

# 唐松草属叶绿体基因组的结构特征解析、分子标记筛选及其系统进化研究

张越<sup>1</sup>, 宋美芳<sup>1</sup>, 杨雯媛<sup>1</sup>, 徐鸟娇<sup>2</sup>, 李贤静<sup>2</sup>, 李海涛<sup>1</sup>, 王云强<sup>1</sup>, 瞿路<sup>1</sup>, 段宝忠<sup>2</sup>, 张忠廉<sup>1,2\*</sup>

1. 中国医学科学院药用植物研究所云南分所, 云南省南药可持续利用重点实验室, 云南 景洪 666100

2. 大理大学药学院, 云南 大理 671000

**摘要:** 目的 对唐松草属 *Thalictrum* L. 植物叶绿体 (chloroplast, CP) 基因组结构及序列特征进行详细解析并筛选出候选分子标记, 为解决唐松草属植物因分布广、表型变异复杂等原因导致的物种分类与鉴定争议问题奠定基础。方法 利用 Illumina HiSeq4000 测序平台首次对滇川唐松草 *Thalictrum finetii* 和高原唐松草 *Thalictrum cultratum* 进行 CP 基因组测序及结构分析; 结合 NCBI 已公布的同属 14 个物种的 CP 基因组数据, 对其简单序列重复 (simple sequence repeats, SSR)、IR 区边界结构、核苷酸多样性 (nucleotide diversity,  $P_i$ ) 等进行分析; 最后结合毛茛科 (Ranunculaceae) 10 个属 45 个物种的 CP 基因组数据, 采用最大似然法 (maximum likelihood, ML) 构建系统发育树并进行系统进化分析。结果 滇川唐松草和高原唐松草 CP 基因组的长度分别为 155 953 bp 及 155 901 bp, 且均呈典型的圆形四分体结构; 共鉴定出 131 个基因; 密码子偏好以 A/U 结尾; 唐松草属物种的 CP 基因组在基因数量和基因组结构上均具有较高的保守性, 但个别物种在 IRs/LSC 和 IRs/SSC 边界的特定基因长度上存在特异性差异; 序列变异方面, 非编码序列变异程度明显高于编码序列, 且 IRs 区明显比 LSC 和 SSC 区更为保守; 筛选出 10 个高变异位点区域 (*ndhF-rpl32*、*ycf1*、*petN-psbM*、*ndhC-trnV*、*trnT-trnL*、*trnS-psbZ*、*ndhG-ndhI*、*ndhD*、*infA*、*rpl16*), 可作为候选 DNA 条形码序列; 通过系统进化分析明确了唐松草属内物种的进化关系及该属在毛茛科中的系统进化位置。结论 首次公布了滇川唐松草和高原唐松草的 CP 基因组, 并深入解析了唐松草属的 CP 基因组结构与序列特征; 所筛选的 10 个高变异位点区域可作为唐松草属物种鉴定的候选 DNA 条形码; 建立了更全面、可靠性更好的毛茛科系统进化树。

**关键词:** 滇川唐松草; 高原唐松草; 叶绿体基因组; 唐松草属; 系统发育

中图分类号: R283 文献标志码: A 文章编号: 0253-2670(2026)06-2289-13

DOI: 10.7501/j.issn.0253-2670.2026.06.023

## Analysis of structural characteristics, screening of molecular markers and phylogenetic study of chloroplast genome in *Thalictrum*

ZHANG Yue<sup>1</sup>, SONG Meifang<sup>1</sup>, YANG Wenyuan<sup>1</sup>, XU Niaoqiao<sup>2</sup>, LI Xianjing<sup>2</sup>, LI Haitao<sup>1</sup>, WANG Yunqiang<sup>1</sup>, QU Lu<sup>1</sup>, DUAN Baozhong<sup>2</sup>, ZHANG Zhonglian<sup>1,2</sup>

1. Yunnan Key Laboratory of Southern Medicine Utilization, Yunnan Branch of Institute of Medicinal Plant Development Chinese Academy of Medical Sciences, Jinghong 666100, China

2. College of Pharmacy, Dali University, Dali 671000, China

**Abstract: Objective** A detailed analysis of the chloroplast (CP) genome structure and sequence characteristics of *Thalictrum* plants was conducted to screen for candidate molecular markers, laying the foundation for addressing the controversies in species classification and identification of *Thalictrum* species caused by their wide distribution and complex phenotypic variations. **Methods** The CP genome sequencing and structural analysis of *Thalictrum finetii* and *Thalictrum cultratum* were performed for the first time using the Illumina HiSeq4000 sequencing platform. Subsequently, a comprehensive analysis was performed simple sequence repeats (SSRs), inverted repeat (IR) region boundary structures, nucleotide diversity ( $P_i$ ) and other characteristics, combined with the chloroplast (CP) genome data of 14 congeneric species published in the NCBI. Finally, by integrating the CP genome data of 45 species from 10 genera within the Ranunculaceae family, an ML phylogenetic tree was constructed and phylogenetic analysis was performed.

收稿日期: 2025-11-03

基金项目: 云南省“兴滇英才支持计划”青年人才专项; 中国医学科学院医学与健康科技创新工程: 药用植物种质资源库建设 (2021-I2M-1-032)

作者简介: 张越, 助理研究员, 主要从事中药资源与分子生药学研究。E-mail: 861613424@qq.com

\*通信作者: 张忠廉, 研究员, 主要从事药用植物资源与分子生药学研究。E-mail: zzl0605@163.com

**Results** The chloroplast genome lengths of *T. finetii* and *T. cultratum* are 155 953 bp and 155 901 bp, respectively, both exhibiting a typical circular quadripartite structure. A total of 131 genes were identified, with codon usage bias showing a predominant preference for A/U endings. While the CP genomes of *Thalictrum* species demonstrate high conservation in both gene count and genomic structure, specific differences are observed in the lengths of particular genes at the IRs/LSC and IRs/SSC boundaries in certain species. In terms of sequence variation, the divergence in non-coding sequences was significantly higher than that in coding sequences, with the IR regions exhibiting markedly lower variability compared to the LSC and SSC regions. Ultimately, ten hypervariable regions (*ndhF-rpl32*, *ycf1*, *petN-psbM*, *ndhC-trnV*, *trnT-trnL*, *trnS-psbZ*, *ndhG-ndhI*, *ndhD*, *infA*, *rpl16*) were identified as candidate DNA barcodes for the *Thalictrum* genus. Through systematic evolutionary analysis, the evolutionary relationships among the species within the genus *Thalictrum* and the systematic evolutionary position of this genus in the Ranunculaceae family were clarified. **Conclusion** This study presents the first report of the chloroplast genomes of *T. finetii* and *T. cultratum*, with a detailed analysis of the genomic structure and sequence characteristics of the *Thalictrum* genus. Ten hypervariable regions were identified as candidate DNA barcodes for species identification within *Thalictrum*. And a more comprehensive and reliable phylogenetic tree of the family Ranunculaceae was established.

**Key words:** *Thalictrum finetii* B. Boivin; *Thalictrum cultratum* Wall.; chloroplast genome; *Thalictrum* L.; phylogenetic analysis

唐松草属 *Thalictrum* L. 植物全球分布约 150 种, 中国分布有 76 种。该属植物具有显著的药用价值, 其地下根茎和地上茎秆均可入药, 有清热、降压、抗菌等功效, 在民间医药体系中常被用于肠炎、痢疾、黄疸、目赤肿痛等疾病的治疗<sup>[1-3]</sup>。另外, 唐松草还具有抗炎、抗癌、抗病毒及清热泻火、燥湿解毒的作用<sup>[4]</sup>。例如, 林翠梧等<sup>[5]</sup>从尖叶唐松草 *Thalictrum acutifolium* (Hand.-Mazz.) B. Boivin 根部分离得到一种双苄基异喹啉类生物碱, 可诱导肺癌细胞发生凋亡, 凋亡率超过 60%, 在相同浓度下对正常细胞未显现毒性或抑制作用。刘冉等<sup>[6]</sup>从长柄唐松草 *T. przewalskii* 中分离得到化合物 2,5,7,9-四羟基-甲氧基-9,10-二氢菲具有抗肿瘤作用。薛晶晶等<sup>[7]</sup>从高原唐松草 *T. cultratum* Wall. 和贝加尔唐松草 *T. baicalense* Turcz. ex Ledeb. 2 种唐松草属植物根的乙醇提取物中分离鉴定了 11 个简单异喹啉类生物碱。当前对唐松草属植物的研究主要聚焦于化学成分分离鉴定及药理作用方面。然而, 唐松草属植物形态变异复杂, 传统形态学手段难以准确界定物种亲缘关系与分类界限, 因此在当地的草药市场上, 该属中近缘物种常因形态相似性而被混淆<sup>[8]</sup>。如滇川唐松草和高原唐松草形态相似度高, 地理分布存在交叉, 药用部位相同, 市场鉴别困难。这一现象进一步导致了中国民间药材市场的混乱局面, 其特点是不同基原药材的交叉混合。这些问题严重阻碍了对唐松草属药用植物的临床应用及科学研究。

由于唐松草属植物的分类学难题仍未解决且很多种群在分类系统中的地位还存在争议, 所以开发唐松草属植物的特异性分子标记鉴定方法, 并深

入研究唐松草类植物的系统进化关系对解决唐松草药材混淆使用及其分类学争议问题有着至关重要的意义。目前, 使用常见 DNA 条形码鉴别唐松草属物种的研究鲜少见刊, 本实验室也曾将 ITS2、*psbA-trnH*、*rbcL*、*matK* 4 个常用的 DNA 分子标记应用与唐松草属部分植物的鉴别, 均未能成功鉴别开属内各个物种。叶绿体 (chloroplast, CP) 有自己完整的遗传体系, 基因组小且构造和成分稳定、分子进化速度适中、测序方便。随着高效基因测序技术不断发展, 大量植物 CP 基因序列被破译, 为植物物种鉴定、探讨植物演化关系、分类学研究提供了新的分子证据<sup>[9]</sup>。研究表明, CP 基因组通常包含 110~130 个功能基因, 并且这些基因的编码区域保守性较强, 与其他质体基因组相比分子演化速度更慢, 结构更加稳定, 基于这些特性, CP 基因组在植物群体遗传学以及系统发育关系的研究中有明显优势<sup>[10-12]</sup>。目前, CP 全基因组作为超级条形码已广泛应用于植物系统发育关系评价或物种鉴定中, 同时从 CP 全基因组高变异区域选择的序列也已被用于物种鉴定研究<sup>[13-14]</sup>。如侯哲等<sup>[15]</sup>对星毛唐松草 *T. cirrhosum* H. Lévl. CP 基因组的碱基组成和基因类型等特征进行分析, 并与同属近缘种的 CP 基因组进行比较和系统发育分析, 分析结果显示 CP 基因组可为唐松属植物的分类及鉴定提供可靠依据。随着 NCBI 数据库已收录的植物 CP 基因序列不断增加, 本研究相较于侯哲等<sup>[15]</sup>的研究增加了更多物种的 CP 基因组数据, 建立了更为完整的系统进化树, 同属植物基因组比较得到的候选高变异区可靠性也有所提高。更多的 CP 基因组数据对解决唐松草

属植物种内和种间的微小差异,有效利用药用植物资源有着重要作用。

本研究利用 Illumina HiSeq4000 测序平台,对滇川唐松草 *T. finetii* B. Boivin 和高原唐松草的 CP 全基因组进行了测序。并研究了它们的基本特征,包括 CP 结构分析、简单重复序列 (simple repeat sequences, SSRs) 和密码子使用情况等。接下来,结合 NCBI 已发布的 14 个唐松草属物种的 CP 基因组进行比较,分析了核苷酸多样性 (nucleotide diversity,  $P_i$ ), 并鉴定了高变区,可作为开发唐松草属 DNA 条形码的潜在标记。最后,本课题组收集了毛茛科 (Ranunculaceae) 45 种植物的 CP 基因组序列,并将其作为超级条形码对该组植物进行了鉴定及系统发育分析。该研究为唐松草属植物的物种鉴定和系统发育分析提供了重要的遗传信息,同时也有助于缓解药材市场上唐松草属药用植物的准确鉴定问题。

## 1 材料

本研究中用于 CP 基因组测序的样品滇川唐松草采自云南省大理州剑川县 (100°01'14", 26°34'01", 3 166 m), 高原唐松草采自西藏昌都市八宿县 (97°20'14", 30°10'07", 4 051 m), 由中国医学科学院药用植物研究所云南分所的张忠廉研究员鉴定为滇川唐松草 *T. finetii* B. Boivin 和高原唐松草 *T. cultratum* Wall., 腊叶标本保存于中国医学科学院药用植物研究所云南分所标本馆。于 2023 年 8 月底在野外采集当年生成健康叶片,收集的叶片用 75%乙醇清洗干净,立即用硅胶干燥, -80 °C 低温保存用于后续实验。

## 2 方法

### 2.1 DNA 提取和测序

在低温条件下,将液氮冷冻的叶片在预冷的研磨器中快速研磨成细粉。将这些植物粉末转移到含有裂解缓冲液的离心管中,在室温条件下静置,确保细胞完全裂解,充分释放 DNA。利用天根生化科技有限公司的植物基因组提取试剂盒进行 DNA 的提取。通过 1%的琼脂糖电泳检查 DNA 完整性好, Nanodrop 2000 分光光度计测定 DNA 的浓度与纯度 ( $A_{260}/A_{280}$  分别为 1.92、2.01), 达到后续建库要求。

利用 Covaris M220 聚焦超声仪 (Covaris, Woburn, MA, 美国) 将 DNA 分解成 300~500 bp 的片段,筛选 500 bp 大小的片段建立文库。DNA 文库使用 Illumina TruSeq™ 纳米 DNA 样品准备试剂盒 (Illumina, San Diego, CA, 美国) 构建。文库

富集通过聚合酶链反应 (polymerase chain reaction, PCR) 扩增进行,目标条带从 2%琼脂糖凝胶中回收。文库使用 Illumina HiSeq 4000 测序平台进行测序,获得 2×150 bp 的末端配对 reads。使用 FastQC Toolkit 检查原始 reads ( $Q \geq 25$ )<sup>[16]</sup>。从原始数据中过滤掉低质量的 reads,去除含有 10% N 的 reads,并在高质量修剪后丢弃 <75 bp 的小片段,以获得高质量的数据 (清洁的 reads) 用于后续分析。

### 2.2 CP 全基因组的组装与注释

用 FileZilla 3.51.0 软件将上述数据上传到服务器,使用 GetOrganelle 软件对 CP 基因组进行组装<sup>[17]</sup>。过滤后的 'gfa' 文件在 Bandage v.0.8.1 中可视化<sup>[18]</sup>。接下来,使用 Geneious v.8.0.2 中的 Bowtie 2 将原始序列与组装好的 CP 基因组比对,验证组装结果<sup>[19]</sup>。最后,利用参考基因组西南唐松草 *T. fargesii* Franch. ex Finet & Gagnep. 对 CP 组装序列的起始位置进行校正,确定 CP 4 个区域 (large single-copy, LSC; small single-copy, SSC; inverted repeat (IRS) regions, IRa/IRb) 的位置和方向,得到组装后的 CP 基因组序列。将组装结果导入 Geneious v.8.0.2 中进行注释,然后在 Geneious v.8.0.2 中手动调整蛋白质编码基因的起始密码子、终止密码子和内含子的位置<sup>[19]</sup>。使用 tRNAscan SE 对 tRNA 基因进行在线验证<sup>[20]</sup>。利用在线网站 (<https://chlorobox.mpimp-golm.mpg.de/OGDraw.html>) 绘制 2 种唐松草的 CP 全基因组图谱<sup>[21]</sup>。最后,获得了 sqn 文件,并向 NCBI 提交了序列。滇川唐松草和高原唐松草的 CP 全基因组序列分别存入 Genbank,序列号分别为 PX556660 和 PX556661。

### 2.3 CP 基因组特征分析

使用 CodonW 软件获取相对同义密码子使用情况 (relative synonymous codon usage, RSCU) 及密码子分布情况,计算 RSCU 值,并可视化<sup>[22]</sup>。使用 MEGA X 分析鸟嘌呤-胞嘧啶 (guanine-cytosine, GC) 含量<sup>[23]</sup>。利用 MicroSarellite 鉴定工具 MISA Perl Script (<http://pgrc.ipk-gatersleben.de/misa/>) 对组装好的唐松草属 CP 基因组序列进行简单序列重复 (Simple Sequence Repeats, SSR) 位点分析<sup>[24]</sup>,参数设置参考同类研究标准: 1-10、2-5、3-4、4-3、5-3 和 6-3<sup>[15,25]</sup>。IRScope 软件则用于检测 IR/SC 区域边界的收缩和扩展<sup>[26]</sup>。

### 2.4 CP 基因组比较及 $P_i$ 分析

结合 NCBI 已发布的唐松草属 14 个物种 (表 1)

CP 基因组使用在线 MAFFT 软件进行比对<sup>[27]</sup>, 鉴定 CP 基因组之间的保守序列。采用 Shuffle-LAGAN 模式下的 mVISTA<sup>[28]</sup>程序, 以丝叶唐松草 *T. foeniculaceum* Bunge CP 基因组为参照, 对唐松草属 16 个物种的 CP 基因组进行比较。然后, 我们使用 DnaSP<sup>[29]</sup>软件确定  $P_i$ , 步长为 200 bp, 窗口长度为 600 bp, 结果通过 Excel 进行可视化。

表 1 NCBI 已发布的唐松草属 14 个物种信息

Table 1 Information of 14 species in genus *Thalictrum* released by NCBI

中文名	拉丁名	Genebank号
欧洲唐松草	<i>T. aquilegifolium</i>	NC_058830.1
贝加尔唐松草	<i>Thalictrum baicalense</i>	MZ962406.1
星毛唐松草	<i>Thalictrum cirrhosum</i>	NC_061927.1
朝鲜唐松草	<i>Thalictrum coreanum</i>	NC_026103.1
小叶唐松草	<i>Thalictrum elegans</i>	NC_085600.1
西南唐松草	<i>Thalictrum fargesii</i>	NC_070058.1
丝叶唐松草	<i>Thalictrum foeniculaceum</i>	NC_053570.1
多叶唐松草	<i>Thalictrum foliolosum</i>	NC_058920.1
亚欧唐松草	<i>Thalictrum minus</i>	NC_041544.1
瓣蕊唐松草	<i>Thalictrum petaloideum</i>	MK253449.1
箭头唐松草	<i>Thalictrum simplex</i>	NC_068627.1
细唐松草	<i>Thalictrum tenue</i>	MK253448.1
唐松莲花	<i>Thalictrum thalictroides</i>	NC_039433.1
黏唐松草	<i>Thalictrum viscosum</i>	NC_058831.1

### 2.5 唐松草属物种系统发育分析

为了确定唐松草属植物的系统进化关系及其在毛茛科植物中的系统发育位置, 利用毛茛科 10 个属 45 个完整 CP 基因组序列进行分析。MAFFT 对这些序列进行比对<sup>[25]</sup>。并使用喜马拉雅大黄 *Rheum webbianum* Royle (PQ963207) 及滇边大黄 *Rheum delavayi* Franch. (NC\_063092) 的 CP 基因组作为外类群, 参考相关研究后选择 MEGA X 软件<sup>[23]</sup>采用最大似然法 (maximum likelihood, ML) 构建了 47 个 CP 基因组序列的系统发育树, bootstrap 值设置为 1 000 次重复。

### 3 结果与分析

#### 3.1 滇川唐松草和高原唐松草 CP 基因组的序列特征分析

本研究测序分析了 2 种唐松草属植物的 CP 基因组, 发现滇川唐松草和高原唐松草长度分别为 155 953、155 901 bp, GC 含量 38.37%、38.39%。这些基因组均展现出典型的四分体结构 (图 1), 包括 2 个 IRs, 1 个 LSC 和 1 个 SSC。滇川唐松草 CP 基因组 4 部分 (LSC、SSC、IRa 和 IRb) 的长度分别为 85 395、17 588、26 485 和 26 485 bp, 高原唐松草 CP 基因组 4 部分 (LSC、SSC、IRa 和 IRb) 的长度分别为 85 343、17 588、26 485 和 26 485 bp。在滇川唐松草和高原唐松草 CP 基因组中, 共鉴定出 131 个基因, 涵盖 86 个蛋白编码基因 (protein coding

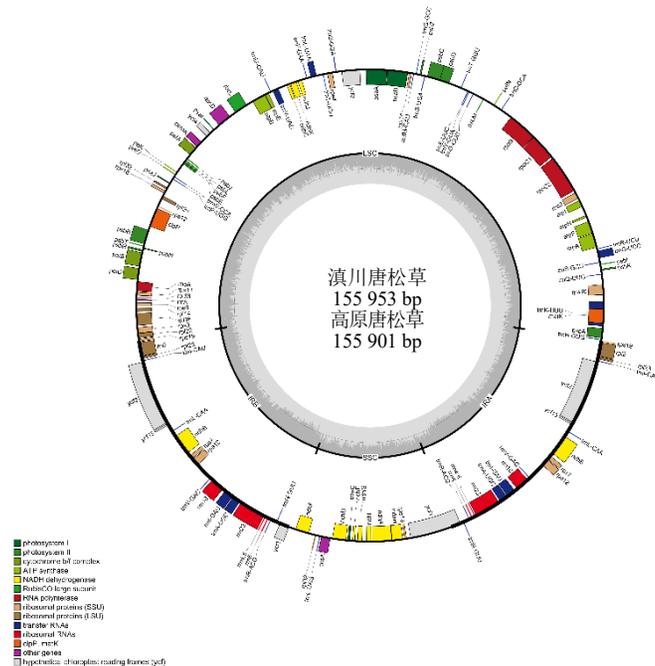


图 1 滇川唐松草和高原唐松草 CP 基因组结构图

Fig. 1 Structure of chloroplast genome of *T. finetii* and *T. cultratatum*

genes, PCGs)、37 个转运 RNA 基因 (transfer RNA, tRNA)和 8 个核糖体 RNA 基因(ribosome RNA, rRNA)。

本实验在分析 2 种唐松草 CP 基因组时, 发现其内含子基因的分布、名称及数量具有显著特征(表 2)。对唐松草属 2 种 CP 基因组的比较分析显示, 其内含子分布呈现规律性特征: 18 个含内含子基因中包含 12 个蛋白编码基因和 6 个 tRNA 基因。值得注意的是, 这些基因中仅 *ycf 3*、*rps12* 和 *clpP* 3 个

基因具有双内含子结构, 其余 15 个基因均为单内含子类型。18 个含内含子基因中长度最长的基因为 *ndhA* 基因 (1 072~1 088 bp), 含内含子长度最短的基因为 *trnL-UAA* 基因 (425~437 bp)。LSC 区集中分布了 9 个基因 (*rps16*、*rpoC1* 等), 占全部内含子基因的 50%; IR 区呈现对称分布特征, 两侧各含 4 个相同基因 (*rpl2*、*ndhB* 等); 而 SSC 区仅检测到 *ndhA* 1 个含内含子基因。

表 2 滇川唐松草和高原唐松草 CP 基因组的基因组成

Table 2 Genetic composition of chloroplast genomes of *T. finetii* and *T. cultratum*

基因类型	基因名	基因数量
光系统I	<i>psaA</i> 、 <i>psaB</i> 、 <i>psaC</i> 、 <i>psaI</i> 、 <i>psaJ</i>	5
光系统II	<i>psbA</i> 、 <i>psbB</i> 、 <i>psbC</i> 、 <i>psbD</i> 、 <i>psbE</i> 、 <i>psbF</i> 、 <i>psbH</i> 、 <i>psbI</i> 、 <i>psbJ</i> 、 <i>psbK</i> 、 <i>psbL</i> 、 <i>psbM</i> 、 <i>psbN</i> 、 <i>psbT</i> 、 <i>psbZ</i>	15
细胞色素b/f复合体	<i>petA</i> 、 <i>petB*</i> 、 <i>petD*</i> 、 <i>petG</i> 、 <i>petL</i> 、 <i>petN</i>	6
ATP 合成酶	<i>atpA</i> 、 <i>atpB</i> 、 <i>atpE</i> 、 <i>atpF*</i> 、 <i>atpH</i> 、 <i>atpI</i>	6
NADH 脱氢酶	<i>ndhA*</i> 、 <i>ndhB*(×2)</i> 、 <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>	12
核糖体大亚基	<i>rbcL</i>	1
RNA 聚合酶 RNA	<i>rpoA</i> 、 <i>rpoB</i> 、 <i>rpoC1*</i> 、 <i>rpoC2</i>	4
核糖体小亚基蛋白质	<i>rps2</i> 、 <i>rps3</i> 、 <i>rps4</i> 、 <i>rps7(×2)</i> 、 <i>rps8</i> 、 <i>rps11</i> 、 <i>rps12***(×2)</i> 、 <i>rps14</i> 、 <i>rps15</i> 、 <i>rps16*</i> 、 <i>rps18</i> 、 <i>rps19</i>	14
核糖体大亚基蛋白质	<i>rpl2*(×2)</i> 、 <i>rpl14</i> 、 <i>rpl16*</i> 、 <i>rpl20</i> 、 <i>rpl22</i> 、 <i>rpl23(×2)</i> 、 <i>rpl33</i> 、 <i>rpl36</i>	10
其他基因	<i>accD</i> 、 <i>clpP**</i> 、 <i>matK</i> 、 <i>ccsA</i> 、 <i>cemA</i> 、 <i>infA</i>	6
未知功能基因	<i>ycf1</i> 、 <i>ycf2(×2)</i> 、 <i>ycf3**</i> 、 <i>ycf4</i> 、 <i>ycf15(×2)</i>	7
转运RNA	37 tRNAs (4 contain an intron, 2 contain two intron 8 in the IRs)	37
核糖体RNA	<i>rrn4.5(×2)</i> 、 <i>rrn5(×2)</i> 、 <i>rrn16(×2)</i> 、 <i>rrn23(×2)</i>	8
合计		131

\*含有 1 个内含子的基因, \*\*含有 2 个内含子的基因; (×2) 含有 2 个拷贝的基因。

\*Gene containing one intron; \*\*gene containing two introns; (×2) gene with two copies.

### 3.2 密码子使用情况分析

在编码序列里, 同义密码子和非同义密码子的选用存在不平衡现象, RSCU 指标用来量化这种不平衡程度, 当 RSCU 低于 1.00 这个阈值时, 密码子的选用频次就达不到常规预期, 若 RSCU 超过 1.00 的界限, 密码子的选用频次就超出了常规预期。滇川唐松草和高原唐松草的 CP 基因组的密码子使用水平见图 2。本研究分析了 2 个物种蛋白质编码基因的相对密码子使用频率, 2 种唐松草蛋白质编码基因的密码子使用频次以及 RSCU 比值经分析后可知, 二者基因组均有 26 183 个密码子, 其中亮氨酸、异亮氨酸和甘氨酸编码序列在这 2 种唐松草 CP 基因组里最常见。滇川唐松草和高原唐松草 CP DNA 序列里密码子以腺嘌呤和尿嘧啶 (A/U) 结尾的占比接近, 分别为 68.59%和 68.57%, 这类密码子相对同义密码子使用频率 (RSCU) 指标往往较高, 如编码亮氨酸的 UUA (RSCU=1.83)、编码精氨酸

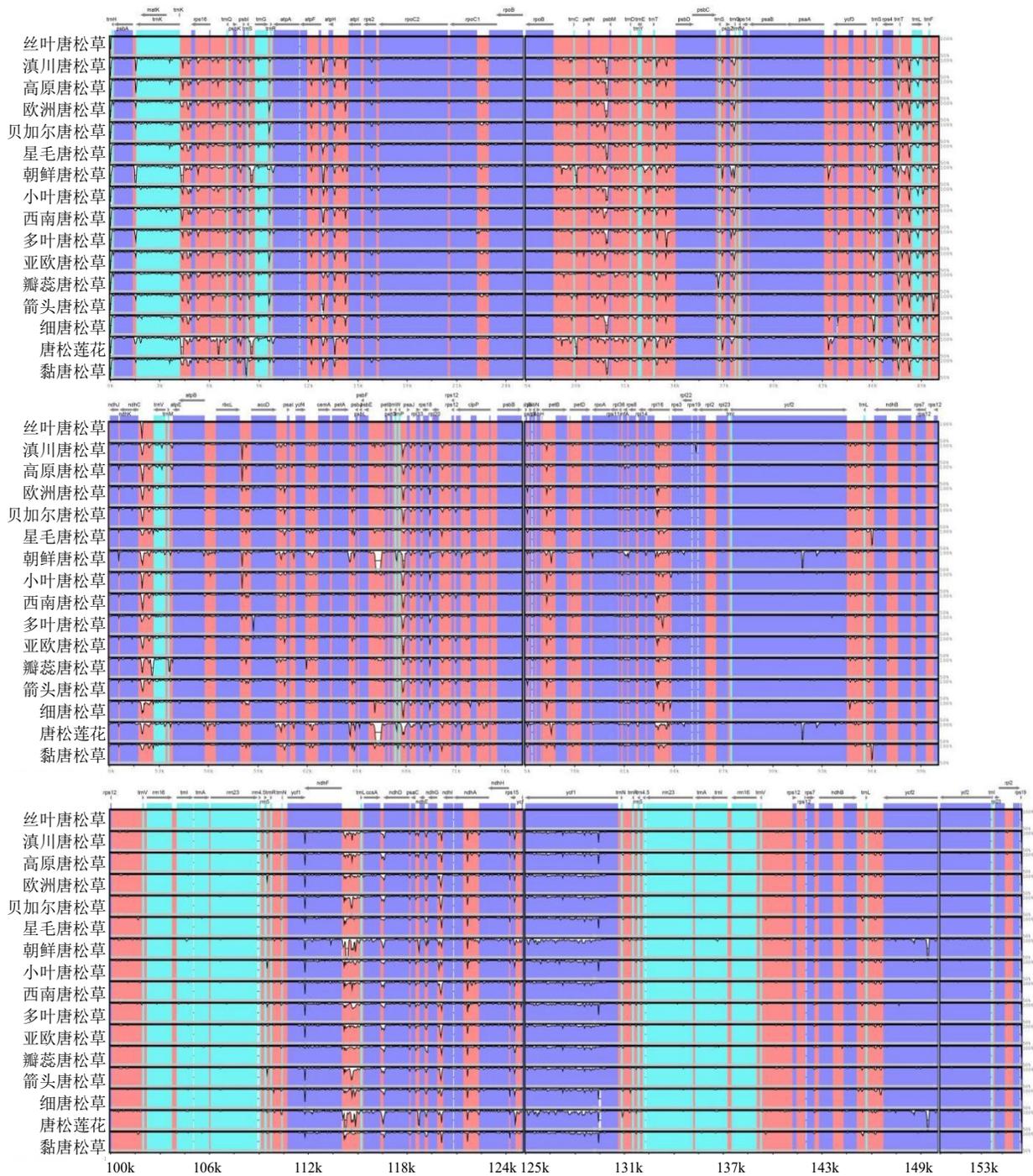
的 AGA (RSCU=1.83)、编码甘氨酸的 GCU (RSCU=1.67)。此外, 分析 A/T 偏差比率能推测密码子应用规律, 一般高等植物的 CP 基因组序列中常用密码子频率大多较高且常偏好 A/T 选择, 同样, 本研究中 2 种唐松草属植物 (除 Leu-UUG 和 Ser-UCC 外) 所有偏好同义密码子 (RSCU 比值超 1.00) 都以 A 或者 U 结尾。

### 3.3 SSR 分析

结合 NCBI 已发表的唐松草属的 14 个物种, 结合本研究的滇川唐松草和高原唐松草, 分析了 16 种唐松草属 CP 基因组中 SSRs 的分布和类型。共检测到 1 311 个 SSRs。其中, 朝鲜唐松草 *T. coreanum* 和细唐松草 *T. tenue* 的 SSRs 数最多, 分别为 91 和 90 个, 而瓣蕊唐松草 *T. petaloideum* 的 SSRs 数最少, 仅为 69 个。16 个物种的 SSRs 主要分布在 CP 基因组的 LSC 区, 其次是 SSC 区和 IRs 区。单核苷酸重复序列最多 (50.55%~63.89%), 其次是二核苷







排列上方的灰色箭头和粗黑线表示基因方向；白色峰表示不同 CP 基因组之间的差异；外显子、内含子和保守的非编码序列（CNSs）以不同的颜色显示；图的相似度临界值为 70%，y 轴表示相似度百分比（50%~100%）。

Gray arrows and thick black lines above the alignments indicate gene orientations; White peaks represent differences among CP genomes; Exons, introns, and conserved non-coding sequences (CNSs) were displayed in different color; A similarity cut-off value of 70% was used for the plots, and the Y-axis represents the percentage similarity (50%—100%).

图 5 16 种唐松草的 CP 全基因组进行比对分析

Fig. 5 Alignment and comparative analysis of whole CP genome for 16 *Thalictrum* species

区域的序列差异。课题组发现，序列差异在非编码区更为突出，而外显子和非翻译区（UTRs）的差异则更为微小。变异程度最高的非编码区有 *ndhF-rpl32-trnL-*

*UAG*、*petN-psbM*、*ndhC-trnV-UAC*、*trnT-UGU-trnL-UAA*，变异程度最高的编码区有 *yef1*、*ndhD*、*rpl16*。非编码区 *ndhF-rpl32-trnL-UAG* 在唐松草属内 16 个

物种间存在较多变异位点,且朝鲜唐松草与唐松莲花在该区间存在较大片段的插入与缺失; *petN-psbM* 与 *trnT-UGU-trnL-UAA* 也存在较多变异位点,且朝鲜唐松草存在小片段的插入; *ndhC-trnV-UAC* 区间则是存在较多变异位点。编码区 *ycf1* 区间同样是存在大量变异位点,同时朝鲜唐松草与唐松莲花有小片段的插入现象。这些高变异区域可作为开发唐松草属物种的特异性分子标记的潜在位点。

### 3.6 $P_i$ 分析

利用 DnaSP 软件检测了唐松草属物种 *CP* 基因组序列的  $P_i$  和高度可变区域(图6)。结果表明,IRs 区的变异性低于 LSC 和 SSC 区。LSC 和 SSC 区域中

相对较高的  $P_i$  值表明,在基因组进化过程中可能发生快速的核苷酸替代,这在物种鉴定和系统发育分析中发挥了重要作用。16 种唐松草 *CP* 基因组的平均  $P_i$  值为 0.0056。共检测到 5 个  $P_i > 0.02$  的核苷酸多态性热点区域,其中 LSC 区有 3 个: *petN-psbM*、*ndhC-trnV*、*trnT-trnL*, SSC 区有 2 个: *ndhF-rpl32*、*ycf1*。基于完整 *CP* 基因组序列的比对,筛选出 10 个高变异位点: *ndhF-rpl32*、*ycf1*、*petN-psbM*、*ndhC-trnV*、*trnT-trnL*、*trnS-psbZ*、*ndhG-ndhI*、*ndhD*、*infA*、*rpl16*, 这些高变异区域可以作为唐松草属物种鉴定的候选标记,可用引物设计工具 Primer BLAST 对其高变异区域设计特异性标记引物。

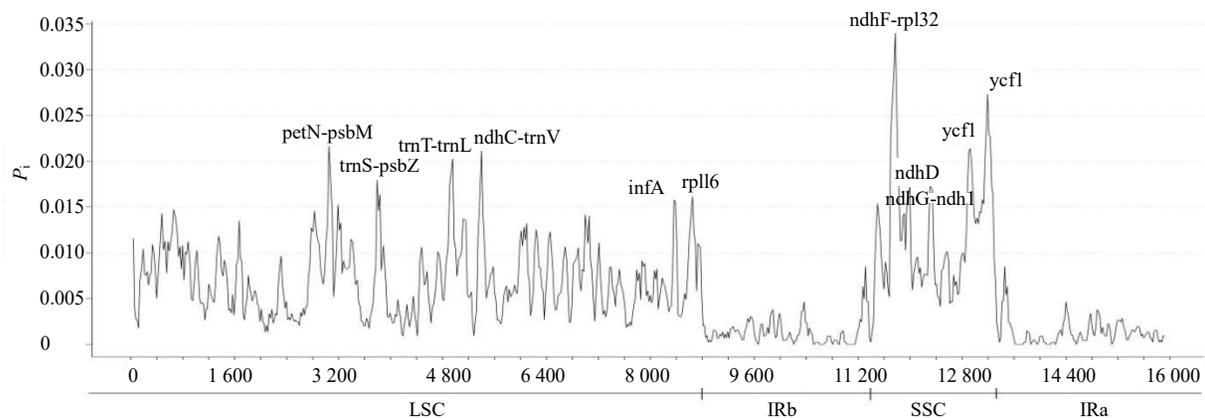


图6 基于16种唐松草属植物完整CP基因组的滑动窗口分析

Fig. 6 Sliding window analysis of *CP* genomes of 16 *Thalictrum* species

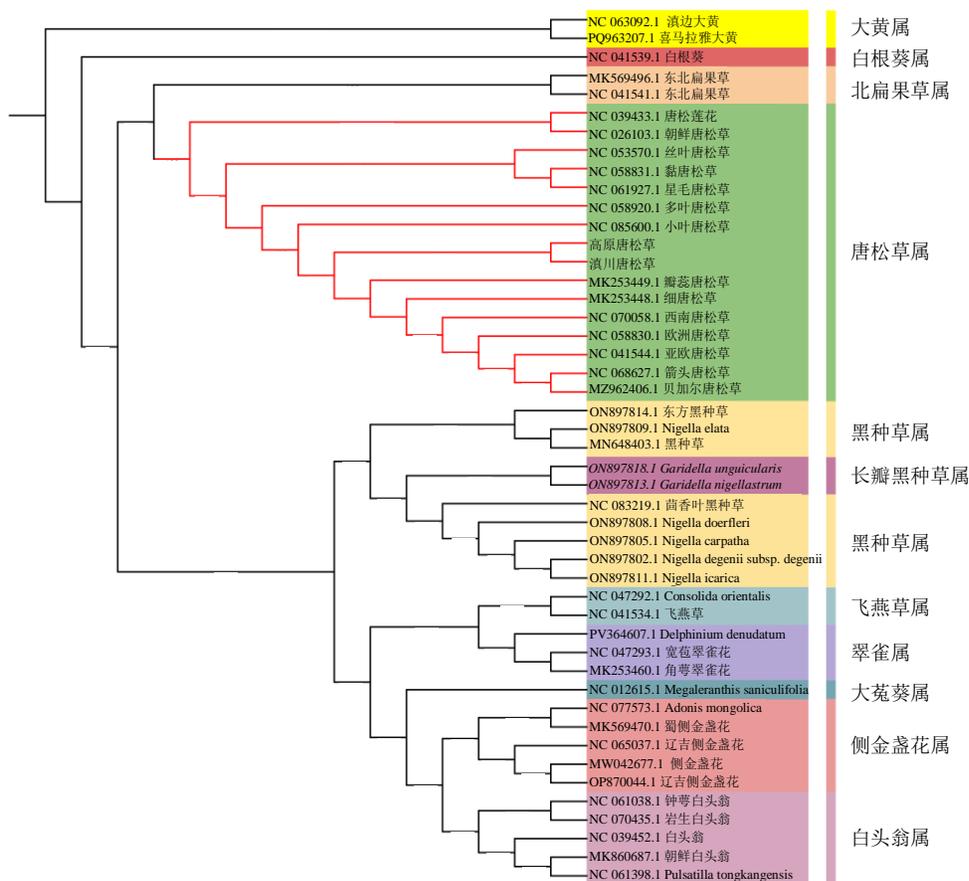
### 3.7 毛茛科植物的系统发育关系

*CP* 基因组序列对研究高等植物的系统发育关系、遗传结构和分类研究具有重要意义。为了研究唐松草属的系统发育关系及其在毛茛科中的系统位置,利用毛茛科 10 个属的 45 个物种 *CP* 基因组序列构建了 ML 系统发育树(图7),所有节点都有较高支持率,分析结果可靠性好。系统发育树显示,喜马拉雅大黄和滇边大黄作为外类群最早分裂成一个独立的分支。随后,白根葵属 *Glaucidium* Siebold & Zucc. 的白根葵分裂为一个独立分支,其余 9 个属聚在了另一个大的分支。其中,飞燕草属 *Consolida* (DC.) S.F. Gray、翠雀属 *Delphinium* L.、黑种草属 *Nigella* L.、长瓣黑种草属 *Garidella* L. 聚为一支,大菟葵属 *Megaleranthis* Ohwi、侧金盏花属 *Adonis* L.、白头翁属 *Pulsatilla* Adans.、北扁果草属 *Isopyrum* L.、唐松草属聚在另一支。黑种草属是一个非单系类群,形成了 2 个进化分支,其中一个分支与长瓣黑种草属的 2 个物种聚在了一起。此外,

北扁果草属与唐松草属聚在了一个分枝,说明唐松草属与北扁果草属的亲缘关系较近,但目前 NCBI 已公布的唐松草属物种 *CP* 基因组数据有限,无法进一步对 2 个属物种的发育关系进行探究,未来可以考虑增加更多样品数据并整合基因组数据来探索 2 个属间的演化关系及分类地位,以弥补目前数据有限和 *CP* 母系遗传特性的不足。系统发育分析强化了对这些物种种属间进化关系的理解,为物种鉴定和系统进化研究提供了重要的分子依据。

### 4 讨论

*CP* 基因组在大部分被子植物中的遗传方式为母系遗传,具有基因组较小、易于测序、结构稳定和组成较为保守、进化速率较慢等优点,因而植物 *CP* 基因组常被用于群体进化、种质资源鉴定及群体遗传学等方面研究,在物种的进化过程中发挥着重要作用<sup>[30-31]</sup>。基于以上背景,本研究对滇川唐松草和高原唐松草的 *CP* 全基因组进行了测序,并分析了它们的基本特征,发现滇川唐松草和高原唐松



支持率显示在系统发育树分支上。

Bootstrap support values are shown on the branches of the phylogenetic tree.

图 7 基于毛茛科 CP 全基因组的 ML 系统发育树

Fig. 7 ML phylogenetic tree based on complete CP of Ranunculaceae

草与其他高等植物一样均展现出典型的四分体结构<sup>[32-33]</sup>。

密码子使用偏好是在 CP 基因组进化中起着重要作用，受自然选择、基因突变、系统发育关系等诸多因素的影响<sup>[34-36]</sup>。此外，密码子使用在基因表达方面发挥着不可或缺的作用，功能基因的转录、翻译及蛋白表达方面均受到密码子影响，因而密码子的使用偏好也会对基因的表达产生重要作用<sup>[37]</sup>。密码子的使用偏好与基因表达密切相关，并影响到基因组中的蛋白质和 mRNA 水平，密码子的使用方法及偏好性会在植物长期的进化进程中逐渐累积，因而在不同植物，甚至同一植物不同基因间，密码子的偏好性也会有所区别。因此，研究 CP 基因组中密码子使用偏好情况对更好地理解功能基因的表达模式及进化方式大有裨益，也能进一步了解不同物种间的进化关系<sup>[37-38]</sup>。滇川唐松草和高原唐松草的 CP 基因组中以腺嘌呤和尿嘧啶 (A/U) 结尾者占

比较高，为 68.59%和 68.57%。这些密码子使用结果与先前报道的虎耳草属 *Saxifraga* Tourn. ex L. 物种、壶瓶碎米荠 *Cardamine hupingshanensis* K. M. Liu, L. B. Chen, H. F. Bai & L. H. Liu、红豆蔻 *Alpinia galanga* (L.) Willd. 和长柄山姜 *A. kwangsiensis* T. L. Wu & S. J. Chen 的结果相似<sup>[34, 39-40]</sup>。以 A/U 结尾的密码子高偏好性可能与植物对生境的适应性选择相关，此外，也可能归因于对应的氨基酸功能以及避免转录时发生错误所需的肽结构相关。这一现象表明，稳定的 CP 基因组进化有助于保护 CP 基因免受有害突变的影响，同时提高对选择压力的适应性<sup>[41-42]</sup>。

简单序列重复是由 1~6 个核苷酸重复单元组成的序列，广泛分布于 CP 基因组中，常用于群体遗传学和分子系统发育研究<sup>[43-44]</sup>。本研究结合 NCBI 已发表的唐松草属 CP 基因组数据，分析了共计 16 个 CP 基因组中检测到的 SSRs 数量及不同

类型核苷酸的分布情况。发现 A 和 T 是最常用的碱基，而含有 C 和 G 碱基的重复序列类型则不常见。这一结果与先前报道叶绿体 SSRs 通常由短 poly-A 或 poly-T 重复序列组成的研究结果一致<sup>[45-46]</sup>。目前，SSR 标记已被广泛应用于遗传多样性和群体结构评估、比较基因组学、遗传图谱开发、标记辅助选择育种等领域<sup>[47-48]</sup>。本研究中唐松草属植物的 SSRs 重复序列信息能为该属植物在物种鉴定和遗传多样性及群体结构研究方面提供宝贵的数据资源。

许多研究者认为，高等植物的 CP 基因组高度保守，CP 基因组大小变化的主要原因是 IRs 区域和 LSC/SSC 区域边界的收缩和扩张<sup>[49-50]</sup>。本研究比较分析了唐松草属 16 个物种的 IRs/LSC 和 IRs/SSC 边界结构。发现 ycf1 的假基因化和出现的位置也在其他高等植物中较为常见<sup>[51-53]</sup>，且 16 个唐松草物种的 CP 基因组虽然在基因数量和基因组结构上具有较高的保守性，但在 IRs、LSC 和 SSC 区域的特定基因长度上存在显著差异。这一现象表明 IR 可能存在扩张和收缩，IR 区的动态变化（包括扩展/收缩）及其与 LSC 和 SSC 区边界的位移也被认为是导致 CP 基因组长度变异的关键因素，是植物 CP 基因组变异的驱动力之一<sup>[50, 54-55]</sup>。

通过对植物 CP 基因组序列的比较分析，可以很容易地识别出高变异区域，这些变异对阐明该属植物的遗传结构和进化关系有着重要作用<sup>[56-57]</sup>。利用在线工具 mVISTA 对 16 种唐松草属植物 CP 基因组序列进行了比较，分析结果显示，2 个 IRs 区域的序列差异小于 LSC 和 SSC 区域的序列差异，且序列差异在非编码区更为突出。这一结果进一步支持了许多学者的观点，即高等植物的编码区比非编码区更保守，IR 区比 LSC 和 SSC 区更保守。Khakhlova 和 Bock 认为，这种现象可能是由于基因转换纠正了 IR 序列中的有害突变<sup>[58]</sup>。同时利用 DnaSP 软件检测 CP 基因组序列的  $P_i$ ，并筛选到 10 个高变异位点区，可作为唐松草属物种鉴定的候选标记。后续研究中可收集更多唐松草属物种样品（重复 $\geq 3$ ），以验证这些候选标记的鉴别能力，最终选择出鉴别能力最优的标记，建立唐松草属物种条形码鉴定体系，在野外样本鉴定及药材市场质量控制中均能发挥重要作用。先前已有研究人员对人参属 *Panax* L.、龙胆属 *Gentiana* (Tourn.) L. 等进行了分子鉴定研究，结果表明 CP 遗传标记具有较高的

物种鉴定能力<sup>[43, 59]</sup>。

基于毛茛科 10 个属 45 种植物及 2 个外类群 CP 基因组的系统发育树显示：唐松草属所有物种聚为 1 个分支，而北扁果草属与唐松草属聚在了一个大分枝，说明唐松草属与北扁果草属的亲缘关系较近。唐松草属的系统发育和分类一直存在争议，该属植物形态特性较为复杂多变，用传统形态学手段来确定唐松草类群间亲缘关系、划分物种界限较为困难<sup>[8, 60]</sup>。CP 基因组序列已被证明可用于被子植物的系统发育研究，尤其是针对亲缘关系较近的属内或种内不同品种间的植物分类鉴定很有价值。本研究中丝叶唐松草、黏唐松草 *T. viscosum* W. T. Wang & S. H. Wang 与星毛唐松草，唐松莲花与朝鲜唐松草，滇川唐松草与高原唐松草亲缘关系最近，均分别聚为一个小支，此结果与 Chen 等<sup>[25]</sup>对唐松草属物种鉴定研究中的结果一致。但本研究中贝加尔唐松草与箭头唐松草、亚欧唐松草亲缘关系最近，Chen 等<sup>[25]</sup>的研究中贝加尔唐松草与西南唐松草、细唐松草、瓣蕊唐松草有较近的亲缘关系，这一不同的原因可能是本研究分析了唐松草属更多物种的 CP 基因组，理论上，所分析数据越全面，得到的系统发育分析结果也更可靠<sup>[25]</sup>。今后对唐松草属系统发育关系及物种间的分类界定研究需要根据其形态类型从野生种群中选择更多的样本，并从核基因组和 CP 基因组中收集更广泛的序列数据，以找出更全面、更可靠的系统发育和物种进化证据。

**利益冲突** 所有作者均声明不存在利益冲突

#### 参考文献

- [1] 王文采, 孙英宝. 《中国唐松草属植物》[J]. 广西植物, 2018, 38(10): 1258.
- [2] 蔡仕宁, 郑佳琪, 李文斯, 等. 滇产高原唐松草的非生物碱类成分研究 [J]. 化学研究与应用, 2025, 37(1): 219-223.
- [3] Jiang S Q, Zhang Y B, Xiao M, et al. Cycloartane triterpenoid saponins from the herbs of *Thalictrum fortunei* [J]. *Carbohydr Res*, 2017, 445: 1-6.
- [4] 刘军. 唐松草药理作用及其机制的研究进展 [J]. 中国民族民间医药, 2010, 19(3): 27-29.
- [5] 林翠梧, 苏镜娱, 曾陇梅, 等. 新的醚链双生物碱尖叶唐松草阿原碱的结构测定 [J]. 高等学校化学学报, 2000, 21(12): 1820-1823.
- [6] 刘冉, 王继红, 夏西超, 等. 长柄唐松草化学成分及其体外抗肿瘤活性 [J]. 中药材, 2021, 44(3): 610-614.
- [7] 薛晶晶, 李建勇, 李冰洁, 等. 2 种唐松草属植物异喹

- 啉类生物碱的研究 [J]. 中国中药杂志, 2022, 47(10): 2676-2680.
- [8] 王文采. 中国唐松草属植物 [M]. 北京: 北京大学出版社, 2018: 56.
- [9] 刘文雅, 张志武, 何彦峰. 贝加尔唐松草叶绿体基因组微卫星序列及密码子偏好性分析 [J]. 分子植物育种, 2023, 42(2): 1-23.
- [10] Jiang Y H, Li H, Wu M, *et al.* Comparative chloroplast genomes of *Incarvillea* species (Bignoniaceae) unveiled genomic diversity and shed light on phylogenetic relationships [J]. *BMC Plant Biol*, 2025, 25(1): 399.
- [11] 毛立彦, 黄秋伟, 龙凌云, 等. 7 种睡莲属植物叶绿体基因组密码子偏好性分析 [J]. 西北林学院学报, 2022, 37(2): 98-107.
- [12] 朱斌, 甘晨晨, 王洪程. 球花石斛 *Dendrobium thysiflorum* 叶绿体基因组特征及亲缘关系解析 [J]. 生物技术通报, 2021, 37(5): 38-47.
- [13] Chen X L, Zhou J G, Cui Y X, *et al.* Identification of *Ligularia* herbs using the complete chloroplast genome as a super-barcode [J]. *Front Pharmacol*, 2018, 9: 695.
- [14] Yang H Y, Wang L Q, Chen H M, *et al.* Phylogenetic analysis and development of molecular markers for five medicinal *Alpinia* species based on complete plastome sequences [J]. *BMC Plant Biol*, 2021, 21(1): 431.
- [15] 侯哲, 娄晓鸣, 李昂, 等. 星毛唐松草叶绿体基因组特征及系统发育分析 [J]. 植物资源与环境学报, 2023, 32(3): 24-32.
- [16] Brown J, Pirrung M, McCue L A. FQC Dashboard: Integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool [J]. *Bioinformatics*, 2017, 33(19): 3137-3139.
- [17] Jin J J, Yu W B, Yang J B, *et al.* GetOrganelle: A fast and versatile toolkit for accurate *de novo* assembly of organelle genomes [J]. *Genome Biol*, 2020, 21(1): 241.
- [18] Wick R R, Schultz M B, Zobel J, *et al.* Bandage: Interactive visualization of *de novo* genome assemblies [J]. *Bioinformatics*, 2015, 31(20): 3350-3352.
- [19] Kearse M, Moir R, Wilson A, *et al.* Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data [J]. *Bioinformatics*, 2012, 28(12): 1647-1649.
- [20] Lowe T M, Chan P P. tRNAscan-SE On-line: Integrating search and context for analysis of transfer RNA genes [J]. *Nucleic Acids Res*, 2016, 44(W1): W54-W57.
- [21] Lohse M, Drechsel O, Kahlau S, *et al.* OrganellarGenomeDRAW: A suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets [J]. *Nucleic Acids Res*, 2013, 41(Web Server issue): W575-W581.
- [22] Sharp P M, Li W H. The codon Adaptation Index: A measure of directional synonymous codon usage bias, and its potential applications [J]. *Nucleic Acids Res*, 1987, 15(3): 1281-1295.
- [23] Kumar S, Stecher G, Li M, *et al.* MEGA X: Molecular evolutionary genetics analysis across computing platforms [J]. *Mol Biol Evol*, 2018, 35(6): 1547-1549.
- [24] Beier S, Thiel T, Münch T, *et al.* MISA-web: A web server for microsatellite prediction [J]. *Bioinformatics*, 2017, 33(16): 2583-2585.
- [25] Chen S X, Safiul Azam F M, Akter M L, *et al.* The first complete chloroplast genome of *Thalictrum fargesii*: Insights into phylogeny and species identification [J]. *Front Plant Sci*, 2024, 15: 1356912.
- [26] Amirouf A, Hyvönen J, Poczai P. IRscope: An online program to visualize the junction sites of chloroplast genomes [J]. *Bioinformatics*, 2018, 34(17): 3030-3031.
- [27] Katoh K, Rozewicki J, Yamada K D. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization [J]. *Brief Bioinform*, 2019, 20(4): 1160-1166.
- [28] Frazer K A, Pachter L, Poliakov A, *et al.* VISTA: Computational tools for comparative genomics [J]. *Nucleic Acids Res*, 2004, 32(Web Server issue): W273-W279.
- [29] Rozas J, Ferrer-Mata A, Sánchez-DelBarrio J C, *et al.* DnaSP 6: DNA sequence polymorphism analysis of large data sets [J]. *Mol Biol Evol*, 2017, 34(12): 3299-3302.
- [30] 侯哲, 娄晓鸣, 李昂, 等. 星毛唐松草叶绿体基因组的测序及其密码子偏好性分析 [J]. 西南林业大学学报: 自然科学, 2023, 43(5): 47-56.
- [31] Shaw J, Lickey E B, Beck J T, *et al.* The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis [J]. *Am J Bot*, 2005, 92(1): 142-166.
- [32] Ahmad W, Asaf S, Khan A, *et al.* Complete chloroplast genome sequencing and comparative analysis of threatened dragon trees *Dracaena serrulata* and *Dracaena cinnabari* [J]. *Sci Rep*, 2022, 12(1): 16787.
- [33] Gu L, Su T, Luo G L, *et al.* The complete chloroplast genome sequence of *Heteropolygonatum Ginfushanicum* (Asparagaceae) and phylogenetic analysis [J]. *Mitochondrial DNA Part B*, 2021, 6(7): 1799-1802.
- [34] Chen Z, Yu X L, Yang Y J, *et al.* Comparative analysis of chloroplast genomes within *Saxifraga* (Saxifragaceae) takes insights into their genomic evolution and adaptation to the high-elevation environment [J]. *Genes*, 2022, 13(9): 1673.
- [35] Xu C, Cai X N, Chen Q Z, *et al.* Factors affecting synonymous codon usage bias in chloroplast genome of *Oncidium Gower Ramsey* [J]. *Evol Bioinform Online*, 2011, 7: 271-278.
- [36] Das S, Paul S, Dutta C. Synonymous codon usage in adenoviruses: Influence of mutation, selection and protein hydrophathy [J]. *Virus Res*, 2006, 117(2): 227-236.

- [37] 李真, 李永光, 薄蕾, 等. 大豆和拟南芥 NFYB 基因家族同义密码子使用偏好性的比较 [J]. 作物杂志, 2014(3): 14-21.
- [38] Lyu X L, Liu Y. Nonoptimal codon usage is critical for protein structure and function of the master general amino acid control regulator CPC-1 [J]. *mBio*, 2020, 11(5): e02605-e02620.
- [39] Zhang Y, Song M F, Li Y, *et al.* Complete chloroplast genome analysis of two important medicinal *Alpinia* species: *Alpinia galanga* and *Alpinia Wangsiensis* [J]. *Front Plant Sci*, 2021, 12: 705892.
- [40] Huang S N, Kang Z J, Chen Z F, *et al.* Comparative analysis of the chloroplast genome of *Cardamine hupingshanensis* and phylogenetic study of *Cardamine* [J]. *Genes*, 2022, 13(11): 2116.
- [41] Gao B M, Yuan L, Tang T L, *et al.* The complete chloroplast genome sequence of *Alpinia oxyphylla* Miq. and comparison analysis within the Zingiberaceae family [J]. *PLoS One*, 2019, 14(6): e0218817.
- [42] Ivanova Z, Sablok G, Daskalova E, *et al.* Chloroplast genome analysis of resurrection tertiary relict *Haberlea rhodopensis* highlights genes important for desiccation stress response [J]. *Front Plant Sci*, 2017, 8: 204.
- [43] George B, Bhatt B S, Awasthi M, *et al.* Comparative analysis of microsatellites in chloroplast genomes of lower and higher plants [J]. *Curr Genet*, 2015, 61(4): 665-677.
- [44] 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(7): 1962.
- [45] Ebert D, Peakall R. Chloroplast simple sequence repeats (cpSSRs): Technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species [J]. *Mol Ecol Resour*, 2009, 9(3): 673-690.
- [46] Kuang D Y, Wu H, Wang Y L, *et al.* Complete chloroplast genome sequence of *Magnolia kwangsiensis* (Magnoliaceae): Implication for DNA barcoding and population genetics [J]. *Genome*, 2011, 54(8): 663-673.
- [47] Flannery M L, Mitchell F J G, Coyne S, *et al.* Plastid genome characterisation in *Brassica* and Brassicaceae using a new set of nine SSRs [J]. *Theor Appl Genet*, 2006, 113(7): 1221-1231.
- [48] Cui Y X, Nie L P, Sun W, *et al.* Comparative and phylogenetic analyses of ginger (*Zingiber officinale*) in the family Zingiberaceae based on the complete chloroplast genome [J]. *Plants*, 2019, 8(8): 283.
- [49] Wang Y F, Wen F, Hong X, *et al.* Comparative chloroplast genome analyses of *Paraboea* (Gesneriaceae): Insights into adaptive evolution and phylogenetic analysis [J]. *Front Plant Sci*, 2022, 13: 1019831.
- [50] Zhang Y, Song M F, Tang D Y, *et al.* Comprehensive comparative analysis and development of molecular markers for *Lasianthus* species based on complete chloroplast genome sequences [J]. *BMC Plant Biol*, 2024, 24(1): 867.
- [51] Lu Q X, Chang X, Gao J, *et al.* Evolutionary comparison of the complete chloroplast genomes in *Convallaria* species and phylogenetic study of Asparagaceae [J]. *Genes*, 2022, 13(10): 1724.
- [52] Szczecińska M, Sawicki J. Genomic resources of three *Pulsatilla* species reveal evolutionary hotspots, species-specific sites and variable plastid structure in the family Ranunculaceae [J]. *Int J Mol Sci*, 2015, 16(9): 22258-22279.
- [53] Zhu J Q, Huang Y, Chai W G, *et al.* Decoding the chloroplast genome of *Tetrastigma* (Vitaceae): Variations and phylogenetic selection insights [J]. *Int J Mol Sci*, 2024, 25(15): 8290.
- [54] Kim K J, Lee H L. Complete chloroplast genome sequences from Korean ginseng (*Panax schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants [J]. *DNA Res*, 2004, 11(4): 247-261.
- [55] Pei J L, Wang Y, Zhuo J, *et al.* Complete chloroplast genome features of dendrocalamusfarinosus and its comparison and evolutionary analysis with other Bambusoideae species [J]. *Genes*, 2022, 13(9): 1519.
- [56] Daniell H, Lin C S, Yu M, *et al.* Chloroplast genomes: Diversity, evolution, and applications in genetic engineering [J]. *Genome Biol*, 2016, 17(1): 134.
- [57] Hong Z, Wu Z Q, Zhao K K, *et al.* Comparative analyses of five complete chloroplast genomes from the genus *Pterocarpus* (fabaceae) [J]. *Int J Mol Sci*, 2020, 21(11): 3758.
- [58] Khakhlova O, Bock R. Elimination of deleterious mutations in plastid genomes by gene conversion [J]. *Plant J*, 2006, 46(1): 85-94.
- [59] Nguyen V B, Park H S, Lee S C, *et al.* Authentication markers for five major *Panax* species developed via comparative analysis of complete chloroplast genome sequences [J]. *J Agric Food Chem*, 2017, 65(30): 6298-6306.
- [60] Xiang K L, Mao W, Peng H W, *et al.* Organization, phylogenetic marker exploitation, and gene evolution in the plastome of *Thalictrum* (Ranunculaceae) [J]. *Front Plant Sci*, 2022, 13: 897843.

[责任编辑 时圣明]