

大苞半蒴苣苔(*Hemiboea magnibracteata* Y. G. Wei & H. Q. Wen) 叶绿体基因组特征和系统发育分析

张梅¹, 汤升虎¹, 杨旭东², 童琪¹, 龙婷^{1*}

(1 贵州省植物园, 贵州 贵阳 550004; 2 清华大学地球系统科学系, 北京 100084)

摘要: 阐明大苞半蒴苣苔的叶绿体基因组结构特征及其在半蒴苣苔属中的系统发育关系对于半蒴苣苔属的分类学研究具有重要意义。为此, 选取大苞半蒴苣苔及其 23 个近缘种为研究对象, 以小石蝴蝶和中华石蝴蝶 2 个种为外类群。通过对叶组织的 DNA 提取、组装、注释和分析, 并与 9 个半蒴苣苔属植物的叶绿体基因组进行比对, 最终选取 2 个叶绿体基因片段(*rps16*, *trnL_F*), 应用贝叶斯法进行了系统发育分析。结果表明: 大苞半蒴苣苔的叶绿体基因组结构呈四分体环状, 包括 1 个大单拷贝区, 1 对反向重复区和 1 个小单拷贝区。叶绿体基因组共编码了 132 个功能基因(即: 87 个蛋白基因、37 个 tRNA 基因、8 个 rRNA 基因)。LSC 区的基因变异度较高, 可作为鉴定种内和种间的 DNA 条形码片段; IRa 区和 IRb 区基因可用于科与科之间和属与属之间的关系鉴定。基于叶绿体基因片段(*rps16*, *trnL_F*) 建立的系统发育关系和形态性状综合分析表明, 大苞半蒴苣苔与拟大苞半蒴苣苔亲缘关系较近, 为姐妹类群, 支持率为 0.95, 其系统发育树可信度较高。本研究揭示出叶绿体基因片段(*rps16*, *trnL_F*) 可作为建立半蒴苣苔属系统发育关系的 DNA 条形码, 明确了大苞半蒴苣苔在半蒴苣苔属植物中的系统位置, 为后续分子遗传标记、大苞半蒴苣苔遗传育种和叶绿体基因工程研究提供了理论参考。

关键词: 半蒴苣苔属; 大苞半蒴苣苔; 叶绿体基因组; DNA 条形码; 系统发育分析

中图分类号: Q949.4

文献标志码: A

The characterization and phylogenetic analysis of chloroplast genome in *Hemiboea magnibracteata* Y. G. Wei & H. Q. Wen

ZHANG Mei¹, TANG Shenghu¹, YANG Xudong², TONG Qi¹, LONG Ting^{1*}

(1 Guizhou Botanical Garden, Guiyang, Guizhou 550004, China; 2 Department of Earth System Science, Tsinghua University, Beijing 100084, China)

Abstract: The structural characteristics of the chloroplast genome of *Hemiboea magnibracteata* are of great significance for reconstructing phylogenetic relationships and advancing taxonomic research within *Hemiboea*. In this study, *H. magnibracteata* and 23 closely related species were selected as research objects, with *Petrocosmea minor* and the *P. sinensis* designated as outgroups. First, the chloroplast genome of this species was extracted, assembled, annotated, and analyzed, and then compared with those of nine *Hemiboea* species; Subsequently, candidate chloroplast gene fragments were screened based on their levels of genetic variation; Finally, two chloroplast genes fragments (*rps16*, *trnL_F*) were selected and subjected to Bayesian phylogenetic analysis using the combined dataset of *H. magnibracteata*, 23 closely related species, and two outgroups. The results showed that the chloroplast genome of *H. magnibracteata* has a quadripartite circular structure, comprising a Large single copy region (LSC), a pair of inverted repeat regions (IRa, IRb) and a small single-copy region (SSC). The chloroplast genome of *H. magnibracteata* encodes 87 protein genes, 37 tRNA genes, 8 rRNA genes. The LSC region, which exhibits high genetic variation, can serve as a DNA barcode for distinguishing intraspecific and interspecific relationships. The IRa and IRb regions are more suitable for resolving phylogenetic relationships at higher taxonomic levels, such as between genera or families. Comprehensive analysis of chloroplast genome sequence and morphological traits revealed that *H. magnibracteata* and *H. pseudomagnibracteata* are closely related and form a well-supported sister clade. This study demonstrates that the chloroplast gene fragment *rps16* and *trnL_F* can serve as effective DNA barcodes for reconstructing the phylogeny of *Hemiboea*. The phy-

收稿日期: 2025-07-21

基金项目: 贵州省科技计划项目(黔科合基础-ZK[2022]一般 291)

作者简介: 张梅(1989—), 女, 硕士, 副研究员, 主要从事植物分类学的研究, e-mail: zhangmei909@126.com。

* 通信作者: 龙婷(1994—), 女, 博士, 助理研究员, 主要从事植物多样性保护方向的研究, e-mail: longting2020@bjfu.edu.cn。

logenetic position of *H. magnibracteata* within *Hemiboea* has been clarified. Overall, this study provides a theoretical foundation for future research on molecular genetic markers, genetic breeding of *H. magnibracteata*, and chloroplast genetic engineering.

Key words: *Hemiboea*; *Hemiboea magnibracteata*; chloroplast genome; DNA barcode; phylogenetic analysis

半蒴苣苔属 (*Hemiboea*) 隶属于苦苣苔科 (Gesneriaceae) 长蒴苣苔族 (Trib. Didymocarpeae)。作为半蒴苣苔属的主要类群,大苞半蒴苣苔 (*Hemiboea magnibracteata*) 是韦毅刚和文和群于 1995 年发现并发表于《广西植物》^[1],为多年生草本植物,茎笔直不分枝,具棱无毛;叶对生;花序为聚伞花序,花期为每年 8 月;果为蒴果,果期每年 9—10 月^[1-2]。该种是一种新兴的喀斯特地区园艺植物,具有观赏和药用价值^[3]。然而,由于大苞半蒴苣苔是草本植物,在分类学上存在基因渐渗及种间杂交等现象,同物异名问题突出。大多数分类学者采用 DNA 条形码来辅助解决这一分类学问题。其中,筛选出能够作为 DNA 条形码的基因片段是关键所在。因此,有必要对大苞半蒴苣苔的叶绿体基因组特征及其在半蒴苣苔属中的系统发育关系进行全面剖析,从而为半蒴苣苔属的分类学研究和大苞半蒴苣苔的育种工作提供必要的科学依据。

叶绿体基因组的表达为单亲遗传,不同性别的后代仅遗传母方的基因,且叶绿体基因可随着无性繁殖产生不对称分离现象。因此,在系统发育关系构建和亲缘物种的种间鉴定中,叶绿体基因组有着广泛应用^[4-6]。例如,在有丝分裂时,其中的一个子细胞可能接受一种类型基因的质体,而其它带有 1 个或者 2 个突变位点、具有不同遗传信息的质体基因组会被随机地分配进入另一个子细胞中。这种营养细胞分裂造成的质体不均匀分配会导致同一组织具有不同的宏观性状表现。叶绿体基因与线粒体基因和核基因之间的进化具有差异性,种属进化速率的特异性是叶绿体基因进化的最大特征^[4]。如,分布于不同时间、不同区域上的物种叶绿体基因组大小不同,主要区别来源于基因间区有长片段插入与缺失的现象^[5-6]。因此,叶绿体基因组中的等位基因常被作为 DNA 条形码用于种间与种内物种的鉴定^[4-7]。

前人针对适用于半蒴苣苔属的 DNA 条形码进行了探索。李佳慧等^[7]利用核基因序列比对结果推测了半蒴苣苔属植物的 DNA 条形码,但未涉及叶绿体基因,且未进行验证。崔艳芳^[8]基于整个质体基因组中的部分基因序列(去掉 IR 区基因组序列、蛋白编码基因和非编码区的基因)构建的半蒴苣苔

属的系统进化树,其拓扑结构与前人研究结果基本一致,但仍未明确半蒴苣苔属的 DNA 条形码序列。上述研究都存在系统进化树分支支持率不高、属内物种关系模糊等问题。在探讨半蒴苣苔属内各亚种间的关系时,这些问题尤其突出。更为重要的是,该属物种基因组和物种分子鉴定研究的欠缺大大限制了半蒴苣苔属、苦苣苔科层面的相关系统发育关系研究。

目前,对苦苣苔科植物叶绿体基因组的研究较多,但主要集中在一些热门的属种,而对半蒴苣苔属的叶绿体基因组关注较少。其中,大苞半蒴苣苔的叶绿体基因组至今未曾有报道。因此,半蒴苣苔属植物的系统发育关系仍然存在相当大的不确定性。为此,本研究利用高通量测序技术对大苞半蒴苣苔进行了叶绿体全基因组测序、组装和分析,拟探讨以下科学问题:1)大苞半蒴苣苔叶绿体基因组结构特征;2)大苞半蒴苣苔及其近缘物种在 IR 和 SC 边界区域碱基分布的基因差异;3)大苞半蒴苣苔在苦苣苔科半蒴苣苔属植物系统发育中的位置。本研究可为后续开展大苞半蒴苣苔物种鉴定、遗传多样性分析、叶绿体基因工程,以及分子育种等研究提供重要的理论基础,也为半蒴苣苔属植物及苦苣苔科植物的分类学研究工作提供丰富的基因组数据。

1 材料与方法

1.1 材料的采集和 DNA 的提取与测序

样品的采集:大苞半蒴苣苔叶片样品采集于贵州省荔波县,采样时间为 2024 年 8 月,为大苞半蒴苣苔盛花期。选取枝叶分布匀称、疏密有致、单株叶片数量多的植株为采样对象,所取叶片样品均呈鲜绿色或苍翠色,色泽均匀且具光泽,叶绿素含量较高,质地柔韧而富有活力。将单个叶片置于透气且独立包装的样品袋中,并用硅胶对其干燥。植物活体保育于贵州省植物园苦苣苔保育中心,植物标本存放于贵州省植物园标本馆(采集号:He001)。

DNA 的提取与测序:本研究运用一种改进后的 CTAB 法从已用硅胶干燥过的植物叶片组织中提取总 DNA^[8-9]。将检测合格的高质量 DNA 采用 Illumina HiSeq 2500 测序平台进行双端测序^[10-14],以构建一个插入片段为 350 bp 的文库,读长为 100 bp 或

150 bp。测序工作依托测序公司(菲沙基因公司)完成。

1.2 叶绿体的组装和注释

获取原始测序数据后,使用 NGSQC Toolkit v2.3.3 软件^[13]按照系统默认参数进行粗略过滤处理,然后去掉头部和尾部未知碱基和低质量序列,最终每个物种获得至少 6 G 左右的测序数据(Clean reads),用于进行后续组装。本研究使用现有的参照方法进行组装^[5,8],选用已发表的半蒴苣苔属物种的完整叶绿体基因组作为参考(OP820508 *Hemiboea ovalifolia*; OR195434 *H. pterocaulis*; OP820512 *H. fangii*; OQ799918 *H. parvibracteata*; OP820511 *H. integra*; OQ799916 *H. subacaulis*; OQ799920 *H. suiyangensis*; OP820509 *H. sinovietnamica*; OP820510 *H. yongfuensis*)。本研究利用 NOVOPlasty^[13]和 GetOrganelle v1.7.5^[13]软件过滤出叶绿体基因组序列的原始 reads,并将原始 reads 组装成 contigs;再用 Bandage 软件将所有 contigs 连接,手动删除多余或者不正确的 contigs,使序列连接成环;最后,将 reads 重新匹配到基因组上进行检查和校对^[13,15]。

使用软件 Geneious v9.0.2 对已准确组装好的叶绿体基因组进行自动注释^[15],即选用已发表的半蒴苣苔属完整叶绿体基因组作为参考序列(OP820508; OR195434; OP820512; OQ799918; OP820511; OQ799916; OQ799920; OP820509; OP820510),将相似度参数设置为 70%。对于初步注释后依然未能成功注释出的基因,使用 BLAST 找出目标片段后进行序列比对,手工添加注释信息。同时,检查已注释的基因,以确保所注释蛋白编码基因(Rotein-coding regions, CDS)的完整性。在注释过程中没有出现完整的开放阅读框(Open Reading Frame, ORF)、在序列 1/2 区域之前提前出现终止子和缺失至少一半以上序列^[13,16-17]均注释为假基因。

1.3 叶绿体基因组特征的分析

将组装和注释好的大苞半蒴苣苔叶绿体基因组数据在 OGDRAW 中绘制叶绿体四分体环状图(图 1)^[16];叶绿体基因组中 3 个区的基因长度(大单拷贝区、小单拷贝区和反向重复区)、GC 含量和基因数量以及基因类型等统计分析采用 Geneious v9.0.2 软件^[15](表 1);大苞半蒴苣苔和 9 个近缘种植物叶绿

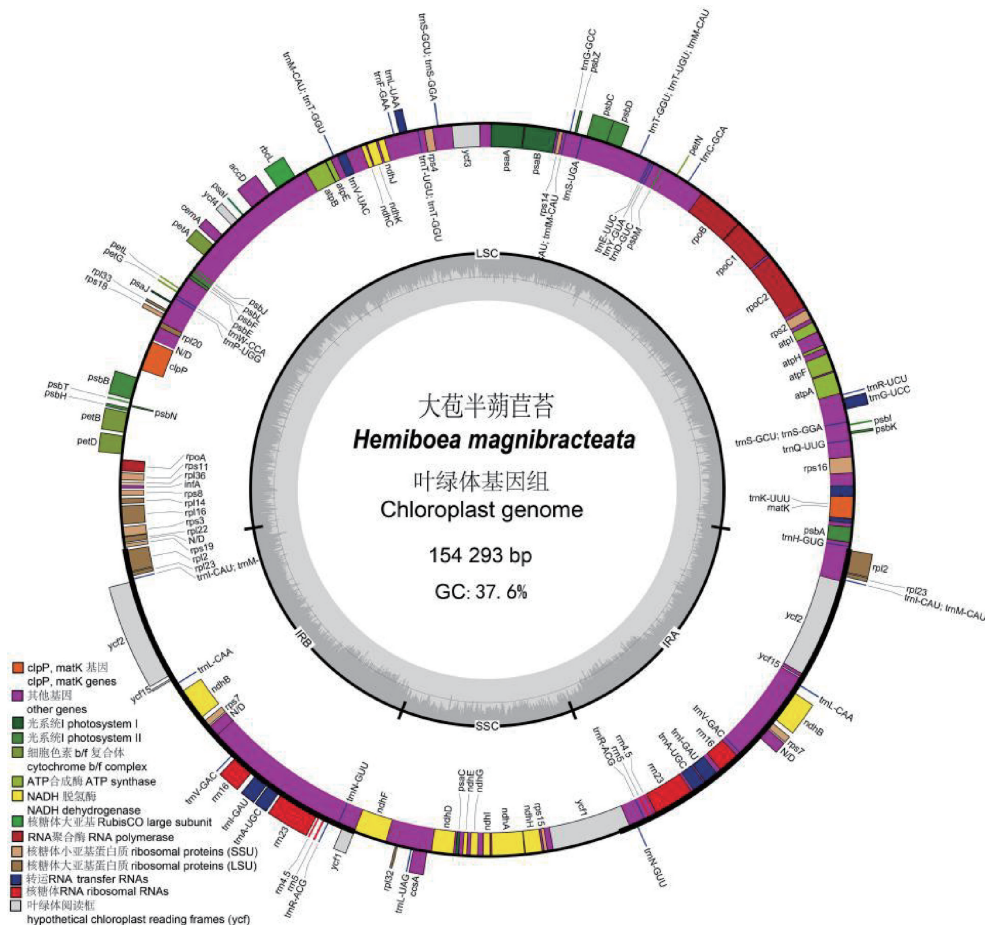


图 1 大苞半蒴苣苔植物叶绿体基因组图谱

Figure 1 Map of the chloroplast genomes of *Hemiboea magnibracteata* assembled in current study

体基因组 IR/SC 边界的扩张与收缩在 IRSCOPE 软件中完成;大苞半蒴苣苔和近缘种植物叶绿体基因组共线性分析采用 MISA (<https://genome.lbl.gov/vista/mvista/submit.shtml>) 软件进行^[14-17]。叶绿体基因组

核苷酸多态性 (Nucleotide polymorphism) 分析在 NADnaSP 6.0 软件中完成^[4]。文中所涉及的专业名词均参考钱俊^[14]、林张翔^[15]、文艺^[17]、胡佳玉^[13]、木志强^[11]、Yong 等^[5]、Panzade 等^[18]的研究。

表 1 大苞半蒴苣苔叶绿体基因组的注释列表

Table 1 List of genes annotated in the chloroplast genome of *Hemiboea magnibracteata* sequenced

基因分类	基因分组	基因名称
自我复制	rRNA	<i>rrn16, rrn23, rrn4.5, rrn5, rrn5, rrn4.5, rrn23, rrn16</i>
	tRNA	<i>trnM-CAU, trnN-GUU, trnL-CAA, trnM-CAU, trnA-UGC, trnV-GAC, trnI-GAU, trnA-UGC, trnR-ACG, trnN-GUU, trnL-UAG, trnS-GCU, trnR-ACG, trnT-GGU, trnI-GAU, trnM-CAU, trnS-GGA, trnI-CAU, trnH-GUG, trnK-UUU, trnQ-UUG, trnS-GCU, trnG-UCC, trnR-UCU, trnC-GCA, trnD-GUC, trnY-GUA, trnE-UUC, trnT-GGU, trnS-UGA, trnG-GCC, trnM-CAU, trnS-GGA, trnT-UGU, trnL-UAA, trnF-GAA, trnF-GGA, trnV-UAC, trnM-CAU, trnW-CCA, trnP-UGG, trnT-UGU, trnL-CAA, trnV-GAC, trnI-CAU, trnM-CAU</i>
	RNA 聚合酶亚基	<i>rpoC1, rpoC2, rpoB, rpoA</i>
	核糖体小亚基	<i>rps7, rps15, rps16, rps2, rps14, rps4, rps18, rps11, rps8, rps3, rps7, rps19</i>
光合作用相关	核糖体大亚基	<i>rpl2, rpl23, rpl32, rpl33, rpl20, rpl36, rpl16, rpl22, rpl14</i>
	翻译起始因子	<i>infA</i>
	ATP 合酶	<i>atpA, atpF, atpH, atpI, atpE, atpB</i>
	光和系统 I	<i>psaB, psaA, psaI, psaJ, psaC</i>
	光和系统 II	<i>psbA, psbK, psbI, psbM, psbD, psbC, psbZ, psbL, psbF, psbE, psbB, psbJ, psbN, psbH, psbT</i>
	二磷酸核酮糖羧化酶亚基	<i>rbcL</i>
	细胞色素复合物	<i>petN, petA, petL, petG, petB, petD,</i>
	NADH 脱氢酶	<i>ndhB, ndhF, ndhD, ndhE, ndhG, ndhI, ndhA, ndhH, ndhB, ndhJ, ndhK, ndhC</i>
	成熟酶	<i>matK</i>
	包膜蛋白	<i>cemA</i>
其他基因	c 型细胞色素合成基因	<i>ccsA</i>
	乙酰辅酶 A 羧化酶亚基	<i>accD</i>
	ATP 依赖的蛋白酶亚基 p	<i>clpP</i>
未知功能基因		<i>yef2, yef15, yef1, yef1, yef2, rpl23, rpl2, yef3, yef4, yef15</i>

1.4 系统发育分析

根据叶绿体基因组分析,筛选出 2 个基因片段 (*rps16, trnL-F*) 用于建立系统发育树。从 GenBank 数据库下载的序列 (表 2) 用 Contig Express 软件进行拼接与检查^[12];随后,利用 Clustal X v. 1.83 软件对已经拼接和检查好的序列排列,再用 BioEdit 中自动比对程序进行比对;最后,手工校对,删除矩阵两端无法确定的区域,用贝叶斯模型法 (Bayesian Inference, BI) 在 MrBayes 3b4 软件中构建所有物种的系统发育树^[12]。

2 结果与分析

2.1 大苞半蒴苣苔叶绿体基因特征

大苞半蒴苣苔的叶绿体基因组结构呈四分体环

状,长度为 154 293 bp,包括 1 个大单拷贝区 (LSC, 85 086 bp), 1 个小单拷贝区 (SSC, 18 097 bp) 和 1 对反向重复区 (IRa, 25 555 bp; IRb, 25 555 bp), 总 GC 含量为 37.6% (图 1)。大苞半蒴苣苔叶绿体基因组编码 132 个功能基因,即 87 个蛋白基因、37 个 tRNA 基因和 8 个 rRNA 基因 (表 1)。

在 LSC 区域和 SSC 区域的基因主要包括与光系统 I *psa* 类基因和光系统 II *psb* 类基因、*ndh* 类质体醌氧化还原酶基因、*atp* 类 ATP 合成酶基因、*rbcL* 类编码 Rubisco 大亚基基因、*rpo* 类 RNA 聚合酶基因和 *trn* 类转运 RNA 基因等。所有的 *rrn* 类核糖体 RNA 都位于反向重复区域,包括 4.5S、5S、16S 和 23S。此外 *trn* 类基因、*rps* 类基因和 *yef* 类基因等也在反向重复区域中被发现 (图 1, 表 1)。

表 2 建立系统发育树的数据
Table 2 Establish phylogenetic tree data

种名	rpS16 序列 GenBank 登录号	trnL_F 序列 GenBank 登录号
华南半蒴苣苔 <i>Hemiboea follicularis</i>	MT184611.1	HQ632885.1
台湾半蒴苣苔 <i>H. bicornuta</i>	MT184736.1	FJ501534.1
贵州半蒴苣苔 <i>H. cavaleriei</i>	MT184737.1	FJ501533.1
齿叶半蒴苣苔 <i>H. fangii</i>	MN367337.1	HQ632882.1
毛果半蒴苣苔 <i>H. flaccida</i>	MN367338.1	JF697579.1
纤细半蒴苣苔 <i>H. gracilis</i>	MT184615.1	FJ501536.1
翅茎半蒴苣苔 <i>H. pterocaulis</i>	MT184623.1	KY607416.1
大苞半蒴苣苔 <i>H. magnibracteata</i>	OK322464.1	HQ632887.1
单座苣苔 <i>H. ovalifolia</i>	MT184618.1	HQ632883.1
峨眉半蒴苣苔 <i>H. omeiensis</i>	MT184617.1	HQ632886.1
红苞半蒴苣苔 <i>H. rubibracteata</i>	OK322468.1	HQ632890.1
麻栗坡半蒴苣苔 <i>H. malipoensis</i>	OK322465.1	KJ948111.1
拟大苞半蒴苣苔 <i>H. pseudomagnibracteata</i>	MT184621.1	-
弄岗半蒴苣苔 <i>H. longgangensis</i>	MN367343.1	HQ632889.1
王氏半蒴苣苔 <i>H. wangiana</i>	MT184626.1	KM232651.1
龙州半蒴苣苔 <i>H. lungzhouensis</i>	MN367344.1	HQ632888.1
柔毛半蒴苣苔 <i>H. mollifolia</i>	MN367346.1	-
降龙草 <i>H. subcapitata</i>	OK322469.1	OK322551.1
小苞半蒴苣苔 <i>H. parvibracteata</i>	OQ799918	OQ799918
全叶半蒴苣苔 <i>H. integra</i>	OP820511	OP820511
短茎半蒴苣苔 <i>H. subacaulis</i>	OQ799916	OQ799916
绥阳半蒴苣苔 <i>H. suiyangensis</i>	OQ799920	OQ799920
中越半蒴苣苔 <i>H. sinovietnamica</i>	OP820509	OP820509
永福半蒴苣苔 <i>H. yongfuensis</i>	OP820510	OP820510
小石蝴蝶 <i>Petrocosmea minor</i>	OL689808.1	OL689837.1
中华石蝴蝶 <i>P. sinensis</i>	OL689809.1	OL689838.1

注:“-”表示 GenBank 数据库缺乏数据。

根据叶绿体各大基因功能分类,可分为 4 类。第一类是与光合作用相关的光合系统基因 45 种,如与光系统 I 相关的 *psa*(*psaA*、*psaB* 等)基因,与光系统 II 相关的 *psb*(*psbA*、*psbB* 等)基因,细胞色素 *f/b6* *pet* 基因,ATP 合成酶基因 *atp*,编码 Rubisco 的 *rbc* 基因,NADH 质体醌氧化还原酶基因 *ndh*,碳吸收相关的 *cem* 基因。第二类是与翻译、转录相关的遗传系统基因 80 种,如核糖体 RNA 基因 *rrn*,转运 RNA 基因 *trn*,RNA 聚合酶基因 *rpo* 等。第三类是与脂肪酸等物质合成相关的生物合成基因,如 *accD* 基因、*cemA* 基因等 5 种。第四类为功能未知的基因 10 种,如编码较长的蛋白质基因 *ycf1* 和 *ycf2* 等(图 1,表 1)。

2.2 大苞半蒴苣苔叶绿体基因重复区边界特征

本研究选取大苞半蒴苣苔近缘种 10 种植物进行叶绿体基因组边界分析(图 2)。从上至下依次为单座苣苔(P820508 *Hemiboea ovalifolia*)、大苞半蒴苣苔(He001 *H. magnibracteata*)、小苞半蒴苣苔

(OQ799918 *H. parvibracteata*)、短茎半蒴苣苔(OQ799916 *H. subacaulis*)、齿叶半蒴苣苔(OP820512 *H. fangii*)、全叶半蒴苣苔(OP820511 *H. integra*)、永福半蒴苣苔(OP820510 *H. yongfuensis*)、中越半蒴苣苔(OP820509 *H. sinovietnamica*)、绥阳半蒴苣苔(OQ799920 *H. suiyangensis*)、翅茎半蒴苣苔(OR195434 *H. pterocaulis*)。

结果显示,所选取植物的叶绿体基因组共有 4 个边界,即 LSC 与 IRb 边界称为 JLB,IRb 与 SSC 形成的边界称为 JSB,SSC 与 IRa 形成的边界称为 JSA,IRa 与 LSC 形成的边界称为 JLA。其中,JLB 均在 *rps19* 基因中。在选取的 10 个种中,超过半数以上物种的 *rps19* 基因均跨越 2 个区(LSC 和 IRb),只有大苞半蒴苣苔和全叶半蒴苣苔中 *rps19* 基因仅在 LSC,且距离 IRb 区的步长为 70 bp 和 91 bp。单座苣苔、短茎半蒴苣苔、永福半蒴苣苔、中越半蒴苣苔、绥阳半蒴苣苔和翅茎半蒴苣苔 6 种的 *rps19* 基因有 243 bp 在 LSC 区,36 bp 在 IRb。小苞半蒴苣苔中 *rps19* 基因有 237 bp

在 LSC 区,42 bp 在 IRb 区。齿叶半蒴苣苔 *rps19* 基因中有 249 bp 在 LSC 区,30 bp 在 IRb 区。JSB 边界均在 *ndhF* 基因和 *ycf1* 基因中,但大苞半蒴苣苔的 *ycf1* 基因仅在 IRb,未跨越到 SSC 区。翅茎半蒴苣苔在 IRb 区和 SSC 区未见到 *ycf1* 基因。JSA 边界没有扩

张和收缩的痕迹,均在 *ycf1* 基因中,即 *ycf1* 基因均跨越 SSC 区和 IRa 区。JLA 界均在 *trnH* 的基因片段上,即 *trnH* 基因跨越 IRa 区和 LSC 区,大部分在 LSC 区,只有 6 bp 左右在 IRa 区。不同的是,大苞半蒴苣苔中 *trnH* 基因仅在 LSC 区,距离边界 92 bp。

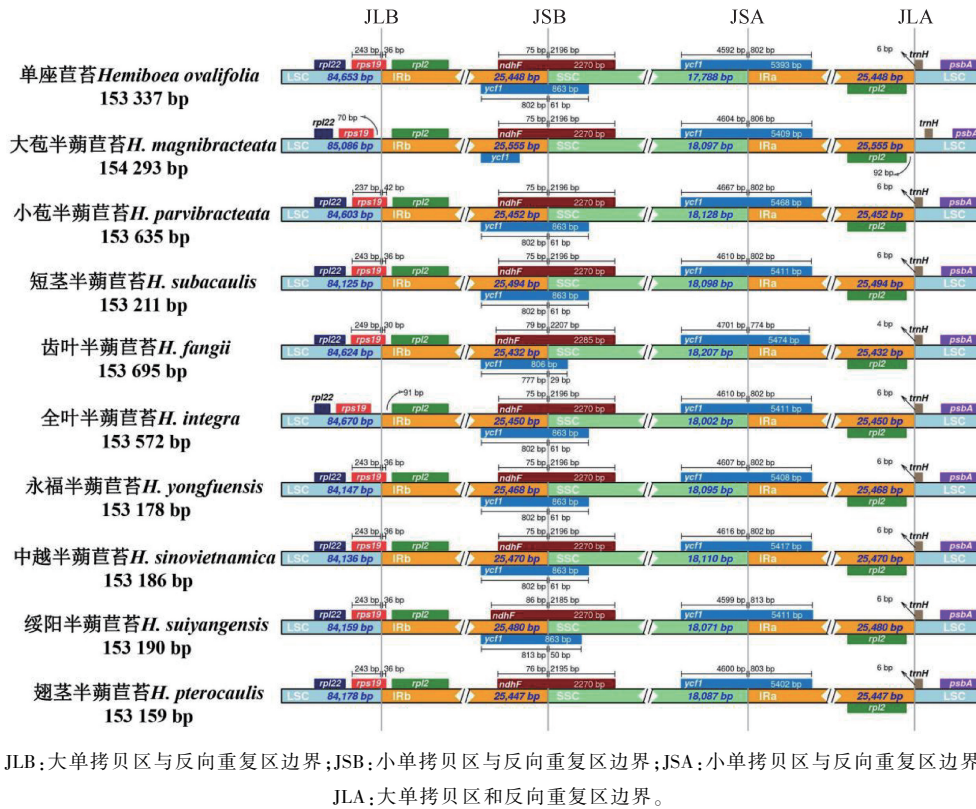


图 2 半蒴苣苔属植物叶绿体基因组 IR/SC 边界的扩张与收缩

Figure 2 Comparison of the border regions of IR/SC among the chloroplast genomes of *Hemiboea*.

从以上分析结果来看,大苞半蒴苣苔与半蒴苣苔属近缘种在进化过程中的 IR 边界区存在不同程度的收缩或扩张,且不同物种之间显示出一定差异。但总体而言,IR 区的变化幅度较小,叶绿体基因组较为保守,该区域的基因可用于作为等级较高(如属等)分类群的划分依据。

2.3 大苞半蒴苣苔叶绿体基因变异分析

本研究以单座苣苔的叶绿体基因组为参考,将大苞半蒴苣苔在内的 9 个种的叶绿体基因组使用 mVISTA 在线工具进行全序列对比分析(图 3)。结果表明,半蒴苣苔属的植物叶绿体基因组的基因差异性较小,不同区域的基因片段相对较保守。单座苣苔和翅茎半蒴苣苔的基因序列相似性较高,而齿叶半蒴苣苔和小苞半蒴苣苔基因序列具有较高的一致性,亲缘关系更近;全叶半蒴苣苔与上述 4 个物种的基因组序列整体上较为相似,说明这 5 个物种的亲缘关系较近。大苞半蒴苣苔、短茎半蒴苣苔和绥阳半蒴苣苔的基因序列具有较

高的相似性(图 3),但与单座苣苔的叶绿体基因组差异性较大。永福半蒴苣苔和中越半蒴苣苔与其它物种叶绿体基因组差异较大,尤其是永福半蒴苣苔,每个信息位点均与其它物种有明显区别(图 3),说明永福半蒴苣苔与单座苣苔的叶绿体基因组差别最大,亲缘关系最远。

结合边界分析(图 2),半蒴苣苔属植物总体上在 LSC 区 4 000~18 000 bp、27 000~33 000 bp、46 000~50 000 bp 等区域变异程度较高;SSC 区在 112 000~114 000 bp 的域变位点较多,IRa 区中 120 000~128 000 bp 变异位点较多。整体上来看,LSC 区变异位点较多,SSC 区变异位点仅次于 LSC 区,IR 区变异位点最少。其中,IRa 区相对于 IRb 区的变异位点要多。因此,LSC 区变异程度较高,该区域的基因可作为鉴定种内和种间的 DNA 条形码片段;IRa 区和 IRb 区相对较为保守,可作为鉴定科与科之间和属与属之间的 DNA 条形码片段。变异程度较高基因片段有 *rps16*、*trnQ-UUG*、*trnQ-UCC*、*trnR-UCU*、

atpI、*petN*、*trnL-UAA*、*trnF-GAA* 等,这些基因片段的基因编码区变异程度较大,存在显著差异,可用于作为鉴定种内和种间物种的 DNA 条形码片段;*matK*、

rpoC2、*rpoB*、*psaB*、*ycf2*、*rrn16*、*rrn23* 等基因编码片段较为保守,可作为鉴别科之间和属之间的 DNA 条形码片段。

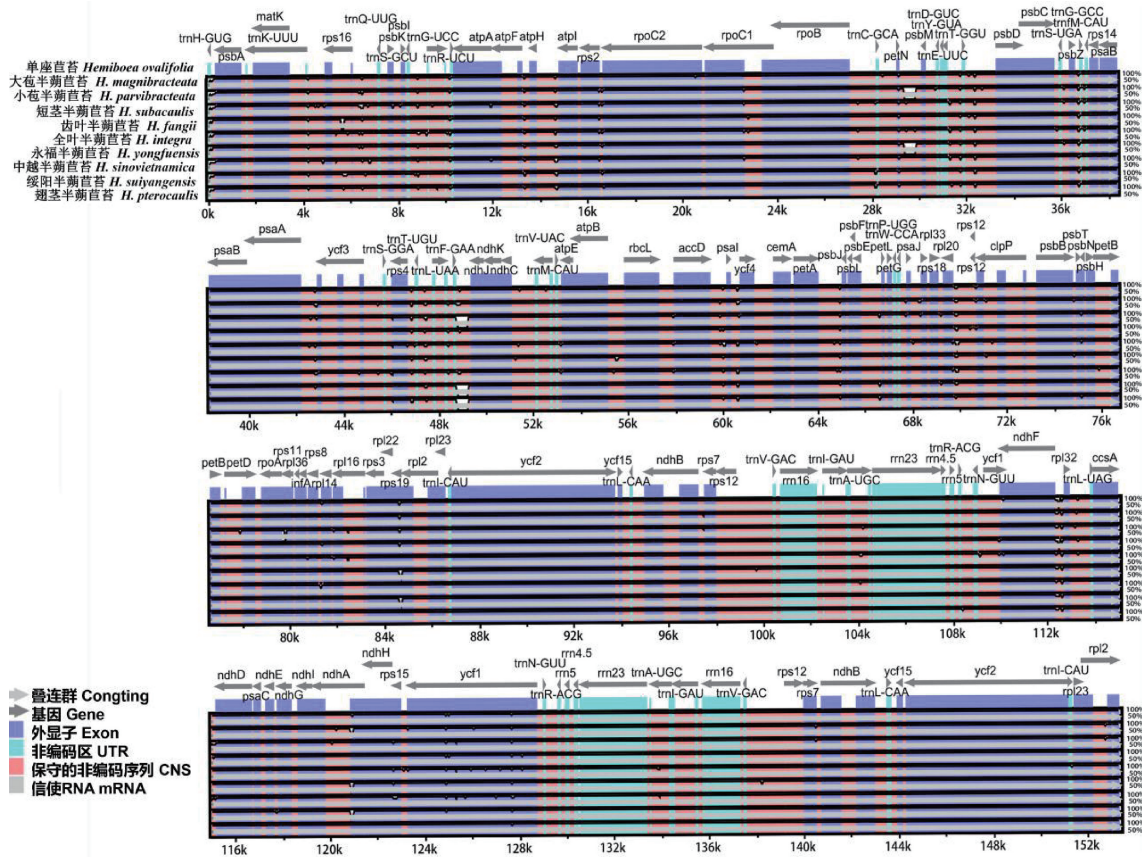


图 3 半蒴苣苔属植物叶绿体基因组共线性分析

Figure 3 Collinearity analysis of chloroplast genome in the *Hemiboea*.

2.4 基因片段 *trnL_F* 和 *rps16* 核苷酸多态性分析

核苷酸多态性可以反映出不同物种在居群水平上核酸序列的变异大小。变异程度较高的区域可为种群遗传多样性的研究提供潜在分子标记,也可用于基因区和基因间区的识别与多样性分析^[19-21],核苷酸差异的均值称为 P_i 值^[8]。

本研究基于 GenBank 数据库和叶绿体基因组差异性分析,筛选出适合进行种间系统发育关系研究的基因片段,并对其核苷酸多态性分析。筛选出变异位点较多的 2 个基因片段(共长 1 960 bp),第一个基因 *trnL_F* 有 3 个核苷酸多态性区,每个区域的差异碱基对都在 100 bp 及以上,且 P_i 值最高约为 0.04;第二个基因 *rps16* 有 5 个核苷酸多态性区,每个区域的差异碱基对普遍高于 50 bp,且 P_i 值最高为 0.06,有 3 个区域的 P_i 值均高于 0.04(图 4)。这说明,大苞半蒴苣苔与近缘种之间的多态性较高,每个核苷酸位点上变异较大。因此,将 *trnL_F* 和

rps16 两个基因片段用于鉴别种间和种内关系是可靠的。

2.5 系统发育分析

叶绿体基因的分析表明,*trnL_F* 和 *rps16* 基因片段适合用于种内和种间的物种鉴定。因此,本研究选择包含大苞半蒴苣苔在内的半蒴苣苔属 24 个种和 2 个石蝴蝶属 (*Petrocosmea*) 的物种作为外类群(表 2),应用贝叶斯模型构建系统发育关系(图 5)。整棵树的拓扑结构和分辨率较高,总体来看,整棵树的树形总体分为两大分枝,其支持率为 1.00,可信度高。其中,齿叶半蒴苣苔单独为一枝,其余所有物种聚为一大枝,支持率为 0.99。在这个大分枝中,小苞半蒴苣苔也单独为一枝,其余又聚为一大枝,支持率为 1.00;在这一大分枝中,总共进化出了 5 个分枝,支持率均在 0.85 以上,结果可信。

在 5 个进化枝中,最顶端的进化枝类群较多,支持率最高为 1.00,可信度最强。该进化枝又分为 2 个小进化枝,底端小分枝支持率为 1.00,可信度较

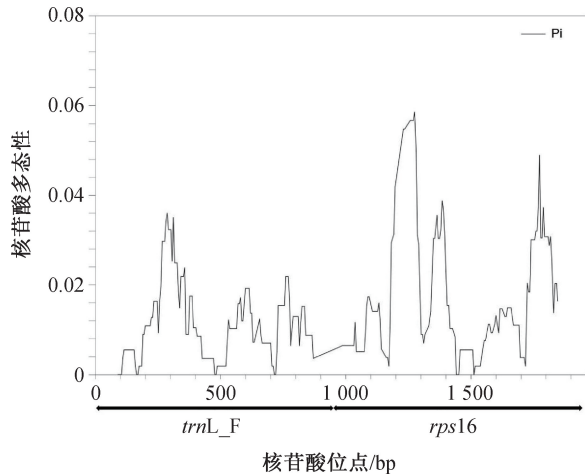


图 4 大苞半蒴苣苔及近缘种植物叶绿体基因 *trnL_F* 和 *rps16* 核苷酸多态性分析

Figure 4 Nucleotide diversities of the chloroplast genomes (*trnL_F* and *rps16*) of *H. magnibracteata* and related species

强,主要类群为华南半蒴苣苔 (*H. follicularis*)、短茎半蒴苣苔和弄岗半蒴苣苔 (*H. longgangensis*),结果可信。顶端的分枝支持率为 0.86,大于 0.85,结果可信。在这个进化枝中,大苞半蒴苣苔和拟大苞半

蒴苣苔 (*H. pseudomagnibracteata*) 亲缘关系较近,互为姐妹类群,支持率为 0.95。最顶端的这个分枝支持率为 1.00,该枝涵盖了王氏半蒴苣苔 (*H. wangiana*)、龙州半蒴苣苔 (*H. lungzhouensis*)、红苞半蒴苣苔 (*H. rubibracteata*)、麻栗坡半蒴苣苔 (*H. malipoensis*)、中越半蒴苣苔和贵州半蒴苣苔 (*H. cavaleriei*)。王氏半蒴苣苔和龙州半蒴苣苔互为姐妹类群,支持率为 0.91,结果可信。

第二个进化枝支持率为 0.86。该进化枝包括台湾半蒴苣苔 (*H. bicornuta*)、全叶半蒴苣苔和毛果半蒴苣苔 (*H. flaccida*)。其中,台湾半蒴苣苔和全叶半蒴苣苔互为姐妹类群,支持率为 1.00,可信。第三个分枝支持率为 0.98,可信度较强。该枝包括的类群有翅茎半蒴苣苔、降龙草 (*H. subcapitata*) 和纤细半蒴苣苔 (*H. gracilis*)。其中,翅茎半蒴苣苔和降龙草互为姐妹类群,支持率为 1.00。第四分枝为绥阳半蒴苣苔、永福半蒴苣苔和峨眉半蒴苣苔 (*H. omeiensis*),支持率为 0.85。第五分枝为单座苣苔和柔毛半蒴苣苔 (*H. mollifolia*),支持率为 0.94,可信度较强。

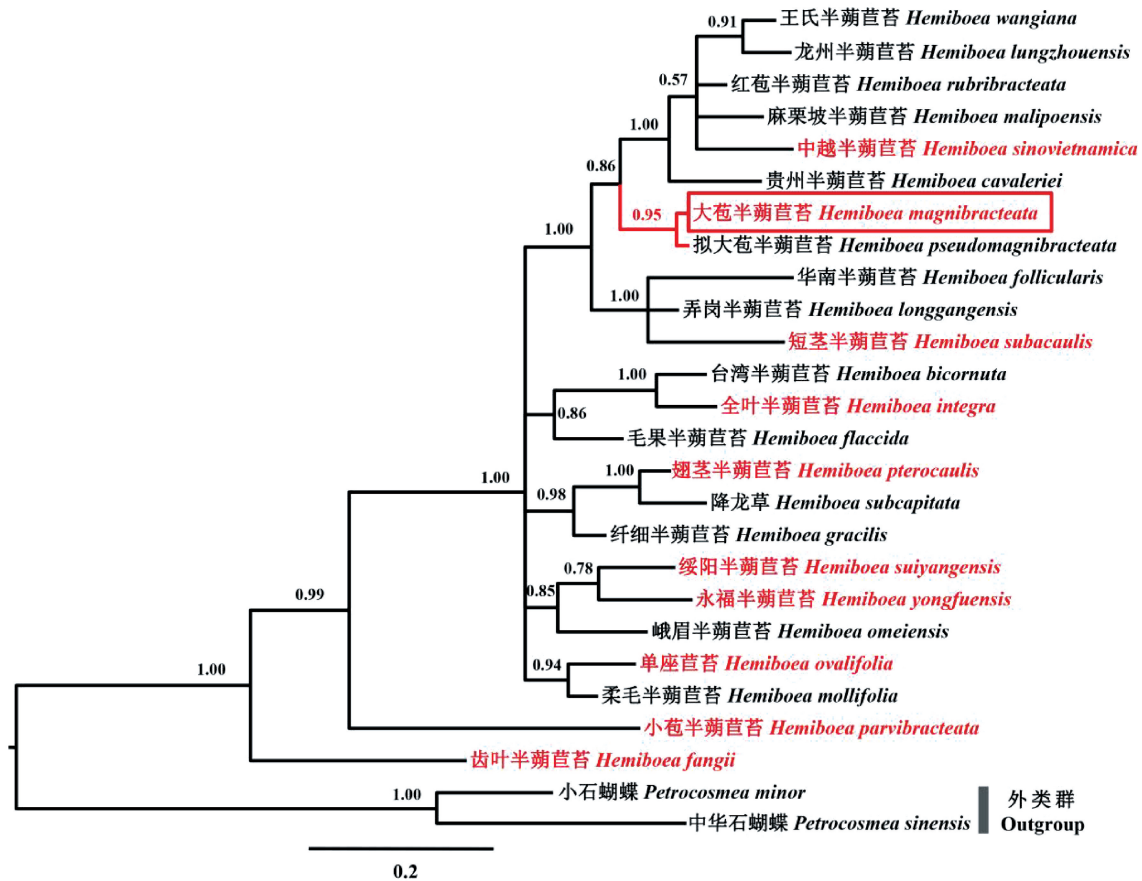


图 5 基于叶绿体基因 *trnL_F*、*rps16* 序列构建半蒴苣苔属的系统发育树

Figure 5 Phylogenetic tree constructed with sequences of *trnL_F* and *rps16* from *Hemiboea*

从整体树形的拓扑结构、支持率和分辨率上来看,选取 *trnL_F* 和 *rps16* 两个基因片段进行建立系统发育树是合理的。整体大分枝上的支持率较高,可信度较强。然而,2 个小分枝支持率不到 0.85,原因在于选取的基因片段上核苷酸多态性高的基因位点偏少,分辨率不够。因此,后期研究中除了选用 *trnL_F* 和 *rps16* 基因以外,可以进一步筛选出核苷酸多态性高的基因片段,再与本研究选用的 2 个叶绿体基因片段结合,以提高建树的分辨率。

3 讨论

3.1 叶绿体基因组特征分析

叶绿体基因组结构复杂,常在同一个 DNA 分子中包含基因组序列的多个拷贝,且不遵循孟德尔遗传法则,表现出非亲本遗传和营养分离。叶绿体有独特的 DNA、mRNA 和蛋白质合成机制。众多叶绿体蛋白质由核基因编码,合成后输送进叶绿体^[19]。在演化过程中,叶绿体基因组不断被简化。一些基因的删除是因为偶然发生的 DNA 复制错误或者异常的 DNA 重组事件导致的,从而造成多余基因的丢失(与其他细胞器基因或者核基因的功能重叠)。基因组减小的趋势在现有植物的线粒体基因组中有所减弱,但在叶绿体基因组中还在继续^[19]。

实际上,叶绿体基因组中的部分基因在不同物种中都普遍存在丢失现象,如 *accD*、*ycf1*、*ycf2*、*ycf4*、*psaI*、*rpoA*、*rpl20*、*rpl22*、*rpl23*、*rpl33* 和 *rps16*;也有部分基因在特定的某一个或少数物种中发现丢失,如 *psbJ*、*rps2*、*rps14* 和 *rps19*^[19]。因此,在半蒴苣苔属的物种中,每个种叶绿体基因数的不同很可能是由于有部分基因丢失导致的。例如,在边界分析中,翅茎半蒴苣苔在 JSB 边界就少了 1 个 *ycf1* 的基因。此外,本研究还发现,包括大苞半蒴苣苔在内的 10 个半蒴苣苔属物种的叶绿体基因组长度并不完全一致,在 153 159~153 695 bp 范围内变化。

与线粒体基因和核基因相比,叶绿体基因结构保守,且进化较慢。除了反向重复区外,所有叶绿体基因均为单拷贝,几乎不存在旁系同源基因和来自线粒体以及核的外源 DNA 的干扰^[20]。核基因包含的遗传信息和复杂性均远远高于叶绿体基因;线粒体基因组的大小也要大于叶绿体基因组,且在各植物类群中变异很大。除此之外,基因组之间的横向转移使线粒体基因组中存在大量外源基因插入现象,从而对其基因结构的稳定性和保守性产生影响^[20]。由此可见,叶绿体基因组的基因结构较为稳

定和保守,适合用于系统发育的分析。其中,叶绿体基因组中的反向重复序列(IR)是最保守的区域。问题在于,其边界区序列可能会向外延伸扩张,也可能向内部收缩,从而导致相关基因拷贝数的变化,或者导致边界区域假基因的产生^[21]。这是叶绿体基因组进化中的共有现象,也能很好的反映本研究中 10 种半蒴苣苔属植物的叶绿体基因组长度不同和基因数不同的主要原因。此外,本研究还比较了同一叶绿体基因组中的反向重复序列(IR)和单拷贝序列(SC)上 DNA 的核苷酸替换率,发现 IR 区较 SC 区序列的核苷酸替代速率要低,即使是 IR 基因间区也是如此^[5]。有研究发现,叶绿体基因组中内含子进化速率要比 CDS 区快,因此,叶绿体蛋白编码基因可分别为不同层级的系统分类学研究提供依据。

3.2 系统发育及遗传育种分析

为了厘清大苞半蒴苣苔在半蒴苣苔属中的系统学位置,本研究在叶绿体基因组特征分析的基础上,选取适合的基因片段构建了半蒴苣苔属的 BI 树。从系统发育树来看,半蒴苣苔属是单系类群,这与前人的研究吻合^[1-3,7-8,22-23]。值得注意的是,大苞半蒴苣苔相对该属近缘物种来说是进化类群,和拟大苞半蒴苣苔亲缘关系较近,为姐妹类群,支持率为 0.95。大苞半蒴苣苔主要分布于广西西北部和贵州南部,花期为每年 8 月;而拟大苞半蒴苣苔主要分布于广西,花期为每年 10 月。从其形态性状上来看,两者的宏观形态性状较为相似,它们的区分需要观察更细的形态性状,如大苞半蒴苣苔茎不分枝具棱,雄蕊为 2,子房偏球型;而拟大苞半蒴苣苔茎具分散的褐色斑点四边形,雄蕊 2~3,子房线形,且两者花期不重合。虽然二者在宏观性状上的区分存在挑战,但分子系统树能很好的将其区分开来。形态性状与分子证据的吻合说明这 2 个种的系统位置可信度较高。除此之外,单座苣苔和柔毛半蒴苣苔的亲缘关系很近,形态性状也较为相似,在该系统发育树上也聚在一起,且支持率为 0.94,可信度强。翅茎半蒴苣苔和降龙草也聚为一枝,支持率为 1.00。从整棵系统树来看,齿叶半蒴苣苔和小苞半蒴苣苔为基部类群,与两者花冠筒喉部均有 2 条龙骨状凸起这一祖征吻合。综上,*rps16* 和 *trnL_F* 基因序列建立的系统发育树可信。

在系统发育树上,每个分枝的所有物种都是由同一个祖先进化而来。因此,齿叶半蒴苣苔为最原始类群,与该种互为姐妹类群的物种进化出了多达

23个种。从图5可以看出,拟大苞半蒴苣苔是从大苞半蒴苣苔中分离出来。它们互为姐妹类群,源于一个共同祖先。近年来,该属新物种还在不断发表,说明该属物种本底数据不明,原因可能是该属为草本,主要生活在喀斯特山区。众多小环境形成不同程度的地理隔离,从而进化出很多新物种。进化的基本过程就是种群或者物种的遗传特性发生变化的过程。这一变化始于生物体携带的遗传物质的变化,我们统称为突变^[24]。突变并非进化的原因,但却是进化的必要条件,是不可或缺的要素。其中,DNA的插入或缺失是一种常见的突变。这一过程涉及单个或者多个碱基对。如果一段编码序列被插入或缺失1个碱基对,那么三联体读码框就会移动1个核苷酸的位置。于是,下游的三联体会被读取为不同的密码子,再翻译成不同的氨基酸^[4]。与近缘种相比,大苞半蒴苣苔叶绿体基因组在JSB边界上的*ycf1*基因比其它物种缺失多个碱基对,而翅茎半蒴苣苔在该边界上缺失了*ycf1*基因。

因此,大苞半蒴苣苔和翅茎半蒴苣苔目前所在节点处预期可能会演化出多个物种或者进化出新的分枝。在分子育种方向也可以从其缺失和插入基因入手,通过功能基因的调控选育出理想的目标品种。从其系统位置可知,该种若要选育出新的品种,可通过与其亲缘关系较近的拟大苞半蒴苣苔进行亲缘杂交,或者与第一进化枝的所有类群(因为它们具有共同的祖先)进行杂交选育出花期长,且花色佳的新品种。

4 结论

本研究首次完成了大苞半蒴苣苔叶绿体基因组序列的测序、组装注释和基础信息分析,揭示了大苞半蒴苣苔叶绿体基因组基本特征。大苞半蒴苣苔的叶绿体基因组结构呈四分体环状,共编码了132个功能基因。基于叶绿体基因片段构建的系统发育树,确定半蒴苣苔属植物为单系类群;随着大苞半蒴苣苔植物的演化,叶绿体基因组变大;通过高分化区筛选出*rps16*和*trnL_F*两个基因序列,可作为DNA条形码序列,用于区分大苞半蒴苣苔与近缘种之间、半蒴苣苔属内的物种鉴定。

此外,基于叶绿体基因片段构建的系统发育树,本研究明确了大苞半蒴苣苔在半蒴苣苔属植物中的系统位置,为后续大苞半蒴苣苔和其它半蒴苣苔属植物的分子遗传标记、叶绿体基因工程研究和遗传育种工作提供了理论指导。

参考文献(References)

- [1] 韦毅刚,文和群. 广西植物二新种[J]. 广西植物, 1995, (3): 216-219.
WEI Y G, WEN H Q. Two new species from Guangxi [J]. *Guihaia*, 1995, (3): 216-219.
- [2] 黄红. 贵州乌蒙山区野生苦苣苔科(Gesneriaceae)植物资源及群落多样性研究[D]. 贵阳: 贵州大学, 2022.
- [3] 任立超,施鹏程,陈招才. 浙江野生苦苣苔科观赏植物利用分类简述[J]. 花卉, 2019(14): 191-192.
REN L C, SHI P C, CHE Z C. A brief introduction to the utilization classification of wild Gesneriaceae ornamental plants in Zhejiang Province [J]. *Flowers*, 2019 (14): 191-192.
- [4] 罗哈睿,王罗云,张建国,等. 胡颓子科叶绿体基因组系统发育分析与演化趋势推断[J]. 西北植物学报, 2024, 44(6): 891-903.
LUO H R, WANG L Y, ZHANG J G, et al. Phylogenetic analysis and evolutionary trend inference of chloroplast genomes in Elaeagnaceae [J]. *Acta Botanica Boreali Occidentalia Sinica*, 2024, 44(6): 891-903.
- [5] YONG Q N, LI M J, LI Z, et al. Complete chloroplast genomes of 13 species of the *Impatiens* genus for genomic features and phylogenetic relationships studies [J]. *Scientific Reports*, 2025, 15: 4258.
- [6] WU X Y, WANG H F, ZOU S P, et al. Comparative analysis of the complete chloroplast genomes of thirteen *Bougainvillea* cultivars from South China with implications for their genome structures and phylogenetic relationships [J]. *PloS One*, 2024, 19(9): e0310091.
- [7] 李佳慧,黄章平,卢永彬,等. 半蒴苣苔复合群的多样性、地理分布和物种界限[J]. 广西植物, 2020, 40(10): 1477-1490.
LI J H, HUANG Z P, LU Y B, et al. Diversity, geographical distribution and species boundary of the *Hemiboea subcapitata* complex [J]. *Guihaia*, 2020, 40(10): 1477-1490.
- [8] 崔艳芳. 半蒴苣苔属的质体基因组比较和系统发育研究[D]. 南昌:南昌大学, 2023.
- [9] 张梅,祖奎玲,童琪,等. 基于DNA条形码对兜兰属(*Paphiopedilum*)植物的亲缘关系分析[J]. 分子植物育种, (2022-05-20) [2025-02-08]. <http://kns.cnki.net/kcms/detail/46.1068.S.20220519.1526.002.html>.
ZHANG M, ZU K L, TONG Q, et al. Phylogenetic relationships of *Paphiopedilum* plants based on DNA barcoding [J/OL]. *Molecular Plant Breeding*, (2022-05-20) [2025-02-08]. <http://kns.cnki.net/kcms/de>

- tail/46.1068.S.20220519.1526.002.html.
- [10] 林张翔. 稗草叶绿体基因组测序及其分子标记开发与利用[D]. 杭州: 浙江大学, 2015.
- [11] 木志强. 河南古槐种质资源遗传多样性研究[D]. 郑州: 河南农业大学, 2023.
- [12] 张梅, 杨加文, 龙成昌, 等. 绣球属(*Hydrangea* L.) 植物分子系统学及系统发育分析[J]. 西北植物学报, 2021, 41(2): 242-253.
- ZHANG M, YANG J W, LONG C C, et al. Molecular systematics and phylogenetic analysis of the *Hydrangea* (Hydrangeaceae)[J]. *Acta Botanica Boreali Occidentalia Sinica*, 2021, 41(2): 242-253.
- [13] 胡佳玉. 石松科马尾杉属的质体谱系基因组学研究[D]. 北京: 中国科学院大学, 2021.
- [14] 钱俊. 丹参的叶绿体和线粒体基因组研究[D]. 北京: 北京协和医学院, 2014.
- [15] 林张翔. 稗草叶绿体基因组测序及其分子标记开发与利用[D]. 杭州: 浙江大学, 2015.
- [16] GREINER S, LEHWARK P, BOCK R. OrganelleGenomeDRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes[J]. *Nucleic Acids Research*, 2019, 47(W1): W59-W64.
- [17] 文艺. 桔梗科质体基因组进化研究[D]. 北京: 中国科学院大学, 2020.
- [18] PANZADE S A, CHIMOTE V P, SHINDE C S, et al. Identification of Soybean (*Glycine max* L.) segregants for kunitz trypsin inhibitor and lipoxygenase free gene using linked molecular markers[J]. *Plant Cell Biotechnology and Molecular Biology*, 2025, 26(1-2): 33-44.
- [19] 陈晓颖, 胡本祥, 史嘉周, 等. 茜草叶绿体全基因组序列及其系统发育分析[J]. 西北植物学报, 2023, 43(11): 1855-1865.
- CHEN X Y, HU B X, SHI J Z, et al. Complete chloroplast genome and phylogenetic analysis of *Rubia cordifolia* [J]. *Acta Botanica Boreali Occidentalia Sinica*, 2023, 43(11): 1855-1865.
- [20] 郭佳星, 黄祥, 杨梅花, 等. 桦木科叶绿体基因组密码子偏好性及系统发育分析[J]. 中国农业科技导报, 2023, 25(10): 74-83.
- GUO J X, HUANG X, YANG M H, et al. Analysis of codon usage bias and phylogenetic in chloroplast genome of Betulaceae[J]. *Journal of Agricultural Science and Technology*, 2023, 25(10): 74-83.
- [21] 高丹, 向小果, 张强, 等. 半蒴苣苔属植物染色体制片优化及染色体数目和倍性研究[J]. 广西植物, 2023, 43(5): 833-845.
- GAO D, XIANG X G, ZHANG Q, et al. Optimization of chromosome preparation and chromosome number and ploidy in *Hemiboea*[J]. *Guihaia*, 2023, 43(5): 833-845.
- [22] 许为斌, 郭婧, 盘波, 等. 中国苦苣苔科植物的多样性与地理分布[J]. 广西植物, 2017, 37(10): 1219-1226.
- XU W B, GUO J, PAN B, et al. Diversity and distribution of Gesneriaceae in China[J]. *Guihaia*, 2017, 37(10): 1219-1226.
- [23] 张金权. 贵州南部区域野生苦苣苔科(Gesneriaceae)植物资源及群落物种多样性研究[D]. 贵阳: 贵州大学, 2022.
- [24] 周云, 韦妍妍. 基于叶绿体基因组解析半枫荷系统位置和进化[J]. 广西植物, 2024, 44(4): 670-681.
- ZHOU Y, WEI Y Y. Plastid phylogenomics resolving phylogeny and evolution of *Semiliquidambar cathayensis* [J]. *Guihaia*, 2024, 44(4): 670-681.

(责任编辑: 唐慧 郭芸婕)