• 药材与资源 •

藜芦属药用植物的叶绿体基因组比较分析和系统发育研究

田　星[1]，刘莹莹[2]，张颖敏[1]，杨从卫[1]，钱子刚[1*]，李国栋[1*]

1. 云南中医药大学 云南省傣医药与彝医药重点实验室，云南 昆明　650500
2. 云南省食品药品审核查验中心，云南 昆明　650106

摘　要：目的　以藜芦属药用植物蒙自藜芦 *Veratrum mengtzeanum*、大理藜芦 *V. taliense*、狭叶藜芦 *V. stenophyllum* 和毛叶藜芦 *V. grandiflorum* 为材料，对其叶绿体基因组进行组装和序列分析。**方法**　采用高通量测序技术对 4 种植物叶绿体基因组进行测序，并且使用 NOVOPlasty 和 Geneious R11 软件分别组装和注释序列，在此基础上进行基本结构和系统发育分析。**结果**　藜芦属 4 种植物叶绿体基因组具有典型的环状四分体结构，全长 151 875～153 711 bp，总 GC 含量 37.7%～37.8%。藜芦属叶绿体基因组均注释到 135 个基因，其中蛋白质编码基因 83～84 个，rRNA 基因 8 个和 tRNA 基因 38 个。通过比较发现，4 种植物 IR 区边界未出现明显的扩张或收缩；毛叶藜芦叶绿体基因组序列较保守，变异位点少；蒙自藜芦、大理藜芦和狭叶藜芦序列变异较大，种间差异较小；此外，在一个大单拷贝区（large single copy region，LSC）和小单拷贝区（small single copy region，SSC）区筛选到 13 个高变片段。系统发育研究表明，毛叶藜芦与蒙自藜芦、大理藜芦和狭叶藜芦亲缘关系较远，后两者亲缘关系最近。**结论**　对藜芦属 4 种植物叶绿体基因组结构和系统发育关系进行了分析，为后续开展分子鉴定及群体遗传学研究提供了数据资料。

关键词：藜芦属；叶绿体基因组；高通量测序；序列变异；系统发育

中图分类号：R282.12　　**文献标志码：**A　　**文章编号：**0253 - 2670(2022)04 - 1127 - 11

DOI: 10.7501/j.issn.0253-2670.2022.04.021

Comparative and phylogeny analysis of four *Veratrum* medicinal plants complete chloroplast genomes

TIAN Xing[1], LIU Ying-ying[2], ZHANG Ying-min[1], YANG Cong-wei[1], QIAN Zi-gang[1], LI Guo-dong[1]

1. Yunnan Key Laboratory of Dai and Yi Medicines, Yunnan University of Chinese Medicine, Kunming 650500, China
2. Center for Food and Drug Inspection of Yunnan, Kunming 650106, China

Abstract: Objective　The chloroplast genomes of *Veratrum mengtzeanum*, *V. taliense*, *V. stenophyllum* and *V. grandiflorum* were sequenced and assembled, laying a foundation for the further development of molecular identification and population genetics. **Methods**　The chloroplast genomes of four *Veratrum* plants were sequenced using high-throughput sequencing technology. NOVOPlasty and Geneious R11 were used to assemble and annotate the chloroplast genomes from the sequence reads respectively. Comparative and phylogenetic analyses were conducted based on gene annotation results. **Results**　The chloroplast genomes of *Veratrum* plants were 151 875—153 711 bp in length, GC content of 37.7%—37.8% and both exhibited the typical quadripartite circular structure. A total of 135 genes were annotated, including 83—84 protein-coding genes, 38 tRNA genes, eight rRNA genes. Comparative analyses of chloroplast genomes in four *Veratrum* plants revealed that no obvious expansion or contraction of the inverted repeat regions. The chloroplast genome sequence of *V. grandiflorum* was relatively conservative, with fewer mutation sites. The sequences variation of *V. mengtzeanum*, *V. taliense* and *V. stenophyllum* were relatively large, and the interspecific differentiation were relatively small. In addition, 13 sequences with higher variation interspecific were detected in the LSC and SSC regions. Phylogenetic trees showed that *V. grandiflorum* had farther relationship with *V. mengtzeanum*, *V. taliense*, *V. stenophyllum*, and the

收稿日期：2021-08-09

基金项目：国家自然科学基金资助项目（81560613）；2018 年中医药公共卫生服务补助专项“全国中药资源普查项目”（财社[2018]43 号）；云南省中青年学术技术带头人后备人才（202105AC160050）；云南省高层次人才培养支持计划“青年拔尖人才”专项（YNWR-QNBJ-2020-278）

作者简介：田　星（1997—），男，硕士，研究方向为中药资源开发与利用。E-mail: 820393356@qq.com

***通信作者：**李国栋，副教授，硕士生导师，研究方向为分子生药学。E-mail: liguodong@ynutcm.edu.cn

钱子刚，教授，博士生导师，研究方向为中药资源学。E-mail: qianzig@aliyun.com

relationship between *V. taliense* and *V. stenophyllum* was the closest. **Conclusion** The complete chloroplast genome of four *Veratrum* medicinal plants and phylogenetic relations were analyzed in this study, and the results will provide data information for the further development of molecular identification and population genetics.

Key words: *Veratrum*; chloroplast genome; high-throughput sequencing; sequence variation; phylogeny

藜芦属 *Veratrum* L.，隶属于百合科，中国分布有 13 种和 1 变种，该属植物多具有显著的药用价值，被广泛用于催吐、化瘀、止痛等。披麻草作为国家保密配方制剂“云南白药”和国家中药保护品种“一粒止痛丸”[1]主要组成药物之一，其基原为藜芦属植物。在《云南省药品标准》[2]中收录的披麻草基原植物为百合科藜芦属蒙自藜芦 *V. mengtzeanum* Loes. f、大理藜芦 *V. taliense* Loes. f、狭叶藜芦 *V. stenophyllum* Diels 和毛叶藜芦 *V. grandiflorum* (Maxim.) Loes. f 的干燥根及根茎；在《中华本草》[3]和《中药大辞典》[4]中记载披麻草为大理藜芦和狭叶藜芦的根；在《云南中药志》[5]中记载披麻草为蒙自藜芦的根；《中药辞海》[6]收录披麻草为大理藜芦的根及根茎；另外在《中国植物志》[7]和《中国高等植物图鉴》[8]中大理藜芦、蒙自藜芦和狭叶藜芦别名均有被称作披麻草，可见在不同资料中记载的披麻草的基原植物存在争议。

披麻草作为药物首载于《云南中草药选》[9]，现收载于 1996 年版《云南省药品标准》[2]，别名小棕包、天蒜、千张纸、大力王等[3]，根据《昆明民间常用草药》[10]中记载：披麻草味麻苦、性凉，有大毒。全草或根及根茎入药，具有内服催吐、撑骨、祛窃；外用止血、止痛、通窍的功效。目前，披麻草除了作为“云南白药”和“一粒止痛丸”的主要组成药物，也被广泛用于其他伤科药物中[11]，其显著的药用价值和经济价值，对于伤科药物的发展具有深远的意义[12]。但近年来由于披麻草未进入规模化人工种植，使用完全依赖自然生长缓慢的野生资源[13]，资源趋向渐危状态，越来越不能满足日常用药需求，从而导致披麻草基原植物使用混乱，市场上混伪品较多。因此采用分子生物学手段来解决基原和用药问题已经是迫在眉睫。

叶绿体是进行光合作用的重要细胞器，存在于植物、藻类和一些原生生物中。叶绿体是半自主性细胞器，拥有自己的基因组，称为叶绿体基因组（chloroplast DNA，cpDNA）[14-15]。随着测序技术进步和测序成本的降低，越来越多的研究以叶绿体基因组为基础展开。如阳春砂[16]、美丽芍药[17]、金铁锁[18]、芸薹[19]等药用植物的系统发育研究。吕瑞华等[20]采用 *psbA-trnH* 和 ITS2 序列对商陆资源及其混伪品进行了识别和鉴定，为商陆资源品种现状与分布研究提供了理论依据。董博然等[21]对龙胆科 5 属 27 个分类群叶绿体基因组进行了系统发育分析，结果表明龙胆属多枝组的短柄龙胆与秦艽组植物聚为一支，在亲缘关系上与秦艽组更近。

本研究通过对 3 种藜芦属植物（蒙自藜芦、大理藜芦、狭叶藜芦）进行叶绿体基因组测序得到的序列以及课题组前期对毛叶藜芦叶绿体基因组测序获得的序列（GenBank 登录号：MN613592）进行分析，揭示其基本特征与系统发育关系，以期为披麻草的基原植物、分子鉴定、资源保护与开发利用等研究提供参考资料。

1 材料

藜芦属 4 种植物新鲜叶片采自 8 个不同居群，采样信息见表 1，经云南中医药大学李国栋副教授鉴定，凭证标本存放于云南中医药大学中药材优良种苗繁育中心。

表 1 藜芦属 4 种植物居群基本信息

Table 1 Information on eight populations of four *Veratrum* plants

种名	拉丁名	采集地	居群编号	经度（E）	纬度（N）	凭证标本 GenBank 登录号
蒙自藜芦	*V. mengtzeanum* Loes. f	云南省宾川县	VmBC	100°25′	25°35′	5329240001
		云南省丽江市	VmLJ	100°14′	27°02′	5307000001
大理藜芦	*V. taliense* Loes. f	云南省师宗县	VtSZ	103°59′	24°45′	5303230001
		云南省漾濞县	VtYB	100°04′	25°43′	5329220001
		云南省宾川县	VtBC	100°25′	25°35′	5329240002
		云南省澄江县	VtCJ	102°31′	24°27′	5304220001
狭叶藜芦	*V. stenophyllum* Diels	云南省鹤庆县	VsHQ	100°17′	26°42′	5329320001
毛叶藜芦	*V. grandiflorum* (Maxim.) Loes. f	云南省昭通市	VgZT	103°05′	27°12′	MN613592

2 方法

2.1 基因组 DNA 的提取和测序

取新鲜幼嫩叶片，利用植物基因组 DNA 提取试剂盒（BioTeke 公司）提取总 DNA。使用琼脂糖凝胶电泳和 Nanodrop One 超微量分光光度计（美国赛默飞世尔科技有限公司）检测 DNA 质量及浓度。将检测合格的 DNA 送至上海美吉生物医药科技有限公司采用 Illumina HiSeq2500-PE150 平台进行建库测序。

2.2 叶绿体基因组的组装、注释及物理图谱绘制

将测序获得的 raw reads 使用 NGS QC Toolkit[22]滤过掉低质量区获得 clean reads。以尖被藜芦 *V. oxysepalum*（NC022715）作为参考序列，采用 NOVOPlasty[23]组装叶绿体基因组。再使用 DOGMA[24]对藜芦属 4 种植物叶绿体基因组进行注释，以蒙自藜芦（NC045300）叶绿体基因组作为参考序列，并且使用 Geneious R11 进行手动修正注释。用在线软件 OGDraw[25]绘制叶绿体基因组物理图谱。

2.3 数据统计

利用 MISA[26]软件分析叶绿体基因组 SSR 位点。设置参数：单核苷酸重复≥10；二核苷酸重复≥5；三核苷酸重复≥4；四、五、六核苷酸重复≥3。通过在线软件 REPuter[27]分析散在重复序列，参数设置：最小重复序列长度为 30（minimal repeat size＝30），最多碱基错配为 3（hamming distance＝3）。使用 IRscope[28]和在线比对工具 mVISTA[29]分别进行 IR 区边界收缩/扩张和全基因组比对分析。

2.4 系统发育分析

为确定 4 种植物在藜芦属中的系统位置，从 NCBI 数据库中下载 9 条藜芦属叶绿体基因组序列，分别是毛穗藜芦 *V. maackii* Regel（MN613590）；藜芦 *V. nigrum* L.（MN613595）；长梗藜芦 *V. oblongum* Loes. f（MN613593）；牯岭藜芦 *V. schindleri* (Baker) Loes. f（MN613588）；黑紫藜芦 *V. japonicum* Loes. f（MN613594）；南川藜芦 *V. nanchuanense* S. Z. Chen & G. J. Xu（MN613591）；兴安藜芦 *V. dahuricum* (Turcz.) Loes. f（MN699635）；尖被藜芦 *V. oxysepalum* Turcz（MW147219）；尖被藜芦 *V. oxysepalum* Turcz（NC022715）；以熊尾草 *Xerophyllum tenax* (Pursh) Nutt（NC027158）和丫蕊花 *Ypsilandra thibetica* Franch.（NC044639）作为外类群。基于 Model Finder[30]的 Akaike Information Criterion（AIC）筛选出核苷酸的最佳替代模型，然后使用 Phylosuite 软件分别构建最大似然（maximum likelihood，ML）系统发育树和贝叶斯（bayesian，BI）系统发育树（bootstrap＝1000）。

3 结果与分析

3.1 叶绿体基因组基本特征及分类

藜芦属 4 种植物的叶绿体基因组均为典型的环状四分体结构（图 1），包括 1 个大单拷贝区（large

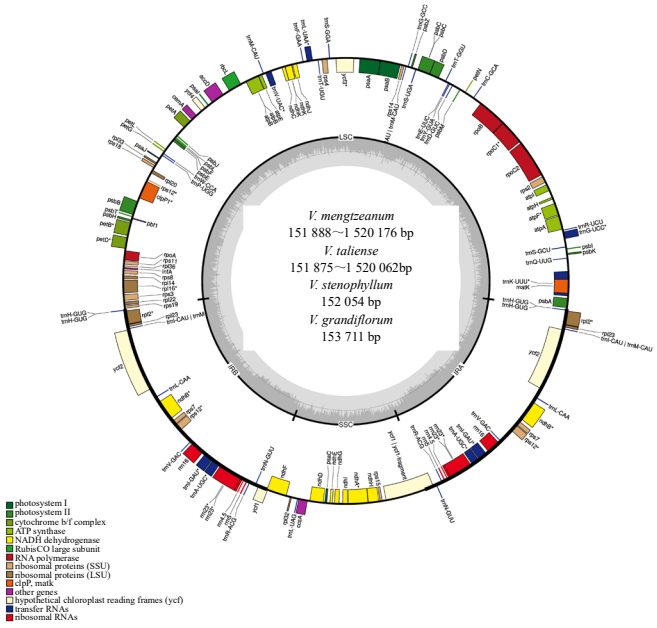

图 1 藜芦属 4 种植物叶绿体基因组图谱

Fig. 1 Gene map of four *Veratrum* plants chloroplast genomes

single copy region，LSC）；1 个小单拷贝区（small single copy region，SSC）和 1 对反向重复区（inverted repeat，IR）。序列总长度 151 875～153 711 bp，总 GC 含量 37.7%～37.8%；LSC、SSC、IR 区长度和 GC 含量分别为：81 967～83 367 bp、17 537～17 628 bp、26 148～26 358 bp；35.7%～35.8%、31.4%～31.5%、42.9%～43.0%。所有序列均注释到 135 个基因，蛋白质编码基因 83～84 个，rRNA 基因 8 个和 tRNA 基因 38 个（表 2）。

4 种植物全部基因分类情况如表 3 所示，根据其功能可以分为 4 大类：与光合作用有关的基因（45 个）、与自我复制有关的基因（74 个）、未知功能的蛋白质基因（10 个）以及成熟酶基因（*matK*）、囊膜蛋白基因（*cemA*）等其他基因（6 个）。在这些基因中含有 21 个双拷贝基因，包括 2 个核糖体大亚基

表 2　藜芦属植物叶绿体基因组基本特征

Table 2　Comparison of chloroplast genomes features of *Veratrum* plants

居群	Whole		LSC		SSC		IR		基因数量	蛋白编码基因	rRNA 基因	tRNA 基因
	长度/bp	GC/%	长度/bp	GC/%	长度/bp	GC/%	长度/bp	GC/%				
VmBC	151 888	37.8	81 967	35.8	17 537	31.5	26 192	43.0	135	84	8	38
VmLJ	152 176	37.8	82 238	35.8	17 542	31.5	26 198	43.0	135	84	8	38
VtSZ	151 875	37.8	82 029	35.8	17 550	31.5	26 148	43.0	135	84	8	38
VtYB	152 037	37.8	82 119	35.8	17 558	31.5	26 180	43.0	135	84	8	38
VtBC	152 062	37.8	82 124	35.8	17 542	31.5	26 192	43.0	135	84	8	38
VtCJ	152 061	37.8	82 123	35.8	17 542	31.5	26 198	43.0	135	84	8	38
VsHQ	152 054	37.8	82 108	35.8	17 550	31.5	26 198	43.0	135	84	8	38
VgZT	153 711	37.7	83 367	35.7	17 628	31.4	26 358	42.9	135	83	8	38

表 3　藜芦属植物叶绿体基因组基因分类

Table 3　List of genes of *Veratrum* plants chloroplast genomes

基因分类	基因分组	基因名称
self-replication	large subunit of ribosomal	*rpl2*[a,c]、*rpl14*、*rpl16*[a]、*rpl20*、*rpl22*、*rpl23*[c]、*rpl32*、*rpl33*、*rpl36*
	small subunit of ribosomal	*rps2*、*rps3*、*rps4*、*rps7*[c]、*rps8*、*rps11*、*rps12*[b,c,d]、*rps14*、*rps15*、*rps18*、*rps19*
	DNA dependent RNA polymerase	*rpoA*、*rpoB*、*rpoC1*[a]、*rpoC2*
	rRNA genes	*rrn4.5*[c]、*rrn5*[c]、*rrn16*[c]、*rrn23*[c]
	tRNA genes	*trnA*$_{UGC}$[a,c]、*trnC*$_{GCA}$、*trnD*$_{GUC}$、*trnE*$_{UUC}$、*trnF*$_{GAA}$、*trnfM*$_{CAU}$、*trnG*$_{GCC}$、*trnG*$_{UCC}$[a]、*trnH*$_{GUG}$[c]、*trnI*$_{CAU}$[c]、*trnI*$_{GAU}$[a,c]、*trnK*$_{UUU}$[a]、*trnL*$_{CAA}$[c]、*trnL*$_{UAA}$[a]、*trnL*$_{UAG}$、*trnM*$_{CAU}$、*trnN*$_{GUU}$[c]、*trnP*$_{UGG}$、*trnQ*$_{UUG}$、*trnR*$_{ACG}$[c]、*trnR*$_{UCU}$、*trnS*$_{GCU}$、*trnS*$_{GGA}$、*trnS*$_{UGA}$、*trnT*$_{GGU}$、*trnT*$_{UGU}$、*trnV*$_{GAC}$[c]、*trnV*$_{UAC}$[a]、*trnW*$_{CCA}$、*trnY*$_{GUA}$
gene for photosynthesis	photosystem I	*psaA*、*psaB*、*psaC*、*psaI*、*psaJ*
	photosystem II	*psbA*、*psbB*、*psbC*、*psbD*、*psbE*、*psbF*、*psbH*、*psbI*、*psbJ*、*psbK*、*psbL*、*psbM*、*psbN*、*psbT*、*psbZ*
	NadH oxidoreductase	*ndhA*[a]、*ndhB*[a,c]、*ndhC*、*ndhD*、*ndhE*、*ndhF*、*ndhG*、*ndhH*、*ndhI*、*ndhJ*、*ndhK*
	cytochrome b6/f complex	*petA*、*petB*[a]、*petD*[a]、*petG*、*petL*、*petN*
	ATP synthase	*atpA*[a]、*atpB*、*atpE*、*atpF*、*atpH*、*atpI*
	rubisco	*rbcL*
other genes	maturase	*matK*
	translational	*infA*
	protease	*clpP*[b]
	envelop membrane protein	*cemA*
	subunit of acetyl-CoA	*accD*
	c-type cytochrome synthesis gene	*ccsA*
unknown function	conserved open reading frames	*ycf1*[c]、*ycf2*[c]、*ycf3*[b]、*ycf4*、*ycf15*[c]、*ycf68*[c]

[a]包含 1 个内含子；[b]包含 2 个内含子；[c]双拷贝基因；[d]反式剪接基因

[a]Gene containing a single intron; [b]Gene containing two introns; [c]Gene with two copies; [d]Trans-splicing gene

（*rpl2*、*rpl23*）、2 个核糖体小亚基（*rps7*、*rps12*）、4 个 rRNA 基因（*rrn4.5*、*rrn5*、*rrn16*、*rrn23*）、8 个 tRNA 基因（$trnA_{UGC}$、$trnH_{GUG}$、$trnI_{CAU}$、$trnI_{GAU}$、$trnL_{CAA}$、$trnN_{GUU}$、$trnR_{ACG}$、$trnV_{GAC}$）、1 个 NADH 脱氢酶（*ndhB*）、4 个未知功能基因（*ycf1*、*ycf2*、*ycf15*、*ycf68*）。总共含有 20 个内含子，其中 *rpl2*、*rpl16*、*rpoC1*、$trnA_{UGC}$、$trnG_{UCC}$、$trnI_{GAU}$、$trnK_{UUU}$、$trnL_{UAA}$、$trnV_{UAC}$、*ndhA*、*ndhB*、*petB*、*petD*、*atpA* 基因各包含 1 个内含子，*rps12*、*clpP* 和 *ycf3* 基因各包含 2 个内含子，并且 *rps12* 基因位于 LSC 区和 IR 区，被分成 2 个独立的转录单元，为反式剪接基因（*trans*-splicing gene）。

3.2 散在重复序列及 SSR 分析

4 种植物叶绿体全基因组序列检测到长度不小于 30 且重复序列间相似度大于 90%的散在重复序列分别为正向重复（forward repeats，F）12～14 条、反向重复（reverse repeats，R）1～3 条、回文重复（palindromic repeats，P）18～21 条、互补重复（complement repeats，C）除 VmBC、VtSZ、VsHQ 居群各 1 条外，其余居群未检测到互补重复序列（表 4）。

在 8 个居群叶绿体基因组序列中，总共检测到 SSR 位点 64～72 个（表 5），其中单核苷酸重复为 41～50 个、二核苷酸重复为 10～13 个、三核苷酸重复 3～4 个、四核苷酸重复 5～6 个、五核苷酸重复 1～5 个，未检测到六核苷酸重复。重复最多的是单核苷酸重复（59.42%～69.70%），且主要是 A/T 的重复（表 6），占总重复类型的 63.93%，紧接着为 AT/AT 重复（16.82%）、C/G 重复（3.36%）、AAAG/CTTT 重复（3.00%）。这些 SSR 主要分布于叶绿体基因组的 LSC 区（66.67%～78.79%），编码基因序列中分布的 SSR 数量仅占总数的 19.44%～25.00%。

表 4 藜芦属植物散在重复序列

Table 4 Interspered repeat sequences information of *Veratrum* plants

居群	重复条数			
	正向重复	反向重复	回文序列	互补序列
VmBC	14	1	18	1
VmLJ	13	1	18	/
VtSZ	13	3	21	1
VtYB	13	3	19	/
VtBC	13	3	19	/
VtCJ	13	3	19	/
VsHQ	12	3	19	1
VgZT	13	3	19	/

3.3 IR 区边界比较分析

在叶绿体基因组中，IR 区与 LSC 和 SSC 区存在 4 个边界，即 LSC/IRb、IRb/SSC、SSC/IRa、IRa/LSC。本研究中藜芦属植物叶绿体基因组 4 个边界相对保守（图 2）。8 个居群中 VmBC、VmLJ、VtYB、VtBC、VtCJ、VsHQ 居群的 LSC/IRb 边界均在 *rps19* 基因内，且 *rps19* 基因均有 277 bp 分布在 LSC 区、2 bp 分布在 IRb 区；VtSZ 和 VgZT 居群的 *rps19* 基因均位于 LSC 区，分别距 LSC/IRb 边界 49、3 bp。在 IRb/SSC 边界的 *ycf1* 基因终止子缺失，为一个假基因，所有居群的 IRb/SSC、SSC/IRa 边界分别处于 *ycf1* 假基因和 *ycf1* 基因内，且 *ycf1* 假基因大部分分布在 IRb 区（960～991 bp），有 1～9 bp 位于 SSC 区；*ycf1* 基因大部分分布在 SSC 区（4401～4398 bp），有 960～991 bp 分布在 IRa 区。IRa/LSC 边界在所有居群中均位于 *trnH* 与 *psbA* 基因之间。

表 5 藜芦属植物叶绿体基因组中 SSR 位点类型及数量

Table 5 Type and number of SSR loci of four *Veratrum* plants chloroplast genomes

居群	SSR 位点数（占比/%）								
	单核苷酸	二核苷酸	三核苷酸	四核苷酸	五核苷酸	IR	SSC	LSC	CDS
VmBC	50(69.44)	10(13.89)	4(5.56)	6(8.33)	2(2.78)	0	16(22.22)	56(77.78)	14(19.44)
VmLJ	41(64.06)	12(18.75)	4(6.25)	6(9.38)	1(1.56)	0	15(23.44)	49(76.56)	14(21.88)
VtSZ	46(69.70)	11(16.67)	3(4.55)	5(7.58)	1(1.52)	0	14(21.21)	52(78.79)	16(24.24)
VtYB	44(68.75)	11(17.19)	3(4.69)	5(7.81)	1(1.56)	0	14(21.88)	50(78.13)	16(25.00)
VtBC	46(69.70)	11(16.67)	3(4.55)	5(7.58)	1(1.52)	0	14(21.21)	52(78.79)	15(22.73)
VtCJ	45(69.23)	11(16.92)	3(4.62)	5(7.69)	1(1.54)	0	14(21.54)	51(78.46)	15(23.08)
VsHQ	47(69.12)	11(16.18)	3(4.41)	5(7.35)	2(2.94)	0	16(23.53)	52(76.47)	15(22.06)
VgZT	41(59.42)	13(18.84)	4(5.80)	6(8.70)	5(7.25)	2(2.90)	21(30.43)	46(66.67)	15(21.74)

表 6　藜芦属植物叶绿体基因组 SSR 重复类型

Table 6　Repeat types of SSR of four *Veratrum* plants

重复类型	居群重复数量								占比/%
	VmBC	VmLJ	VtSZ	VtYB	VtBC	VtCJ	VsHQ	VgZT	
A/T	47	38	44	42	44	43	45	39	63.93
C/G	3	3	2	2	2	2	2	2	3.36
AT/AT	10	12	11	11	11	11	11	13	16.82
AAG/CTT	1	1	1	1	1	1	1	2	1.68
AAT/ATT	1	1	/	/	/	/	/	1	0.56
ACT/AGT	1	1	1	1	1	1	1	1	1.50
AGG/CCT	1	1	1	1	1	1	1	/	1.31
AAAG/CTTT	2	2	2	2	2	2	2	2	3.00
AAAT/ATTT	2	2	1	1	1	1	1	2	2.06
AATG/ATTC	2	2	2	2	2	2	2	2	3.00
AAAGT/ACTTT	1	1	1	1	1	1	1	/	1.31
AATAT/ATATT	1	1	/	/	/	/	/	/	0.37
AAAAT/ATTTT	/	/	/	/	/	/	/	1	0.19
AATAG/ATTCT	/	/	/	/	/	/	/	2	0.37
AATGG/ATTCC	/	/	/	/	/	/	1	/	0.19
ACTAT/AGTAT	/	/	/	/	/	/	/	1	0.19
ATATC/ATATG	/	/	/	/	/	/	/	1	0.19

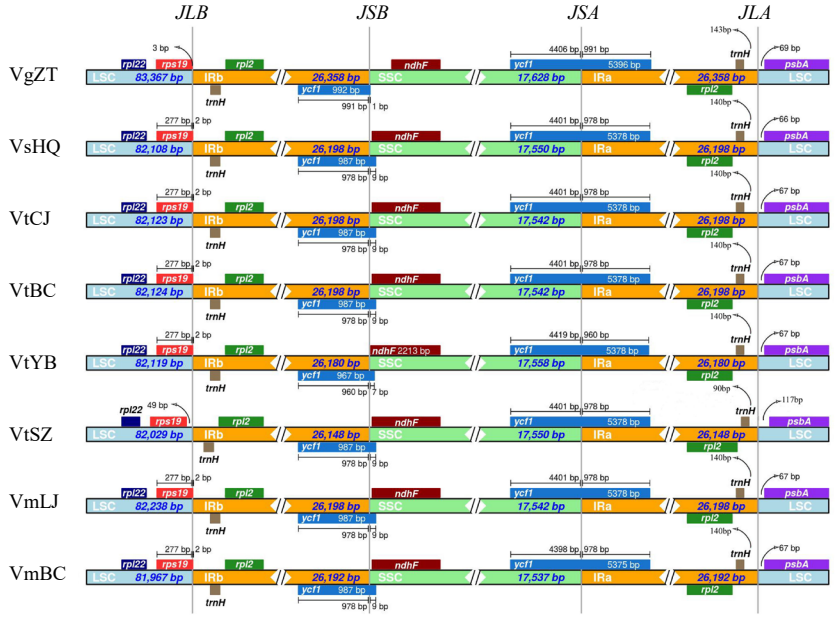

图 2　藜芦属植物叶绿体全基因组 IR 边界的比较示意图

Fig. 2　Comparison of boundaries of four *Veratrum* plants chloroplast genomes

3.4　叶绿体基因组比对及特异性 DNA 条形码筛选

利用 mVISTA 在线软件对藜芦属 8 个居群叶绿体基因组进行了比较分析。结果显示，在居群水平上，毛叶藜芦居群叶绿体基因组序列较保守，变异位点少，其他 7 个居群在种内序列变异较大，种间差异较小，在 *atpH-atpI*、*trnD*$_{GUC}$-*trnY*$_{GUA}$、*psaA-ycf3*、*trnT*$_{UGU}$-*trnL*$_{UAA}$、*petD*、*ndhB-trnL*$_{CAA}$ 区存在不同程度的变异（图 3）。整体上，变异较高的区域基本集中在 LSC 和 SSC 区域的非编码区，编码区比较保守。

使用 DnaSP 6 软件的多态性分析检测藜芦属 4 种植物叶绿体基因组中的高度可变区。核苷酸多样性（Pi）整体变化范围为 0～0.008，且 LSC 区和 SSC 区核苷酸多样性高于 IR 区。在图 4 中标记了 Pi＞0.006 的基因片段，分别是 LSC 区（*trnK*$_{UUU}$-*psbI*、*trnL*$_{UAA}$-*ndhJ*、*trnV*$_{UAC}$-*atpE*、*accD-psaI*）、SSC 区（*ndhF-trnL*$_{UAG}$、*ndhD-ndhE*、*ycf1*）。

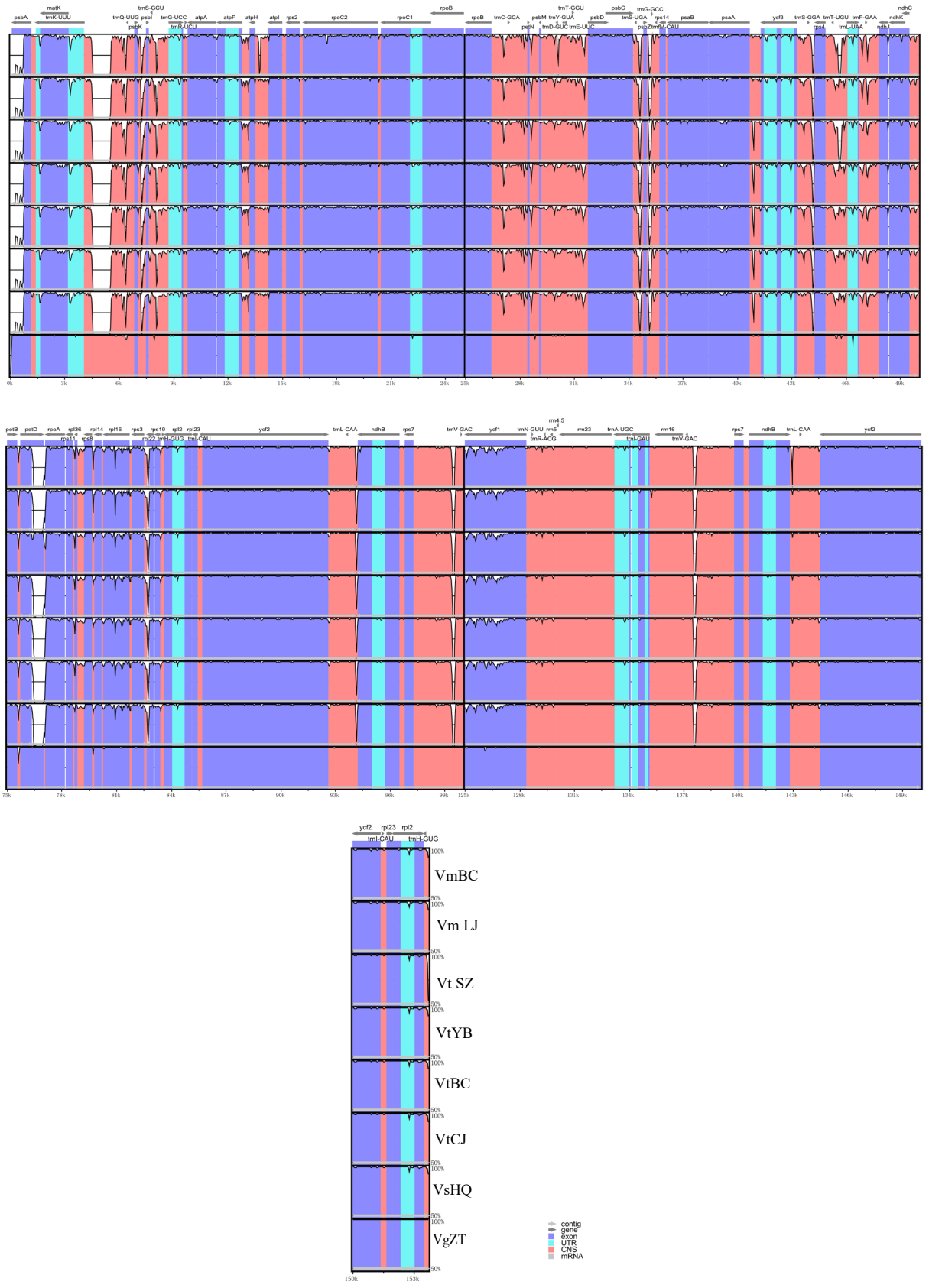

图 3　藜芦属植物叶绿体基因组全局比对

Fig. 3　Global comparison of *Veratrum* plants chloroplast genomes

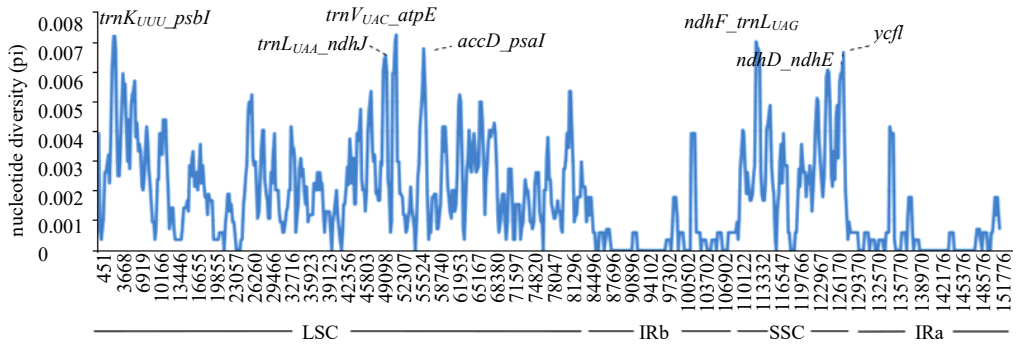

图 4　藜芦属植物叶绿体基因组核苷酸多样性

Fig. 4　Nucleotide diversity of *Veratrum* plants chloroplast genomes

3.5　系统发育分析

筛选出核苷酸的最佳替代模型为 GTR＋I＋G，使用 Phylosuite 软件对测序获得的藜芦属 8 个居群叶绿体基因组以及 NCBI 下载的 9 条藜芦属序列和 2 条其他属外类群分别构建最大似然（maximum likelihood，ML）和贝叶斯（bayesian，BI）系统发育树。如图 5 所示，藜芦属被明显分为了 5 个进化枝，毛叶藜芦在 A 节点和蒙自藜芦、大理藜芦、狭叶藜芦分化开（支持率 100%），并且与南川藜芦 *V. nanchuannens* 亲缘关系最近；蒙自藜芦在 B 节点与大理藜芦和狭叶藜芦分化开（支持率 100%）形成进化枝 III；大理藜芦和狭叶藜芦在 C 节点处分别形成了进化枝 I 和进化枝 II（支持率 100%），亲缘关系较近。

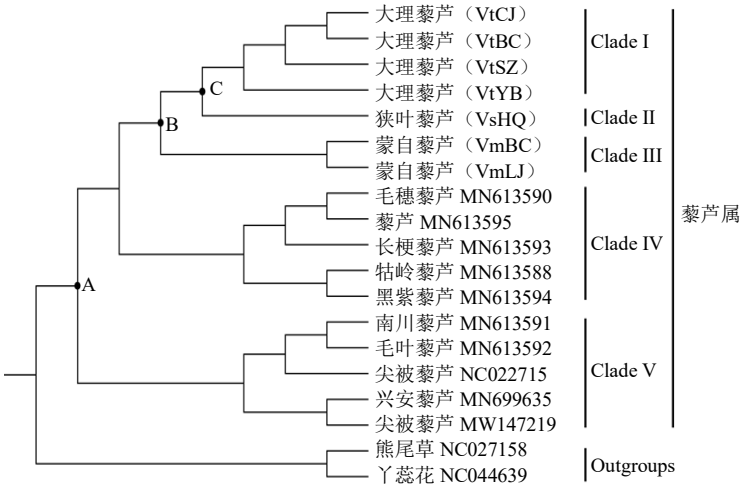

图 5　基于 19 个叶绿体基因组构建的藜芦属 ML 和 BI 系统发育树

Fig. 5　ML and BI phylogenetic tree based on 19 *Veratrum* chloroplast genomes

4　讨论

披麻草作为“云南白药”和“一粒止痛丸”的主要组成药物，由于未进入规范化人工种植和过度采挖，导致野生居群数量急剧下降。目前，研究者们为了保护和合理开发披麻草野生资源，对其进行了植物分类学[12]、生药学鉴定[1]、化学成分[31]、药理作用[32-34]等方面的研究。然而关于披麻草基原植物叶绿体基因组的研究还尚未见报道。

叶绿体作为植物进行光合作用的场所，在植物的整个生命过程中具有重要意义，其基因组序列高度保守。本研究对披麻草基原植物叶绿体基因组进行了分析，结果表明其基本结构均为环状四分体结构，基因组大小差异较小，最大差仅 1836 bp，LSC 区差异最大（1400 bp）。总 GC 含量 37.7%～37.8%；IR 区的 GC 含量最高（42.9%～43.0%），所有居群均注释到 135 个基因，蛋白质编码基因 83～84 个，

rRNA 和 tRNA 基因分别为 8、38 个。这与金铁锁[18]、湖北麦冬[35]、三叶崖爬藤[36]等的研究结果相似，体现出叶绿体基因组结构稳定，总体进化速率较慢。

SSR 分子标记由于其数量丰富、多态性高、信息含量高、不受外界环境等因素的影响，并且是共显性遗传已被广泛用于药用植物的群体遗传学研究[37-38]。本研究共检测到 SSR 位点 64～72 个，重复最多的是单核苷酸重复（59.42%～69.70%），其次为二核苷酸重复（13.89%～18.84%），主要重复类型是 A/T 重复，紧接着为 AT/AT 重复，共占比 80.75%。张明英等[39]报道了柴胡属叶绿体基因组 SSR 主要由 polyA 或 polyT 构成，此结果与本实验的研究结果一致。此外，在 8 个居群中检测到散在重复序列 32～38 条。这些 SSR 以及散在重复序列为今后开发披麻草 SSR 标记以及群体遗传学研究提供了候选分子标记。

在叶绿体基因组中，其差异主要体现在 IR 区边界的收缩和扩张上。本研究中藜芦属 4 种植物叶绿体基因组在 4 个边界上除 LSC/IRb 边界存在微小的差异外，其余边界均较保守。在 IRb/SSC 边界的 *ycf1* 基因终止子缺失，为 1 个假基因。此现象在石豆兰属[40]、菜头肾[41]等植物中也有报道。基于 mVISTA 和 Pi 分析结果显示，在 *atpH-atpI*、$trnD_{GUC}$-$trnY_{GUA}$、*psaA-ycf3*、$trnT_{UGU}$-$trnL_{UAA}$、*petD*、*ndhB*-$trnL_{CAA}$、$trnK_{UUU}$-*psbI*、$trnL_{UAA}$-*ndhJ*、$trnV_{UAC}$-*atpE*、*accD-psaI*、*ndhF*-$trnL_{UAG}$、*ndhD-ndhE*、*ycf1* 区存在明显变异，这些片段可为披麻草分子鉴定的特异性 DNA 条形码的筛选提供参考。

基于 ML 和 BI 系统发育树结果表明，毛叶藜芦与其他 3 个种亲缘关系较远，蒙自藜芦、大理藜芦和狭叶藜芦亲缘关系较近。披麻草作为药用价值较高的伤科药物，在不同资料中记载的基原植物存在争议。本研究结果显示：从叶绿体基因组角度分析，毛叶藜芦同其他 3 个种叶绿体基因组序列差异较大，变异较小；ML 和 BI 系统发育树一致表明毛叶藜芦与其他 3 个种亲缘关系较远。本课题组前期通过《中国植物志》[7]中的描述对披麻草基原植物进行鉴定时发现蒙自藜芦、大理藜芦、狭叶藜芦形态学特征（花序、叶片、果实）极其相似，仅存在微小的差异，如蒙自藜芦花被片上存在明显的腺体，大理藜芦茎多分枝，花序相对密集，狭叶藜芦茎几乎没有分支，花序相对扩展，而毛叶藜芦与前三者形态差异较大，其叶片、花被片和花序较前 3 者宽大；在对披麻草药材鉴定时发现蒙自藜芦、大理藜芦、狭叶藜芦药材外观性状十分相似，但毛叶藜芦药材外观性状和大小与前三者差异较大。

目前，研究者们从蒙自藜芦中分离了藜芦胺、藜芦明宁、藜芦胺-*N*-氧化物、3-当归酰基计明碱、胡萝卜苷等 11 个化合物[42]；从大理藜芦中分离了大理藜芦碱 A～D、介藜芦胺、介芬胺、狭叶藜芦碱乙、3-藜芦酰棋盘花胺、计明碱、新大理藜芦碱 A 和 B 等 14 个化合物[43-45]；从狭叶藜芦中分离了藜芦胺、棋盘花胺、介藜芦胺、介芬胺、狭叶藜芦碱甲～丁、β-查茄碱、原藜芦碱 B、季明碱等 22 个化合物[46-47]；从毛叶藜芦中分离并鉴定了藜芦胺、表红介芬胺、介芬胺、3-当归酰基棋盘花胺、藜芦托素、表红介藜芦碱、异玉红芥芬胺、桑皮苷 A、等 21 个化合物[48-49]。从分离出的化学成分可以看出虽然 4 种植物化学成分有所差异，但基本结构类型都涉及了藜芦胺型、介藜芦碱型的甾体生物碱。杨崇仁等[42]报道了蒙自藜芦须根中的甾体生物碱以藜芦胺的含量较高，并且药理实验表明藜芦胺有镇痛和降压的生理活性；Li 等[50]采用醋酸致小鼠扭体法和角叉菜胶致小鼠足肿胀法得出大理藜芦中介芬胺型生物碱具有抗炎和镇痛作用；马丽焱等[34]发现狭叶藜芦碱甲有较强且持久的降压和减慢心率的生理活性。刘莹莹[11]对披麻草基原植物中藜芦胺和介芬胺含量进行了测定，结果表明 4 种植物间藜芦胺和介芬胺含量差异较大，不同药用部位含量差异也较大，其中根及根茎含量最高，此结果表明了披麻草以根及根茎入药的合理性。

综上所述，本研究从叶绿体基因组角度认为蒙自藜芦、大理藜芦、狭叶藜芦为披麻草基原植物比较合理，为披麻草基原问题提供了叶绿体基因组方面的资料。但披麻草化学成分和药理研究表明 4 种植物均含有藜芦胺型和介藜芦碱型甾体生物碱，且是镇痛和降压作用的主要活性成分，说明 4 种植物间差异不大，该结果没有支持本研究结果，原因可能是由于目前药理研究仅是针对 1 种植物进行研究，而没有将 4 种植物一起进行药理比对所致。因此，后期还需从药理活性、化学等方面进一步深入研究。

利益冲突 所有作者均声明不存在利益冲突

参考文献

[1] 尹子丽, 谢晖, 张洁. 披麻草生物学特性初步研究 [J]. 云南中医中药杂志, 2016, 37(4): 54-57.

[2] 云南省卫生厅. 云南省药品标准: 1996 年版 [M]. 昆明: 云南大学出版社, 1998: 26.
[3] 国家中医药管理局《中华本草》编委会. 中华本草: 精选本（上、下册） [M]. 上海: 上海科学技术出版社, 1998: 182-183.
[4] 江苏新医学院. 中药大辞典(上) [M]. 上海: 上海人民出版社, 1977: 56.
[5] 黎光南. 云南中药志-I [M]. 昆明: 云南科技出版社, 1990: 263.
[6] 严永清. 中药辞海-第二卷 [M]. 北京: 中国医药科技出版社, 1996: 26.
[7] 中国科学院中国植物志编辑委员会. 中国植物志 （第十四卷）[M]. 北京: 科学出版社, 1980: 123.
[8] 中国科学院植物研究所. 中国高等植物图鉴-第二册: 补编 [M]. 北京: 科学出版社, 1983: 427.
[9] 昆明军区后勤部卫生部. 云南中草药选 [M]. 昆明: 昆明军区后勤部卫生部, 1970: 126.
[10] 昆明市卫生局. 昆明民间常用草药 [M]. 昆明: 昆明市卫生局, 1970: 126.
[11] 刘莹莹. 披麻草药材及饮片的质量标准研究 [D]. 昆明: 云南中医药大学, 2019.
[12] 尹子丽, 谭文红, 杨仙雨, 等. 披麻草三种基源植物的研究进展 [J]. 云南中医中药杂志, 2014, 35(3): 60-62.
[13] 尹子丽, 冯泽辉, 陈静怡, 等. 披麻草野生居群表型多样性分析研究 [J]. 云南中医中药杂志, 2020, 41(7): 77-82.
[14] McFadden G I. Primary and secondary endosymbiosis and the origin of plastids [J]. *J Phycol*, 2001, 37(6): 951-959.
[15] Chen Q, Wu X B, Zhang D Q. Phylogenetic analysis of *Fritillaria cirrhosa* D. Don and its closely related species based on complete chloroplast genomes [J]. *PeerJ*, 2019, 7: e7480.
[16] 马孟莉, 孟衡玲, 张薇, 等. 阳春砂叶绿体全基因组解析及系统发育研究 [J]. 西北植物学报, 2020, 40(6): 978-986.
[17] 张明英, 王西芳, 高静, 等. 美丽芍药叶绿体全基因组解析及系统发育分析 [J]. 药学学报, 2020, 55(1): 168-176.
[18] 高亚芳, 刘莹莹, 杨从卫, 等. 金铁锁叶绿体基因组序列及其系统发育分析 [J]. 中草药, 2019, 50(22): 5532-5536.
[19] Li P R, Zhang S J, Li F, *et al*. A phylogenetic analysis of chloroplast genomes elucidates the relationships of the six economically important *Brassica* species comprising the triangle of U [J]. *Front Plant Sci*, 2017, 8: 111.
[20] 吕瑞华, 冯昭, 马添翼, 等. 陕西关中野生商陆资源的 ITS2 和 psbA-trnH 条形码序列研究 [J]. 药学学报, 2020, 55(8): 1951-1956.
[21] 董博然, 赵志礼, 倪梁红, 等. 龙胆科叶绿体基因组结构特征及其药用植物鉴定意义 [J]. 中草药, 2020, 51(6): 1641-1649.
[22] Patel R K, Jain M. NGS QC Toolkit: A toolkit for quality control of next generation sequencing data [J]. *PLoS One*, 2012, 7(2): e30619.
[23] Dierckxsens N, Mardulyn P, Smits G. NOVOPlasty: de novo assembly of organelle genomes from whole genome data [J]. *Nucleic Acids Res*, 2017, 45(4): e18.
[24] Wyman S K, Jansen R K, Boore J L. Automatic annotation of organellar genomes with DOGMA [J]. *Bioinformatics*, 2004, 20(17): 3252-3255.
[25] Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes [J]. *Nucleic Acids Res*, 2019, 47(W1): W59-W64.
[26] Beier S, Thiel T, Münch T, *et al*. MISA-web: a web server for microsatellite prediction [J]. *Bioinformatics*, 2017, 33(16): 2583-2585.
[27] Kurtz S, Choudhuri J V, Ohlebusch E, *et al*. REPuter: the manifold applications of repeat analysis on a genomic scale [J]. *Nucleic Acids Res*, 2001, 29(22): 4633-4642.
[28] Amiryousefi A, Hyvönen J, Poczai P. IRscope: an online program to visualize the junction sites of chloroplast genomes [J]. *Bioinformatics*, 2018, 34(17): 3030-3031.
[29] Frazer K A, Pachter L, Poliakov A, *et al*. VISTA: computational tools for comparative genomics [J]. *Nucleic Acids Res*, 2004, 32: W273-W279.
[30] Kalyaanamoorthy S, Minh B Q, Wong T K F, *et al*. ModelFinder: Fast model selection for accurate phylogenetic estimates [J]. *Nat Methods*, 2017, 14(6): 587-589.
[31] 张琴, 范琦, 秦少容. 披麻草根的研究进展 [J]. 中国药业, 2005, 14(7): 85-86.
[32] Zhou C X, Tan R, Ye W, *et al*. Steroidal alkaloids from *Veratrum taliense* [J]. *J Phycol*, 2000, 36(6): 922-929.
[33] 乐开礼. 蒙自藜芦的降压作用及机制 [J]. 药学学报, 1963(11): 668-672.
[34] 马丽焱, 周远鹏, 江京俐, 等. 狭叶藜芦碱甲对犬血流动力学的影响 [J]. 中草药, 1998, 29(2): 105-107.
[35] 王梦涵, 胡帅军, 杨楚虹, 等. 湖北麦冬、川麦冬及杭麦冬叶绿体基因组分析 [J]. 中国实验方剂学杂志, 2020, 26(8): 182-191.
[36] 蒋明, 王军峰, 应梦豪, 等. 三叶崖爬藤叶绿体基因组的组装与序列分析 [J]. 中草药, 2020, 51(2): 461-468.
[37] 葛淑俊, 孟义江, 李广敏, 等. 我国药用植物遗传多样性研究进展 [J]. 中草药, 2006, 37(10): 1584-1589.

[38] Akkaya M S, Bhagwat A A, Cregan P B. Length polymorphisms of simple sequence repeat DNA in soybean [J]. *Genetics*, 1992, 132(4): 1131-1139.

[39] 张明英, 张雨曲, 李依民, 等. 北柴胡、紫花阔叶柴胡叶绿体全基因组解析及柴胡属药用植物叶绿体基因组比较与系统发育分析 [J]. 药学学报, 2021, 56(2): 618-629.

[40] 杨嘉鹏, 朱紫乐, 范雅娟, 等. 三种石豆兰属药用植物的叶绿体基因组比较分析及其在物种鉴定中的意义 [J]. 药学学报, 2020, 55(11): 2736-2745.

[41] 赵祺, 余佳兴, 秦宇雯, 等. 基于高通量测序的菜头肾叶绿体基因组的组装及序列分析 [J]. 中草药, 2021, 52(6): 1744-1750.

[42] 杨崇仁, 刘润民, 周俊, 等. 蒙自藜芦的甾体生物碱研究 [J]. 云南植物研究, 1987, 9(3): 359-364.

[43] 谭仁祥, 闵知大. 大理藜芦中甾体和异甾体生物碱的研究 [J]. 南药译丛, 1986, 10(1): 87-88.

[44] 闵知大, 谭仁祥, 郑启泰, 等. 大理藜芦碱 B 的结构及其绝对构型的确定 [J]. 药学学报, 1988, 23(8): 584-587.

[45] Mizuno M, Tan R X, Zhen P, *et al*. Two steroidal alkaloid glycosides from *Veratrum taliense* [J]. *Phytochemistry*, 1990, 29(1): 359-361.

[46] 梁光义, 孙南君. 狭叶藜芦中活性成分的化学研究 III. β1-查茄碱的结构和狭叶藜芦碱丙、碱丁的部分结构研究 [J]. 药学学报, 1984, 19(6): 431-436.

[47] 吕轶峰. 狭叶藜芦的化学成分研究与部分生物碱的生物转化 [D]. 沈阳: 沈阳药科大学, 2008.

[48] 张谨. 毛叶藜芦根部化学成分研究 [D]. 武汉: 华中科技大学, 2016.

[49] 时嘉敏, 刘兴鸿, 何翠林, 等. 毛叶藜芦中 1 个新的生物碱 [J]. 中草药, 2020, 51(13): 3365-3371.

[50] Li Q, Yang K X, Zhao Y L, *et al*. Potent anti-inflammatory and analgesic steroidal alkaloids from *Veratrum taliense* [J]. *J. Ethnopharmacol*, 2016, 179: 274-279.

[责任编辑 时圣明]