL, Xu HX, et al. Chloroplast genome structures in *Gentiana* (Gentianaceae), based on three medicinal alpine plants used in Tibetan herbal medicine [J]. Curr Genet, 2017, 63: 241-252.
- [19] Dong BR, Zhao ZL, Ni LH, et al. Molecular markers based upon whole chloroplast genomes and identifying alpine *Gentiana waltonii* and *G. lhassica* (Gentianaceae) [J]. Acta Pharm Sin (药学报), 2021, 56: 2584-2591.
- [20] Tao Z, Jian W, Yun J, 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.
- [21] Sun SS, Fu PC, Zhou XJ, et al. The complete plastome sequences of seven species in *Gentiana* sect. *Kudoa* (Gentianaceae): insights into plastid gene loss and molecular evolution [J]. Front Plant Sci, 2018, 9: 493.
- [22] Gao Z, Liu Y, Wang X, et al. Derivative technology of DNA barcoding (nucleotide signature and SNP double peak methods) detects adulterants and substitution in Chinese patent medicines [J]. Sci Rep, 2017, 7: 5858.
- [23] Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf material [J]. Phytochem Bull, 1987, 19: 11-15.
- [24] Wu LW, Cui YX, Nie LP, et al. The characteristics of complete chloroplast genome sequence and phylogenetic analysis of *Dendrobium moniliforme* [J]. Acta Pharm Sin (药学报), 2020, 55: 1056-1066.
- [25] 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.
- [26] Kuang DY, Wu H, Wang YL, et al. Complete chloroplast genome sequence of *Magnolia kwangsiensis* (Magnoliaceae): implication for DNA barcoding and population genetics [J]. Genome, 2011, 54: 663-673.
- [27] Shen Y, Chen K, Gu C, et al. Comparative and phylogenetic analyses of 26 Magnoliaceae species based on complete chloroplast genome sequences [J]. Can J Forest Res, 2018, 48: 1456-1469.
- [28] Song Y, Dong W, Bing L, et al. Comparative analysis of complete chloroplast genome sequences of two tropical trees *Machilus yunnanensis* and *Machilus balansae* in the family Lauraceae [J]. Front Plant Sci, 2015, 6: 662.
- [29] Chen X, Liao B, Song J, et al. A fast SNP identification and analysis of intraspecific variation in the medicinal *Panax* species based on DNA barcoding [J]. Gene, 2013, 530: 39-43.
- [30] Merrill KR, Coleman CE, Meyer SE, et al. Development of single-nucleotide polymorphism markers for *Bromus tectorum* (Poaceae) from a partially sequenced transcriptome [J]. Appl Plant Sci, 2016, 4: 1600068.
- [31] Liu ZX, Xu J, Sun W, et al. Application of DNA metabarcoding in species identification of Chinese herbal medicines [J]. China J Chin Mater Med (中国中药杂志), 2019, 44: 1-8.
- [32] Xin T, Xu Z, Jing J, et al. Biomonitoring for traditional herbal medicinal products using DNA metabarcoding and single molecule, real-time sequencing [J]. Acta Pharm Sin B, 2018, 8: 488-497.
- [33] Liu J, Mu W, Shi M, et al. The species identification in traditional herbal patent medicine, Wuhu San, based on shotgun metabarcoding [J]. Front Pharmacol, 2021, 12: 607200.