Research
Traditional Chinese Medicine—Article

4 种鼠尾草药用植物叶绿体基因组的对比分析

梁从莲^{a,b,c}, 王磊^{a,b,d}, 雷隽^e, 段宝忠^f, 马维思^g, 肖水明^b, 齐海军^c, 王振^c, 刘尧奇^c, 沈晓凤^b, 郭帅^b, 胡灏禹^b, 徐江^{b,*}, 陈士林^{b,*}^a College of Pharmacy, Shandong University of Traditional Chinese Medicine (TCM), Jinan 250355, China^b Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China^c State Key Laboratory of Innovative Natural Medicine and TCM Injections, Jiangxi Qingfeng Pharmaceutical Co. Ltd., Ganzhou 341000, China^d Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing 100045, China^e School of Biological Science, The University of Queensland, Brisbane, QLD 4072, Australia^f College of Pharmacy, Dali University, Dali 671000, China^g Institute of Medicinal Plants, Yunnan Academy of Agricultural Sciences, Kunming 650205, China

ARTICLE INFO

Article history:

Received 23 July 2018

Revised 27 November 2018

Accepted 7 January 2019

Available online 16 August 2019

关键词

鼠尾草属
叶绿体基因组
对比分析

摘要

本草基因组学是中药研究与开发的一个新兴领域。它将中药研究与基因组学相结合,有助于建立中药的科学有效性,使其在医学领域得到更广泛的应用。鼠尾草属 (*Salvia*, Linn.) 是唇形科的一个大属,包括很多重要的药用植物。在本草基因组学研究中,我们对两种被用作丹参替代品的鼠尾草属物种——甘西鼠尾草 (*S. przewalskii*)、戟叶鼠尾草 (*S. bulleyana*)——的叶绿体 (cp) 全基因组进行了测序,并与另外两种已有叶绿体全基因组公布的鼠尾草——丹参 (*S. miltiorrhiza*) 和鼠尾草 (*S. japonica*)——的叶绿体全基因组进行了比较。比较内容为基因组结构、基因数目、类型和重复序列。注释结果表明,本研究中的两种鼠尾草均含有114个独有基因,其中蛋白编码基因80个,转运RNA (tRNA) 基因30个,核糖体RNA (rRNA) 基因4个。重复序列分析显示,两种鼠尾草叶绿体基因组都含有21条正向重复序列和22条回文重复序列,甘西鼠尾草和戟叶鼠尾草分别有17条和21条串联重复序列。对4种鼠尾草叶绿体基因组的同源性比较表明,其编码区具有高度的序列相似性,而基因间隔区域具有较高的差异。序列比对和单核苷酸多态性 (SNP) 分析发现了一些可以鉴定鼠尾草属物种的候选片段,如 *trnV-ndhC*、*trnQ-rps16*、*atpI-atpH*、*psbA-ycf3* 等基因间区和 *ycf1*、*rpoC2*、*ndhF*、*matK*、*rpoB*、*rpoA*、*accD* 等基因。所有的结果——包括重复序列和SNP位点,反向重复 (IR) 区域边界,以及系统发育分析——都表明,从遗传学角度来看,甘西鼠尾草和戟叶鼠尾草是极其相似的。本文报道的两种鼠尾草的叶绿体基因组序列将为鼠尾草属药用植物的育种、物种鉴定、系统进化和叶绿体基因工程研究奠定基础。

© 2019 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. 引言

鼠尾草属是唇形科的一个大属,在全世界约有1000个物种;其中中国约有84种,主要分布在西南地区 [1]。在这84种里,近30种被记载为药用植物 [2]。丹参是一

种具有代表性的中药 [3],对心脑血管疾病有显著疗效 [4,5]。《中国药典》表明,中药丹参的唯一来源是丹参 (*Salvia miltiorrhiza*) 这种植物 [6]。然而,许多其他鼠尾草属植物具有与丹参相似的化学成分和相应的药理作用,并且在民间和地方用药被当作丹参来使用。例如,

* Corresponding author.

E-mail address: jxu@icmm.ac.cn (J. Xu), slchen@icmm.ac.cn (S. Chen).

甘西鼠尾草是一种多年生草本植物，广泛分布于中国西南部[7]，它作为丹参的替代品已有至少300年历史。这种植物在四川省还被当作秦艽的替代品。许多研究表明，甘西鼠尾草与丹参具有相似的成分；但甘西鼠尾草根中的脂溶性成分含量比丹参高数倍，具有很大的应用价值[8–10]。戟叶鼠尾草又称“紫丹参”，主要分布于云南大理，常被用作民间丹参。

叶绿体是绿色植物重要的细胞器，在自养光合作用中起着关键的作用[12–14]。叶绿体基因组是细胞中典型的环状多拷贝DNA分子[15,16]。植物叶绿体基因组结构十分保守，一般可以分为四个片段，即两个反向重复（IR）区（IRa和IRb），以及把IR区隔开的一个大的单拷贝（LSC）区域和一个小的单拷贝（SSC）区域[17,18]。不同物种叶绿体基因组的大小差异主要是由于IR区的收缩或扩张造成的[19,20]。随着新测序技术的出现，测序时间越来越短、成本越来越低；目前，已有超过2000个叶绿体基因组在美国国家生物技术信息中心（NCBI）上公布[21]。使用DNA条形码来区分亲缘关系较近的物种和种下类群的方法由于基因片段短、系统发育信息数量少而受到阻碍。因此，一些研究人员提出利用整个叶绿体基因组作为物种鉴定的“超级条形码”[22]。此外由于叶绿体转化技术的表达高效性、定点整合性和母体遗传性，它已显示出相当大的遗传改良潜力[23–26]。

基因组数据，包括细胞器基因组数据，可以为中药基原物种的研究提供分子基础。在之前的工作中，我们研究了很多基因组序列[27–29]，包括丹参叶绿体全基因组序列[30]，以及丹参核基因组的序列草图和分析[31]。在本研究中，我们使用二代测序平台获得了甘西鼠尾草和戟叶鼠尾草的叶绿体全基因组序列。然后，我们将这两个叶绿体基因组与另外两种叶绿体基因组已公布的鼠尾草属物种的叶绿体基因组的基因组结构、重复序列和IR区长度方面进行了比较分析。最后，我们对16种被子植物的叶绿体全基因组进行了系统发育分析。这些工作为构建鼠尾草属物种的叶绿体基因组数据库提供了额外的信息，将有助于鼠尾草属物种的鉴定，并为其进化起源提供新的依据。

2. 结果

2.1. 叶绿体基因组结构

我们分别获得了约 5.9×10^9 个和 1.372×10^{10} 个甘西

鼠尾草和戟叶鼠尾草的原始数据，以及约 5.2×10^9 个和 1.343×10^{10} 个修整数据。两种鼠尾草——甘西鼠尾草（登录号：MH603953）和戟叶鼠尾草（登录号：MH603954）的叶绿体基因组组装结果已在NCBI上提交。二者的叶绿体基因组总长度分别为151319 bp和151547 bp。两种鼠尾草的叶绿体基因组都具有典型的四段式结构，与大多数有花植物一样，由一对IR（50982 bp和51098 bp）和两个单拷贝区（LSC：82732 bp和82853 bp；SSC：17605 bp和17596 bp）组成，如图1和表1所示。

在4种鼠尾草序列中，甘西鼠尾草的叶绿体基因组最短，而戟叶鼠尾草的叶绿体基因组长度在丹参和鼠尾草之间。4个物种的叶绿体基因组长度从151319到53995 bp不等，而LSC区长度从82695到84573 bp不等。IR区的长度范围是50982~51832 bp，SSC区长度范围是17555~17605 bp。甘西鼠尾草和戟叶鼠尾草叶绿体基因组的GC含量分别为37.96%和37.99%，IR区的GC含量（43.11%和43.12%）高于LSC区（36.08%和36.12%）和SSC区（31.88%和31.89%）。一般，IR区有4个高GC含量的核糖体RNA（rRNA）基因[32]。甘西鼠尾草和戟叶鼠尾草叶绿体基因组的AT含量高于GC；另外两种鼠尾草和其他陆生植物的叶绿体基因组也是如此[33–36]。

在本研究中，我们在两种鼠尾草叶绿体基因组中注释到134个基因，其中114个基因是独有的，包括80个蛋白编码基因、30个转运RNA（tRNA）基因和4个rRNA基因（图1和补充资料中的Table S1）。有18个基因在IR区，包括7个蛋白编码基因、7个tRNA基因和4个rRNA基因。LSC区共存在61个蛋白编码基因和22个tRNA基因，SSC区有12个蛋白编码基因和1个tRNA基因。1个*ycf1*假基因和1个*rps19*假基因位于两个IR边界区域（补充材料中的Fig. S1）。

本研究中的两种鼠尾草叶绿体基因组中基因的数量和类型与另外两种鼠尾草相同。4种鼠尾草叶绿体基因组的编码序列（CDS）中的核苷酸组成有微小的差异（表2）。本研究中的两种鼠尾草CDS区，第一、第二、第三密码子位点的GC含量分别为45.75%/44.91%、38.19%/38.49%和30.32%/30.89%（表2）。在另外两种鼠尾草中，第一、第二、第三密码子位点的GC含量分别是45.77%/45.73%、38.21%/38.13%和30.38%/30.17%。从4种鼠尾草的叶绿体基因组中可以看出，第三密码子位点具有AT偏好性，这个特征在其他被子植物中也有所体现[37–40]。

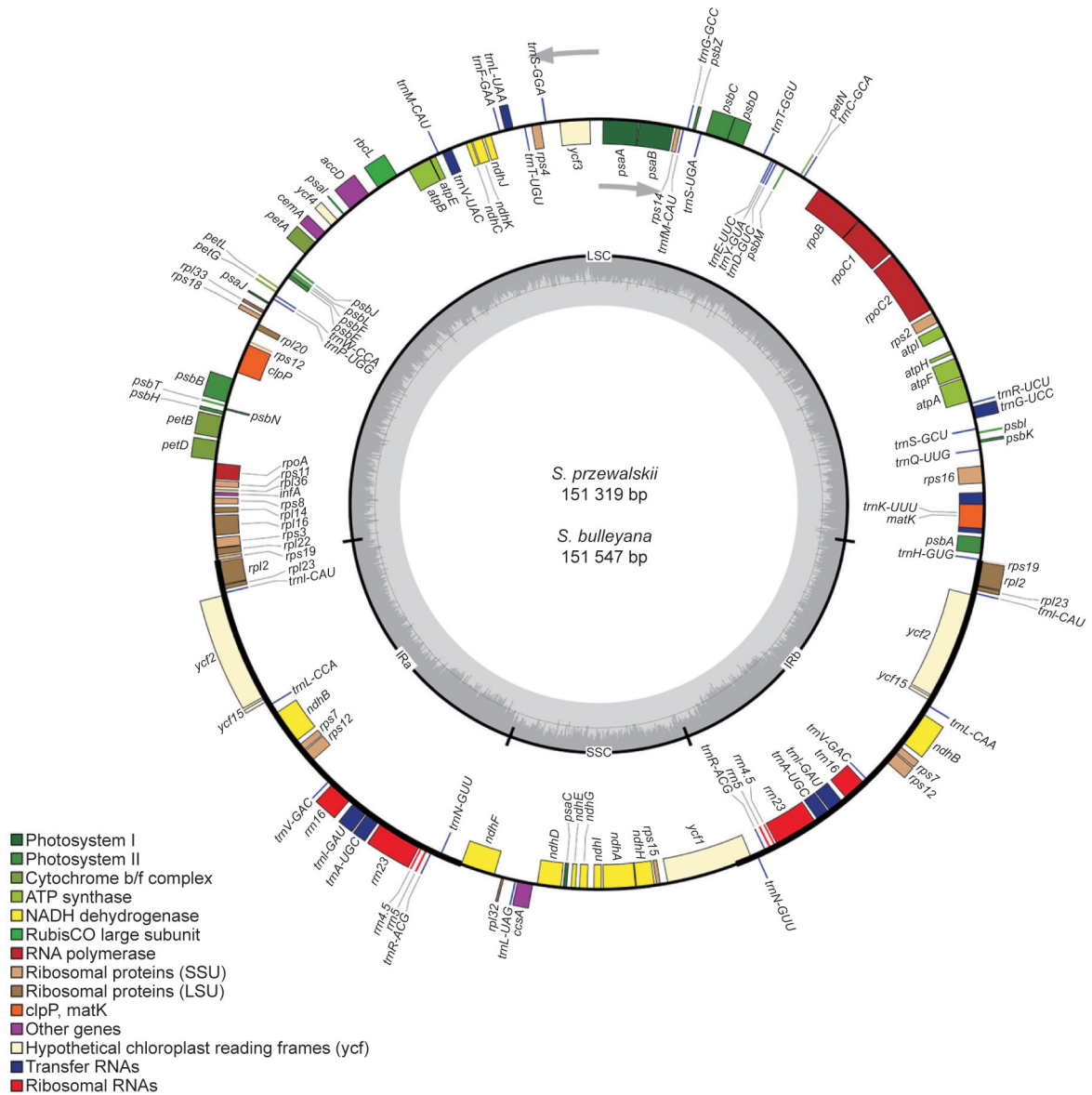


图1. 甘西鼠尾草和戟叶鼠尾草叶绿体基因组的基因图谱。外圆内外的基因按照顶部内外灰色箭头的方向转录。基因按其生物学功能分为14组，以不同颜色的方框显示。在内圆中，深灰色表示GC含量，浅灰色表示AT含量。

表1 4种鼠尾草叶绿体基因组碱基组成

Species	Total cp genome size (bp)	LSC region (bp)	IR region (bp)	SSC region (bp)	GC content (%)	GC content in LSC region (%)	GC content in IR region (%)	GC content in SSC region (%)
<i>S. przewalskii</i>	151 319	82 732	50 982	17 605	37.96	36.08	43.11	31.88
<i>S. bulleyana</i>	151 547	82 853	51 098	17 596	37.99	36.12	43.12	31.89
<i>S. miltiorrhiza</i>	151 328	82 695	51 078	17 555	36.20	36.20	43.10	32.00
<i>S. japonica</i>	153 995	84 573	51 832	17 590	36.24	36.24	43.03	31.67

甘西鼠尾草和戟叶鼠尾草叶绿体基因组中共有87个蛋白编码基因，分别有26 439个和26 432个密码子，而丹参和鼠尾草分别有26 483个和26 485个密码子（补充材料中的Table S2）。其中，在甘西鼠尾草和戟叶鼠尾草

中使用最多的是编码异亮氨酸(Ile)的AUU(1096/1106)，使用最少的是编码半胱氨酸的UGC(76/114)。在另外两个鼠尾草叶绿体基因组中，使用最多和最少的也分别是AUU和UGC(1100/1102和70/71)。4种叶绿体基因

表2 4种鼠尾草叶绿体基因组的基因数及CDS核苷酸组成

Species	Number of		ATCG contents in CDS (%)				ATCG contents in 1st position (%)				ATCG contents in 2nd position (%)				ATCG contents in 3rd position (%)				
	unique genes	protein-coding genes	tRNAs		rRNAs		T	C	A	G	T	C	A	G	T	C	A	G	
			30	4	30	4													
<i>S. przewalskii</i>	114	80	30	4	17.80	30.58	20.29	23.71	18.98	30.54	26.77	32.56	20.38	29.24	17.81	37.73	14.03	31.96	16.29
<i>S. bulleyana</i>	114	80	30	4	17.80	30.56	20.30	24.25	18.84	30.83	26.07	32.26	20.22	29.25	18.27	37.51	14.34	31.59	16.55
<i>S. miltiorrhiza</i>	114	80	30	4	17.82	30.54	20.30	23.69	19.01	30.54	26.76	32.56	20.40	29.23	17.81	37.79	14.05	31.84	16.33
<i>S. japonica</i>	114	80	30	4	17.77	30.60	20.24	23.68	18.98	30.59	26.75	32.68	20.35	29.19	17.78	37.81	13.99	32.02	16.18

组的同义密码子相对使用度 (RSCU) 略有不同。除了甲硫氨酸和色氨酸外, 所有氨基酸都由多个密码子编译 (图2)。精氨酸 (Arg)、亮氨酸 (Leu) 和丝氨酸 (Ser) 有6个同义密码子; 缬氨酸 (Val)、脯氨酸 (Pro)、苏氨酸 (Thr)、丙氨酸 (Ala) 和甘氨酸 (Gly) 有4个同义密码子; Ile和终止密码子都有3个同义密码子; 剩下的每个氨基酸都有两个同义密码子。同义密码子通常只在第三位点发生突变, 从而减少有害突变。

内含子在基因表达调控中起着至关重要的作用, 可以增强外源基因在植物各部位上的表达, 从而产生理想的农艺性状[41,42]。在4个鼠尾草叶绿体基因组中共观察到18个含有内含子的基因, 其中15个含有1个内含子, 3个含有2个内含子。*trnK-UUU*基因的内含子最长, 并且包含*matK*基因。*rps12*基因的5'端位于LSC区, 3'端位于IR区, 为反式剪接基因。除了*ndhB*基因 (补充材料中的Table S3) 外, 4种鼠尾草的其他外显子长度都是保守的。

将4种鼠尾草的18个基因与其他4种同科物种——欧薄荷、罗勒、紫苏、黄芩——进行比较。观察到8个唇形科物种的叶绿体基因组的外显子长度存在微小差异; 部分基因仅在4种鼠尾草和欧薄荷中是保守的。*trnV-UAC*、*rpoC1*和*ycf3*基因的内含子相位在8种唇形科植物叶绿体基因组中是不同的。*rps16*基因的外显子长度在4种鼠尾草中是特定的。此外, 还观察到基因的内含子长度比外显子长度差异更大, 虽然一些基因的内含子长度, 如*trnL-UAA*和*rpl2*, 在除鼠尾草之外的3种鼠尾草中都一样。

RNA编辑参与质体转录调控, 这可以丰富蛋白多样性[43-45]。本研究预测了4种鼠尾草叶绿体基因组的35个基因的潜在RNA编辑位点。共预测了43个RNA编辑位点; 其中37个为这4个物种的共同位点 (补充材料中的Table S4)。在这35个基因中, 有16个基因未检测到潜在的RNA编辑位点, 包括*atpA*、*atpB*、*clpP*、*petD*、*petG*、*petL*、*psaB*、*psaI*、*psbB*、*psbE*、*psbF*、*psbL*、*rpl23*、*rpoC1*、*rps8*和*ycf3*。*rps16*基因在鼠尾草中并未检测到的潜在RNA编辑位点, 但在其他3个种鼠尾草中检测到一个RNA编辑位点。在43个潜在RNA编辑位点中, 11个位于密码子的首位, 32个位于第二位, 第三位未发现潜在的RNA编辑位点, 且碱基转换类型均为C到T。这个结果与其他陆生植物相似[46,47]。从Ser到Leu的氨基酸转化最频繁, 而从Pro到Ser和Thr到Ile的转化最少。

2.2. 重复序列和简单重复序列标记 (SSR) 分析

利用REPuter进行重复序列分析, 在两种鼠尾草中发现了43个重复, 包括21个正向重复和22个回文重复 (图3、补充材料中的Table S5)。大多数重复序列位

于 $ycf2$ 编码区; 4个存在于 $ycf3$ 内含子区和 $rps12$ 与 $trnV-GAC$ 的基因间隔区 (IGS) 之间, 4个存在于 $rrn4.5$ 和 $rrn5$ 的IGS中。在甘西鼠尾草和戟叶鼠尾草中分别检测到17个和21个串联重复 (图3、补充材料中的Table

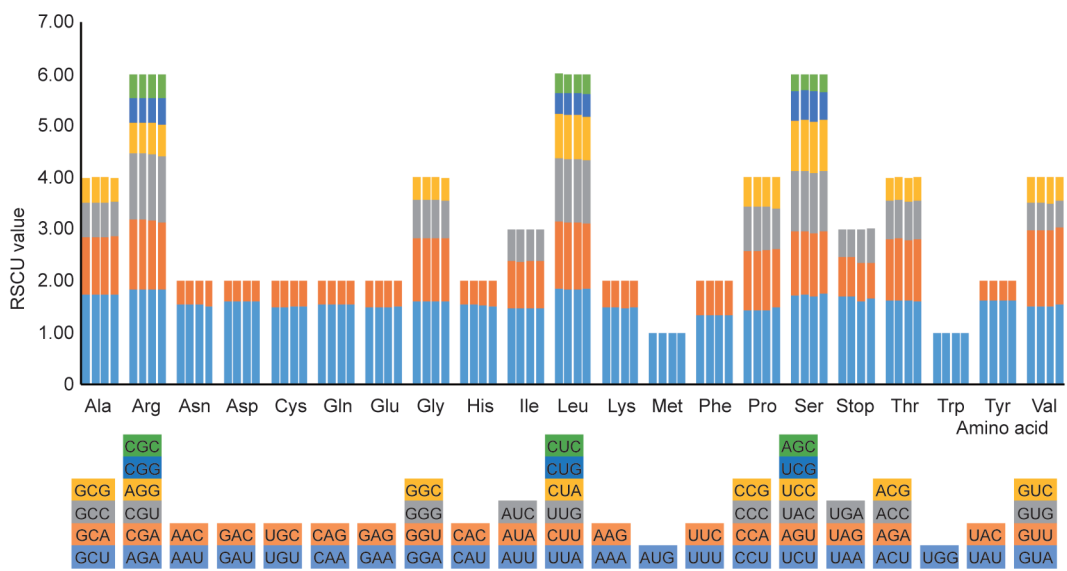


图2. 4种鼠尾草叶绿体基因组中CDS的密码子含量。横坐标表示20个氨基酸和终止子, 纵坐标表示RSCU值。对于每种氨基酸, 从左至右对应的物种分别为: 甘西鼠尾草、戟叶鼠尾草、丹参、鼠尾草。每个氨基酸的不同颜色对应于下面相同颜色的密码子。Asn: 天冬酰胺; Asp: 天门冬氨酸; Cys: 半胱氨酸; Gln: 谷氨酰胺; Glu: 谷氨酸; His: 组氨酸; Lys: 赖氨酸; Met: 甲硫氨酸; Phe: 苯丙氨酸; Trp: 色氨酸; Tyr: 酪氨酸。

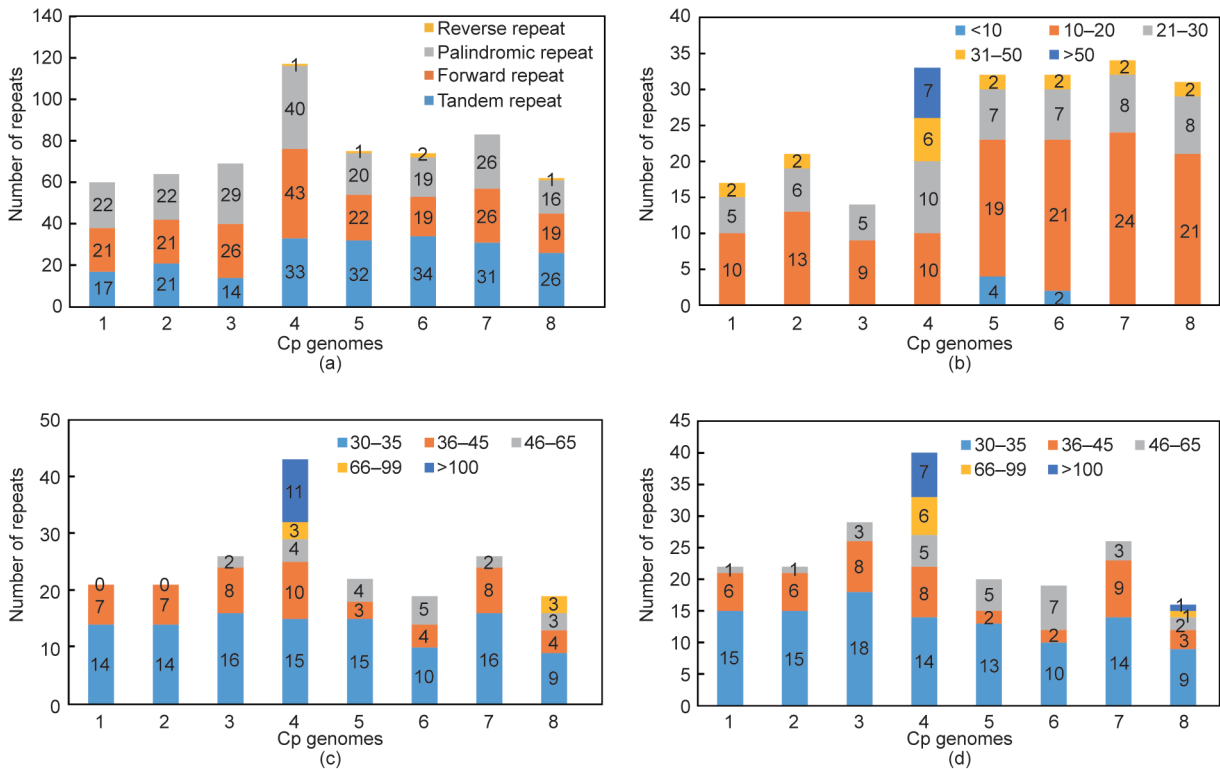


图3. 8个叶绿体基因组的重复序列分析。(a) 8个叶绿体基因组的重复类型; (b) 8个叶绿体基因组的串联重复; (c) 8个叶绿体基因组的正向重复; (d) 8个叶绿体基因组的回文重复。(a)中, 不同的颜色表示不同的重复类型; 在(b-d)中, 不同的颜色表示不同的长度。纵坐标表示重复的次数; 在横坐标中, 数字1~8表示如下: 1为甘西鼠尾草, 2为戟叶鼠尾草, 3为丹参, 4为鼠尾草, 5为欧薄荷, 6为罗勒, 7为紫苏, 8为黄芩。

S6)。这些重复中约有一半位于 $ycf1$ 和 $ycf2$ 基因中，另一半位于IGS区。两个最长的重复约长90 bp，存在于两种鼠尾草的 $ycf2$ 基因中。

对8种唇形科植物叶绿体基因组的重复序列进行对比分析，发现甘西鼠尾草和戟叶鼠尾草与其他叶绿体基因组在重复类型上有相似之处，而鼠尾草则比另外几种拥有更多的长片段重复。8种唇形科植物叶绿体基因组的串联重复长度在10~30 bp之间，而正向和回文重复长度大多在30~45 bp之间。

SSR广泛分布于基因组的不同位置[48]。叶绿体基因组具有单亲遗传的特点，SSR在同一物种内具有较高的变异水平。因此，叶绿体SSR作为分子标记被广泛应用于遗传图谱构建、目标基因校准和定位等方面的研究[49–51]。在甘西鼠尾草叶绿体基因组中共观察到178个SSR，包括134个单核苷酸、35个二核苷酸、7个四核苷酸和2个五核苷酸SSR；在戟叶鼠尾草叶绿体基因组中共观察到177个SSR，包括136个单核苷酸、32个二核苷酸和9个四核苷酸SSR（图4、补充材料中的Table S7）。SSR的主要类型是单核苷酸；绝大多数的单核苷酸SSR都是A/T型SSR（132/129）。

对8个叶绿体基因组SSR的比较表明，8个叶绿体基因组是相似的。在8个叶绿体基因组中，大部分SSR由单核苷酸和二核苷酸重复单元构成。单核苷酸重复数量从109（罗勒）到136（戟叶鼠尾草）不等，二核苷酸重复数量从25（丹参）到39（紫苏）不等。而3次或以上的寡核苷酸重复数量相对较少，但类型丰富（补充材料中的Table S7）。

2.3. 同步比较和 SNP 分析

以带注释的甘西鼠尾草叶绿体基因组为参考，利用mVISTA对甘西鼠尾草和戟叶鼠尾草与另外两种鼠尾草的叶绿体基因组进行配对比对（补充材料中的Table S2）。对4种鼠尾草叶绿体基因组和80个基因编码区的变异位点进行分析（图5、补充材料中的Table S8）。对4个叶绿体基因组的比对和单核苷酸多态性（SNP）分析表明，IR区比LSC和SSC区更为保守，在后两个区域中存在更多的变异。非编码区域比编码区域具有更多的变异。在这些叶绿体基因组中，IGS具有更多的变异位点，包括LSC区域的 $rps16-trnQ$ 、 $trnG-trnS$ 、 $atpH-atpI$ 、 $psbA-ycf3$ 、 $ycf4-cenA$ 、 $thrnV-ndhC$ ，SSC区域的 $ycf1-rps15$ 、 $rpl32-trnL$ 、 $ndhI-ndhG$ 。在4个叶绿体基因组的 $rpoC2$ 、 $ndhF$ 和 $ycf1$ 基因的编码区域也发现了一些变异位点。与鼠尾草相比，甘西鼠尾草和戟叶鼠尾草与丹参更为相似。

2.4. 系统发育分析

基于8种唇形科植物、4种茄科植物和另外3种菊目植物的叶绿体基因组，使用最大似然（ML）法，以拟南芥为外类群，构建系统发育树（图6）。在ML树中，所有分支具有100%的自展值，并且4种鼠尾草形成一个稳健的单系分支。甘西鼠尾草和戟叶鼠尾草聚集在一个末端分支，代表弧隔鼠尾草亚属。这8个唇形科植物都聚集在一个单系类群中，并嵌套在芝麻（芝麻科）、旋蒴苣苔（苦苣苔科）和木犀榄（橄榄科）的分支中。这个结果与之前的报告的结果一致[30]。

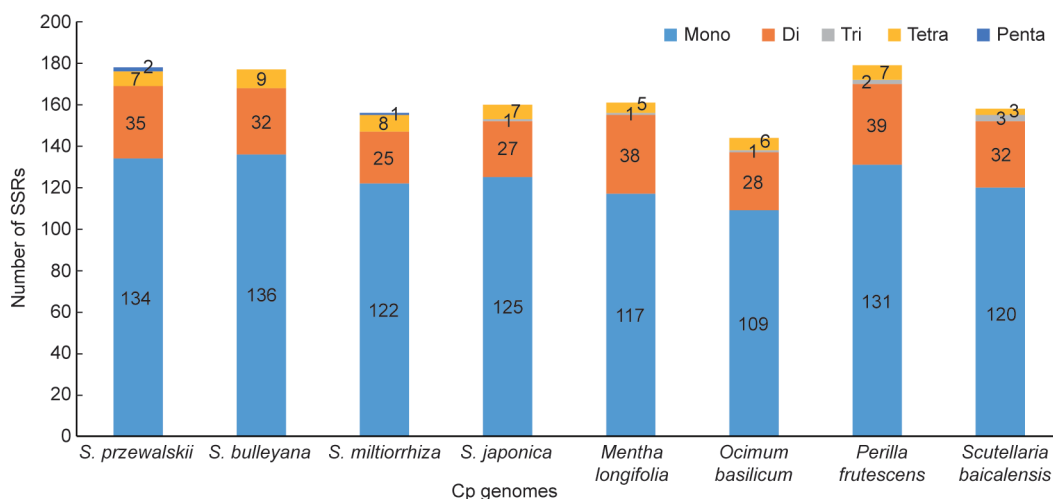


图4. 8个叶绿体基因组的SSR分析。纵坐标表示SSR的数量。

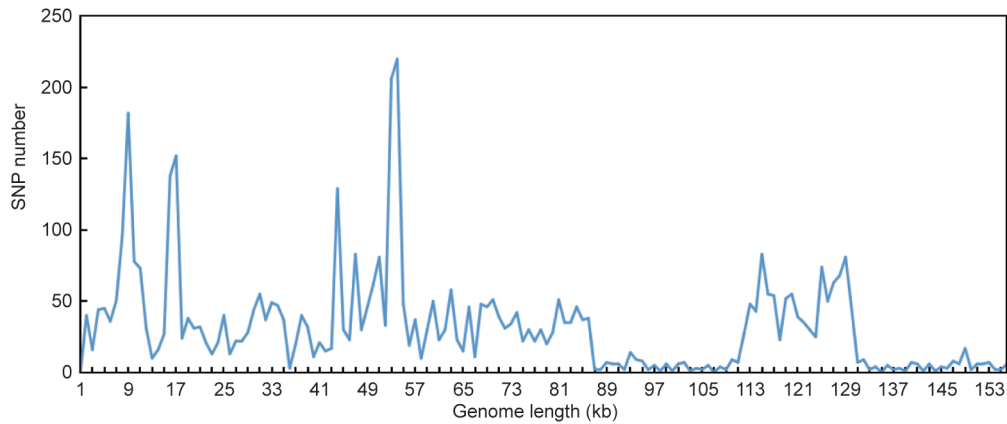


图5. 4个叶绿体基因组中SNP的数据。

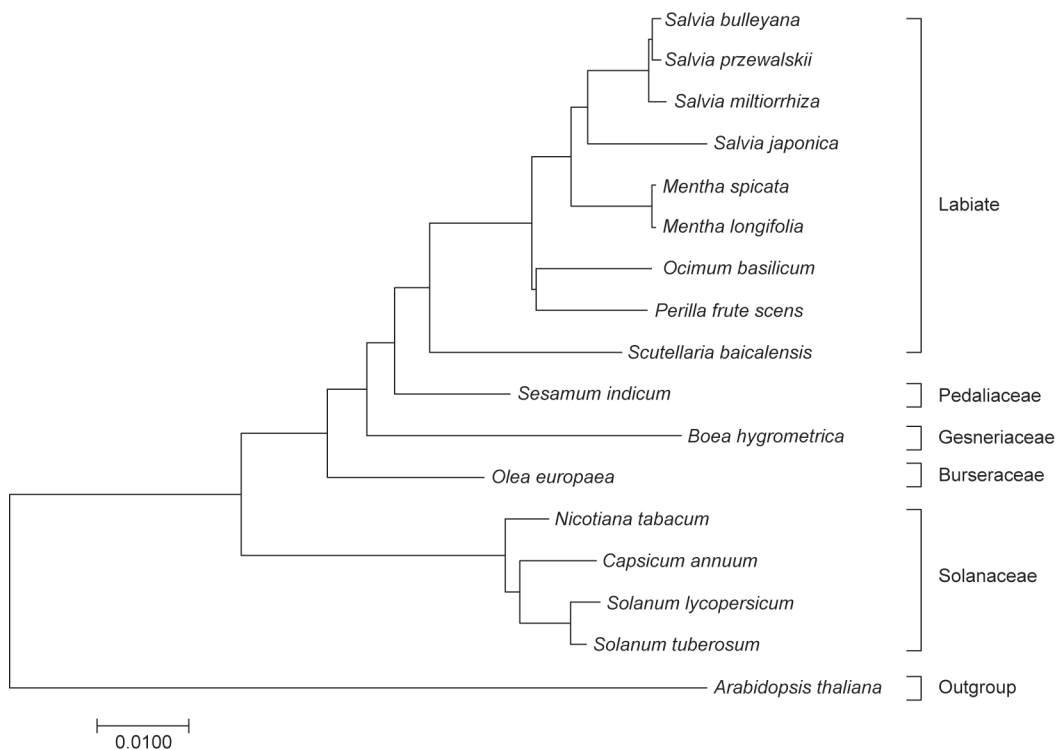


图6. 包含16种植物叶绿体基因组的ML系统发育树重建。拟南芥设为外类群。

3. 方法

3.1. 植物材料和 DNA 提取

甘西鼠尾草采自云南省丽江市玉龙县，戟叶鼠尾草采自云南省大理白族自治州宾川县乌龙坝村。甘西鼠尾草的全基因组DNA是利用植物DNA提取试剂盒（德国QIAGEN股份有限公司）从100 mg新鲜叶片中提取的，而戟叶鼠尾草的全基因组DNA是利用改良的十六烷基三甲基溴化铵（CTAB）法，从100 mg新鲜叶片中

提取的[52]。基因组DNA的质量和浓度用琼脂糖凝胶电泳和NanoDrop 2000c分光光度计（美国赛默飞世尔科技有限公司）进行评价。合格的DNA用于文库构建。用Illumina HiSeq 1500平台（美国Illumina公司）测序。

3.2. 两种鼠尾草的组装与注释

从全基因组序列中提取叶绿体基因组。首先，使用FastQC对原始数据（reads）进行评估，并使用Trimmomatic进行裁剪，去除低质量碱基[$Q < 30$, $Q =$

-10 lg P 型误差, $Q < 30$ 意味着错误率 < 0.001 , 长度 < 50]和接头序列[53,54]。接下来, 以丹参(JX312195)的叶绿体基因组作为参考序列, 使用BLAST从修剪过的数据中提取叶绿体基因组的数据[55]。最后, 使用SOAPdenovo将叶绿体的数据进行基因组组装[56]。通过对不同 k -mer (k -mer的意思是将数据分成含有 k 个碱基的字符串, 然后进行组装)值下重叠群(contig)的长度和数量比较, 发现 k -mer大小分别为127和77时, 甘西鼠尾草和戟叶鼠尾草的结果较好。这两个参数的重叠群用于最终的组装。使用SSPACE和GapCloser获得支架(scaffolds)并填充缺口(gap)[56,57]。使用CPGAVAS进行叶绿体基因组注释[58]。从CPGAVAS获得的注释结果要用Apollo对起始密码子、终止密码子和内含子、外显子边界进行手动修正[59]。tRNAscanSE用来识别tRNA基因[60]。使用OGDRAW得到这两种鼠尾草叶绿体基因组的圈图[61]。利用MEGA6分析密码子使用情况和叶绿体基因组特征[62]。使用预测RNA编辑器PREP suite预测两个叶绿体基因组中基因编码蛋白的RNA编辑位点[63], 其cut-off值为0.8。

3.3. 重复序列和 SSR 的表征

使用PERuter进行正向、反向和回文3种类型的重复序列的检测, 其中汉明距离设为3和最小重复长度设为30 bp[64]。使用Tandem Repeats Finder分析串联重复序列[65]。使用MISA[66]检测SSR, 其阈值如下: 单核苷酸SSR有8个重复单位, 二核苷酸和三核苷酸重复SSR有4个重复单位, 四核苷酸、五核苷酸和六核苷酸重复SSR有3个重复单位[15]。

3.4. 基因组比较

在Shuffle-LAGAN模式下使用mVISTA[67]将甘西鼠尾草和戟叶鼠尾草的叶绿体基因组与另外两种鼠尾草叶绿体基因组进行比较(即丹参[JX312195]和鼠尾草[NC_035233])。

3.5. 系统发育分析

从NCBI细胞器基因组和核苷酸资源数据库中下载共14个叶绿体全基因组序列进行系统发育分析。用MAFFT软件进行叶绿体基因组的比对, 用MEGA6[68]基于ML方法构建系统发育树。通过1000个进化树对分重接(TBR)分支交换进行自展分析, 并将拟南芥设为外类群。

4. 讨论

4.1. 叶绿体基因组结构

在本研究中, 我们首次报道了甘西鼠尾草和戟叶鼠尾草的叶绿体全基因组。两个物种的叶绿体基因组长度范围是151 319~151 547 bp, 具有典型的四段式结构。这两个叶绿体基因组共编码114个单基因, 包括80个蛋白编码基因、30个tRNA基因和4个rRNA基因。在这些基因中, *ycf2*基因的CDS最大, 而*petN*基因的CDS最小。*rps12*基因的5'端位于LSC区域, 3'端位于IR区, 这种现象在被子植物叶绿体基因组中比较常见。综上所述, 甘西鼠尾草和戟叶鼠尾草的叶绿体基因组结构、GC含量、基因数量和类型、密码子的使用与另外两种鼠尾草的叶绿体基因组相似。

4.2. 基因对照

某些基因, 如*matK-trnK*、*atpB-atpE*、*psbC-psbD*和*rps3-rpl22*, 在8种唇形科植物叶绿体基因组中相互重叠。通过重叠长度的比较可知, *atpB-atpE*、*psbC-psbD*和*rps3-rpl22*基因的重叠长度相同, 而黄芩中*rps3*和*rpl22*基因无重叠。在甘西鼠尾草、戟叶鼠尾草和丹参中, *matK*基因长度(1563 bp)相同, 但有12个变异位点。*matK*基因的长度(1530 bp)在鼠尾草、欧薄荷和黄芩中是相同的。*matK*基因在2009年被生命条码联盟选为陆生植物条形码[69]。先前有研究发现, *matK*基因相较于内部的转录间隔区(*ITS*)基因能更好地识别唇形科的不同物种[70]。

这里, 我们在8种唇形科植物中发现了两个不同的*ndhB*基因类型; 一个编码511个氨基酸, 另一个编码493个氨基酸。当根据丹参叶绿体基因组注释*ndhB*基因外显子1的长度时, 两种鼠尾草叶绿体基因组中的*ndhB*基因外显子1的长度比丹参(JX312195)短54 bp, 在丹参的叶绿体基因组第8个碱基位置A取代了G。这一现象导致终止子的出现, 这与紫苏(NC_030756)*ndhB*基因的情况一致。该位点经聚合酶链反应(PCR)被证实, 正向引物为AACAAACGAAAAGGAAACG, 反向引物为CTCCATAGGAACAATAGGG。先前研究报道, 蕨类*ndhB*基因的两个外显子具有独特的基因内拷贝数变异模式[71]。虽然*rps16*基因在不同的豆科植物谱系中功能缺失, 但有报告显示, 即使在异养条件下, *rps16*基因也是必不可少的, 但可以被核基因功能取代[72,73]。在本研究中, 我们发现该基因长度在4种鼠尾草叶绿体基

因组中是保守的，且在该基因CDS中仅观察到3个变异位点（补充材料中的Table S8）。

4.3. 属内差异

Hebert等[74]最早提出的DNA条形码技术，可以通过*ITS2*、*matK*、*psbA-trnH*、*rbcL*等DNA序列进行物种识别。然而，近缘种——特别是同一属中形态相近的物种——鉴定仍然存在一些困难。因此，为这些物种找到合适的DNA标记是至关重要的。叶绿体基因组与核基因组相比进化较慢，因此常被用于系统发育研究和物种鉴定[75]。本研究通过对4种鼠尾草叶绿体基因组的配对比对和SNP分析发现，*trnV-ndhC*、*trnQ-rps16*、*atpI-atpH*、*psbA-ycf3*、*ycf1*、*rpoC2*、*ndhF*、*matK*、*rpoB*、*rpoA*和*accD*基因的IGS中变异位点数量较多。因此，这些区域可以作为新的候选片段来鉴定鼠尾草属物种。研究显示，*ycf1a*或*ycf1b*是最易变的质体基因组区域，可以作为陆生植物的核心条形码[76]。然而，这还需要更多的鼠尾草叶绿体序列数据的支持，这一问题在未来的研究中应该加以解决。

4.4. 4种鼠尾草的亲缘关系

甘西鼠尾草和戟叶鼠尾草都属于弧隔鼠尾草亚属，而丹参属于荔枝草亚属。鼠尾草属于鼠尾草亚属。这种遗传关系反映在系统发育树中，甘西鼠尾草和戟叶鼠尾草聚集在一个终端分支上。在外观和特征方面，许多研究表明甘西鼠尾草和戟叶鼠尾草具有相似性[77,78]。成分上，甘西鼠尾草和戟叶鼠尾草含有丹参的大部分化学成分，常被作为丹参的替代品；相比之下，鼠尾草只含有部分与丹参相同的成分[79,80]。在系统发育树中，甘西鼠尾草和戟叶鼠尾草先与丹参，然后与鼠尾草，形成了一个稳健的单系分支。甘西鼠尾草和戟叶鼠尾草基因组之间共存在297个SNP位点；在甘西鼠尾草、戟叶鼠尾草和丹参基因组之间，这个数据是785。而4种鼠尾草叶绿体基因组间有4982个变异位点。从重复次数、重复类型、IR区域边界、系统发育树等分析表明，甘西鼠尾草和戟叶鼠尾草亲缘关系密切，与鼠尾草的亲缘关系比与丹参的更远。

5. 结论

本文报道了我国西南地区长期以来被作为丹参替代品的甘西鼠尾草和戟叶鼠尾草的叶绿体全基因组。我们

还将这两个叶绿体基因组与另外两种公布在NCBI上的鼠尾草属物种的叶绿体基因组进行了比较。本文研究的两种鼠尾草的基因序列和基因组结构与已公布的两种鼠尾草叶绿体基因组相似。此外，将含有SSR的重复序列与其他6个叶绿体基因组进行比较，发现了细微的区别。变异位点、IR区域边界和系统发育分析结果表明，甘西鼠尾草和戟叶鼠尾草之间，比与另两个物种之间的相似度更大。本文的数据和分析可能为深入了解鼠尾草属内各物种间的系统发育关系提供依据，且叶绿体全基因组可能对今后的育种和进一步的生物学发现有帮助。

致谢

本课题由国家自然科学基金（QFSL2018004、2017YFC1702100、81741060）、中央级公益性科研院所基本科研业务费专项资金（ZXKT17004）资助。

作者贡献

陈士林和徐江构思并设计研究框架；段宝忠、马维思对样品进行采集和鉴定；梁从莲和王磊进行实验；梁从莲、沈晓凤、郭帅分析数据；梁从莲和王磊撰写论文。肖水明、齐海军、王振、刘尧奇润色语言，胡灏禹、雷隽、徐江对最终稿进行修改。所有的作者都阅读并确认了最终的手稿。

Compliance with ethics guidelines

Conglian Liang, Lei Wang, Juan Lei, Baozhong Duan, Weisi Ma, Shuiming Xiao, Haijun Qi, Zhen Wang, Yaoqi Liu, Xiaofeng Shen, Shuai Guo, Haoyu Hu, Jiang Xu, and Shilin Chen declare that they have no conflict of interest or financial conflicts to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.eng.2019.01.017>.

References

- [1] Xiao XH, Fang QM, Xia WJ, Yin GP, Su ZW, Qiao CZ. Numerical taxonomy of

- medicinal *Salvia* L. and the genuineness of Danshen. *J Plant Resour Environ* 1997;6(2):17–21. Chinese.
- [2] Wang Y, Li D, Zhang Y. Analysis of ITS sequences of some medicinal plants and their related species in *Salvia*. *Yao Xue Xue Bao* 2007;42(12):1309–13. Chinese.
 - [3] Li ZM, Xu SW, Liu PQ. *Salvia miltiorrhiza* Burge (Danshen): a golden herbal medicine in cardiovascular therapeutics. *Acta Pharmacol Sin* 2018;39(5):802–24.
 - [4] Wang L, Ma R, Liu C, Liu H, Zhu R, Guo S, et al. *Salvia miltiorrhiza*: a potential red light to the development of cardiovascular diseases. *Curr Pharm Des* 2017;23(7):1077–97.
 - [5] Chen W, Chen GX. Danshen (*Salvia miltiorrhiza* Bunge): a prospective healing sage for cardiovascular diseases. *Curr Pharm Des* 2017;23(34):5125–35.
 - [6] Chinese Pharmacopoeia Commission. Pharmacopoeia of the People's Republic of China. 2015 ed. Beijing: People's Medical Publishing; 2015. Appendix 76.
 - [7] Ren H, Hu X, Liu Y, Dai D, Liu X, Wang Z, et al. *Salvia przewalskii* extract of total phenolic acids inhibit TLR4 signaling activation in podocyte injury induced by puromycin aminonucleoside *in vitro*. *Ren Fail* 2018;40(1):273–9.
 - [8] Yang Y, Wang ZP, Gao SH, Ren HQ, Zhong RQ, Chen WS. The effects of *Salvia przewalskii* total phenolic acid extract on immune complex glomerulonephritis. *Pharm Biol* 2017;55(1):2153–60.
 - [9] Wang HQ, Yang LX, Chen XY, Yang PF, Chen RY. Chemical constituents from *Salvia przewalskii* root. *J Chin Med Mater* 2015;38(6):1197–201. Chinese.
 - [10] Liu X, Liu Y, Yang Y, Xu J, Dai D, Yan C, et al. Antioxidative stress effects of *Salvia przewalskii* extract in experimentally injured podocytes. *Nephron* 2016;134(4):253–71.
 - [11] Li SF, Li JQ, Li XJ, Wang XY, Du PJ, Zhou KP, et al. Pharmacognostical studies on *Salvia buleyana* Diels. *Chin J Ethnomed Ethnopharm* 2008;7:18–22. Chinese.
 - [12] Brunkard JO, Runkel AM, Zambryski PC. Chloroplasts extend stromules independently and in response to internal redox signals. *Proc Natl Acad Sci USA* 2015;112(32):10044–9.
 - [13] Poczar P, Hyvönen J. The complete chloroplast genome sequence of the CAM epiphyte Spanish moss (*Tillandsia usneoides*, Bromeliaceae) and its comparative analysis. *PLoS ONE* 2017;12(11):e0187199.
 - [14] Douglas SE. Plastid evolution: origins, diversity, trends. *Curr Opin Genet Dev* 1998;8(6):655–61.
 - [15] Lin CP, Wu CS, Huang YY, Chaw SM. The complete chloroplast genome of *Ginkgo biloba* reveals the mechanism of inverted repeat contraction. *Genome Biol Evol* 2012;4(3):374–81.
 - [16] Mo JS, Kim K, Lee MH, Lee JH, Yoon UH, Kim TH. The complete chloroplast genome sequence of *Perilla citriodora* (Makino) Nakai. *Mitochondrial DNA Part A* 2017;28(1):131–2.
 - [17] Wicke S, Schneeweiss GM, DePamphilis CW, Müller KF, Quandt D. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol Biol* 2011;76(3–5):273–97.
 - [18] Wang MX, Cui LC, Feng KW, Deng PC, Du XH, Wan FH, et al. Comparative analysis of Asteraceae chloroplast genomes: structural organization, RNA editing and evolution. *Plant Mol Biol Report* 2015;33(5):1526–38.
 - [19] Raubeson LA, Peery R, Chumley TW, Dziubek C, Fourcade HM, Boore JL, et al. Comparative chloroplast genomics: analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. *BMC Genomics* 2007;8(1):174.
 - [20] Wang RJ, Cheng CL, Chang CC, Wu CL, Su TM, Chaw SM. Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. *BMC Evol Biol* 2008;8(1):36.
 - [21] Zhu A, Guo W, Gupta S, Fan W, Mower JP. Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. *New Phytol* 2016;209(4):1747–56.
 - [22] Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S. Plant DNA barcoding: from gene to genome. *Biol Rev Camb Philos Soc* 2015;90(1):157–66.
 - [23] Daniell H, Lin CS, Yu M, Chang WJ. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol* 2016;17(1):134.
 - [24] Cheng L, Li HP, Qu B, Huang T, Tu JX, Fu TD, et al. Chloroplast transformation of rapeseed (*Brassica napus*) by particle bombardment of cotyledons. *Plant Cell Rep* 2010;29(4):371–81.
 - [25] Svab Z, Hajdukiewicz P, Maliga P. Stable transformation of plastids in higher plants. *Proc Natl Acad Sci USA* 1990;87(21):8526–30.
 - [26] R'epková J. Potential of chloroplast genome in plant breeding. *Czech J Genet Plant Breed* 2010;46(3):103–13.
 - [27] Chen S, Xu J, Liu C, Zhu Y, Nelson DR, Zhou S, et al. Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. *Nat Commun* 2012;3(1):913.
 - [28] Huang ZH, Xu J, Xiao SM, Liao BS, Gao Y, Zhai CC, et al. Comparative optical genome analysis of two pangolin species: *Manis pentadactyla* and *Manis javanica*. *GigaScience* 2016;5(1):1–5.
 - [29] Xu J, Chu Y, Liao B, Xiao S, Yin Q, Bai R, et al. *Panax ginseng* genome examination for ginsenoside biosynthesis. *GigaScience* 2017;6(11):1–15.
 - [30] Qian J, Song J, Gao H, Zhu Y, Xu J, Pang X, et al. The complete chloroplast genome sequence of the medicinal plant *Salvia miltiorrhiza*. *PLoS ONE* 2013;8(2):e57607.
 - [31] Xu H, Song J, Luo H, Zhang Y, Li Q, Zhu Y, et al. Analysis of the genome sequence of the medicinal plant *Salvia miltiorrhiza*. *Mol Plant* 2016;9(6):949–52.
 - [32] He Y, Xiao H, Deng C, Xiong L, Yang J, Peng C. The complete chloroplast genome sequences of the medicinal plant *Pogostemon cablin*. *Int J Mol Sci* 2016;17(6):820.
 - [33] Li XW, Gao HH, Wang YT, Song JY, Henry R, Wu HZ, et al. Complete chloroplast genome sequence of *Magnolia grandiflora* and comparative analysis with related species. *Sci China Life Sci* 2013;56(2):189–98.
 - [34] Shen X, Wu M, Liao B, Liu Z, Bai R, Xiao S, et al. Complete chloroplast genome sequence and phylogenetic analysis of the medicinal plant *Artemisia annua*. *Molecules* 2017;22(8):1330.
 - [35] Jiang D, Zhao Z, Zhang T, Zhong W, Liu C, Yuan Q, et al. The chloroplast genome sequence of *Scutellaria baicalensis* provides insight into intraspecific and interspecific chloroplast genome diversity in *Scutellaria*. *Genes* 2017;8(9):227.
 - [36] Vining KJ, Johnson SR, Ahkami A, Lange I, Parrish AN, Trapp SC, et al. Draft genome sequence of *Mentha longifolia* and development of resources for mint cultivar improvement. *Mol Plant* 2017;10(2):323–39.
 - [37] Guo H, Liu J, Luo L, Wei X, Zhang J, Qi Y, et al. Complete chloroplast genome sequences of *Schisandra chinensis*: genome structure, comparative analysis, and phylogenetic relationship of basal angiosperms. *Sci China Life Sci* 2017;60(11):1286–90.
 - [38] Mariotti R, Cultrera NG, Díez CM, Baldoni L, Rubini A. Identification of new polymorphic regions and differentiation of cultivated olives (*Olea europaea* L.) through plastome sequence comparison. *BMC Plant Biol* 2010;10(1):211.
 - [39] Yi DK, Kim KJ. Complete chloroplast genome sequences of important oilseed crop *Sesamum indicum* L. *PLoS ONE* 2012;7(5):e35872.
 - [40] Lu C, Shen Q, Yang J, Wang B, Song C. The complete chloroplast genome sequence of Safflower (*Carthamus tinctorius* L.). *Mitochondrial DNA, Part A* 2016;27(5):3351–3.
 - [41] Wu JY, Xiao JF, Wang LP, Zhong J, Yin HY, Wu SX, et al. Systematic analysis of intron size and abundance parameters in diverse lineages. *Sci China Life Sci* 2013;56(10):968–74.
 - [42] Xu J, Feng D, Song G, Wei X, Chen L, Wu X, et al. The first intron of rice EPSP synthase enhances expression of foreign gene. *Sci China Life Sci* 2003;46(6):561.
 - [43] Yan J, Zhang Q, Yin P. RNA editing machinery in plant organelles. *Sci China Life Sci* 2018;61(2):162–9.
 - [44] Chen C, Bundschuh R. Systematic investigation of insertional and deletional RNA-DNA differences in the human transcriptome. *BMC Genomics* 2012;13(1):616.
 - [45] Knoop V. When you can't trust the DNA: RNA editing changes transcript sequences. *Cell Mol Life Sci* 2011;68(4):567–86.
 - [46] Grennan AK. To thy proteins be true: RNA editing in plants. *Plant Physiol* 2011;156(2):453–4.
 - [47] Lutz KA, Maliga P. Chapter 23—transformation of the plastid genome to study RNA editing. *Methods Enzymol* 2007;424:501–18.
 - [48] Asaf S, Khan AL, Khan MA, Waqas M, Kang SM, Yun BW, et al. Chloroplast genomes of *Arabidopsis halleri* ssp. *gemmifera* and *Arabidopsis lyrata* ssp. *petraea*: structures and comparative analysis. *Sci Rep* 2017;7(1):7556.
 - [49] Yin D, Wang Y, Zhang X, Ma X, He X, Zhang J. Development of chloroplast genome resources for peanut (*Arachis hypogaea* L.) and other species of *Arachis*. *Sci Rep* 2017;7(1):11649.
 - [50] Zhou T, Wang J, Jia Y, Li W, Xu F, Wang X. Comparative chloroplast genome analyses of species in *Gentiana* section *Cruciata* (Gentianaceae) and the development of authentication markers. *Int J Mol Sci* 2018;19(7):1962.
 - [51] Flannery ML, Mitchell FJ, Coyne S, Kavanagh TA, Burke JI, Salamin N, et al. Plastid genome characterisation in *Brassica* and Brassicaceae using a new set of nine SSRs. *Theor Appl Genet* 2006;113(7):1221–31.
 - [52] Shi QH, Yao ZP, Zhang H, Xu L, Dai PH. Comparison of four methods of DNA extraction from chickpea. *J Xinjiang Agric Univ* 2009;1:64–7. Chinese.
 - [53] Kaila T, Chaduvla PK, Rawal HC, Saxena S, Tyagi A, Mishra S, et al. Chloroplast genome sequence of clusterbean (*Cyamopsis tetragonoloba* L.): genome structure and comparative analysis. *Genes* 2017;8(9):212.
 - [54] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30(15):2114–20.
 - [55] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990;215(3):403–10.
 - [56] Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, et al. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *GigaScience* 2015;4(1):30. Corrected and republished from *GigaScience* 2012;1(1):18.
 - [57] Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. Scaffolding preassembled contigs using SSPACE. *Bioinformatics* 2011;27(4):578–9.
 - [58] Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, et al. CpGAVAS, an integrated web server for the annotation, visualization, analysis and GenBank submission of completely sequenced chloroplast genome sequences. *BMC Genomics* 2012;13(1):715.
 - [59] Misra S, Harris N. Using Apollo to browse and edit genome annotations. *Curr Protoc Bioinformatics* 2006;12(1):9.5.1–28.
 - [60] Schattner P, Brooks AN, Lowe TM. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res* 2005;33(Suppl 2):W686–9.
 - [61] 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. *Curr Genet* 2007;52(5–6):267–74.
 - [62] Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 2013;30(12):2725–9.
 - [63] Jansen RK, Kaittanis C, Sasaki C, Lee SB, Tomkins J, Alverson AJ, et al. Phylogenetic analyses of *Vitis* (Vitaceae) based on complete chloroplast

- genome sequences: effects of taxon sampling and phylogenetic methods on resolving relationships among rosids. *BMC Evol Biol* 2006;6(1):32.
- [64] Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res* 2001;29(22):4633–42.
- [65] Benson G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 1999;27(2):573–80.
- [66] Thiel T, Michalek W, Varshney RK, Graner A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor Appl Genet* 2003;106(3):411–22.
- [67] Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. VISTA: computational tools for comparative genomics. *Nucleic Acids Res* 2004;32(Suppl 2):W273–9.
- [68] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011;28(10):2731–9.
- [69] Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, et al. A DNA barcode for land plants. *Proc Natl Acad Sci USA* 2009;106(31):12794–7.
- [70] Yang P, Shen WH, Shi JM, Chen XY, Zhang KK, Guan P. Identification of DNA barcoding of common medicinal plants in Lamiaceae. *Chin Tradit Herb Drugs* 2017;7:1397–402. Chinese.
- [71] Gao L, Wang B, Wang ZW, Zhou Y, Su YJ, Wang T. Plastome sequences of *Lygodium japonicum* and *Marsilea crenata* reveal the genome organization transformation from basal ferns to core leptosporangiates. *Genome Biol Evol* 2013;5(7):1403–7.
- [72] Fleischmann TT, Scharff LB, Alkatib S, Hasdorf S, Schöttler MA, Bock R. Nonessential plastid-encoded ribosomal proteins in tobacco: a developmental role for plastid translation and implications for reductive genome evolution. *Plant Cell* 2011;23(9):3137–55.
- [73] Ueda M, Nishikawa T, Fujimoto M, Takanashi H, Arimura S, Tsutsumi N, et al. Substitution of the gene for chloroplast RPS16 was assisted by generation of a dual targeting signal. *Mol Biol Evol* 2008;25(8):1566–75.
- [74] Hebert PD, Cywinska A, Ball SL, DeWaard JR. Biological identifications through DNA barcodes. *Proc R Soc B* 2003;270(1512):313–21.
- [75] Song Y, Chen Y, Lv J, Xu J, Zhu S, Li MF, et al. Development of chloroplast genomic resources for species discrimination. *Front Plant Sci* 2017;8:1854.
- [76] Dong W, Xu C, Li C, Sun J, Zuo Y, Shi S, et al. ycf1, the most promising plastid DNA barcode of land plants. *Sci Rep* 2015;5(1):8348.
- [77] Zhang L, Yang Z, Huang X, Li J, Wan D. Study on the leaf epidermal structural characters of *Salvia miltiorrhiza* and *Salvia* from Sichuan. *J Sichuan Univ Nat Sci* 2008;45(3):674–80. Chinese.
- [78] Wang T, Liu SY, Wang L, Wang HY, Zhang L. Anatomical characteristics of laminae and petioles of 11 species of *Salvia* and their taxonomic significance. *China J Chin Mater Med* 2014;39(14):2629–34. Chinese.
- [79] Wang Y, Jiang K, Wang L, Han D, Yin G, Wang J, et al. Identification of *Salvia* species using high-performance liquid chromatography combined with chemical pattern recognition analysis. *J Sep Sci* 2018;41(3):609–17.
- [80] Skała E, Wysokiński H. Tanshinone production in roots of micropropagated *Salvia przewalskii* maxim. *Z Naturforsch C* 2005;60(7–8):583–6.