

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

赵 祺^{1,2}, 余佳兴^{1#}, 秦宇雯^{1,2}, 任仙樱^{1,3}, 高海东⁴, 吴志刚^{1*}, 姜程曦^{1,3*}

1. 温州医科大学药学院, 浙江 温州 325035

2. 浙江省生物医药协同创新中心, 浙江 温州 325035

3. 安徽省九华山佛教医药研究所, 安徽 池州 242700

4. 南京集思慧远生物科技有限公司, 江苏 南京 210033

摘要: 目的 以药用植物菜头肾 *Championella sarcorrhiza* 为材料, 对其叶绿体基因组进行组装和序列分析, 为进一步开展菜头肾遗传与鉴定学研究奠定基础。方法 采用高通量测序技术首次对菜头肾叶绿体基因组进行测序, 以板蓝叶绿体基因组为参考, 进行特征分析和聚类关系研究。结果 生物信息学分析表明, 菜头肾叶绿体基因组总长为 144 713 bp, 具有典型的被子植物叶绿体基因组环状四分体结构, GC 含量为 38.35%; 共注释得到 129 个基因, 包括 84 个蛋白编码基因, 37 个 tRNA 基因和 8 个 rRNA 基因; 蛋白编码基因对应的密码子偏好使用 A/T 碱基。系统发育研究表明, 菜头肾与板蓝亲缘关系最近, 提示它们在进化上具有共线性。结论 建立了适于菜头肾植物完整叶绿体基因组组装及其特征分析的方法, 为菜头肾的鉴定、系统发育研究奠定了基础。

关键词: 菜头肾; 板蓝; 叶绿体; 基因组; 高通量测序

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2021)06-1744-07

DOI: 10.7501/j.issn.0253-2670.2021.06.024

Assembly and sequence analysis of chloroplast genome of *Championella sarcorrhiza* based on high-throughput sequencing

ZHAO Qi^{1,2}, YU Jia-xing¹, QIN Yu-wen^{1,2}, REN Xian-ying^{1,3}, GAO Hai-dong⁴, WU Zhi-gang¹, JIANG Cheng-xi^{1,3}

1. School of Pharmacy, Wenzhou Medical University, Wenzhou 325035, China

2. Zhejiang Biomedical Collaborative Innovation Center, Wenzhou 325035, China

3. Jiuhua Mountain Buddhist Medicine Research Institute of Anhui Province, Chizhou 242700, China

4. Nanjing Jisihuiyuan Biotechnology Co., Ltd., Nanjing 210033, China

Abstract: Objective The chloroplast genome of *Championella sarcorrhiza* was assembled and sequenced, laying a foundation for the further development of the genetics and identification of *C. sarcorrhiza*. **Methods** The chloroplast genome of *C. sarcorrhiza* was sequenced using high-throughput sequencing technology for the first time, taking the chloroplast genome of *Strobilanthes cusia* as a reference, the feature analysis and clustering relationship study were carried out. **Results** Bioinformatics analysis showed that the total length of the chloroplast genome of *C. sarcorrhiza* is 144,713bp, containing a typical angiosperm chloroplast genome circular tetrad structure, with a GC content of 38.35%. A total of 129 genes were annotated, including 84 protein-coding genes, 37 tRNA genes and eight rRNA genes. The codons corresponding to protein coding genes are preferred with A/T bases. Phylogenetic studies showed that the relationship between *C. sarcorrhiza* and *S. cusia* was the closest, suggesting that they were evolutionarily co-linear. **Conclusion** The research provided a method suitable for the assembly of the complete chloroplast genome and the analysis of characteristics of the whole chloroplast genome of the *C. sarcorrhiza*, which also could lay the foundation for the research on the identification and phylogeny of *C. sarcorrhiza*.

Key words: *Championella sarcorrhiza* C. Ling; *Strobilanthes cusia* (Nees) Kuntze; chloroplast; genome; high-throughput sequencing

收稿日期: 2020-12-01

基金项目: 国家中医药管理局中药饮片标准化建设项目 (ZYBZH-Y-SC-40); 云南大理药业股份有限公司课题 (KJHX1603); 温州医科大学人才科研启动项目 (OTJ18037)

作者简介: 赵 祺 (1994—), 女, 湖北襄樊人, 硕士, 研究方向为中药鉴定学。Tel: (0577)86591685 E-mail: 1297909219@qq.com

*通信作者: 姜程曦, 博士, 研究员, 硕士生导师, 研究方向为中药学。Tel: 18969715696 E-mail: jiangchengxi@126.com

吴志刚, 副教授, 研究方向为中药次生代谢产物研究。Tel: 15867176058 E-mail: wuzhigang177@126.com

#并列第一作者: 余佳兴 (1997—), 女, 四川成都人, 硕士, 研究方向为中药学。Tel: (0577)86591685

菜头肾 *Championella sarcorrhiza* C. Ling 来源于爵床科马蓝亚族黄猄草属多年生草本植物^[1], 别名土太子参、肉根马蓝, 是温州民间著名补肾方剂“七肾汤”的君药, 在临床上可与其他药物配伍或单独使用^[2-3]。由于菜头肾具有显著的补肾功能, 其在民间广泛使用, 并已实现了人工种植^[4]。目前, 菜头肾的相关研究主要集中于栽培、资源调查^[5]、化学成分鉴定^[6-7]等方面研究, 有关该植物叶绿体基因组组装、基因组成和序列分析等未见报道。

叶绿体是植物细胞进行光合作用的重要场所, 拥有一套完整的基因组, 可进行自主遗传。20 世纪 60 年代初, 在烟草 *Nicotiana tabacum* L.^[8]、地钱 *Marchantia polymorpha* L.^[9]中首次报道了叶绿体全序列。由于叶绿体基因组包含大量遗传信息, 不同物种之间具有良好的共线性, 被广泛应用于玉米 *Zea mays* L.^[10]、水稻 *Oryza sativa* L.^[11]、大豆 *Glycine max* (Linn.) Merr.^[12]等植物分子进化及系统发育的研究, 并在药用植物的遗传转化、基因工程和分子育种等方面发挥着重要作用^[13-17]。鉴于此, 本研究利用高通量测序技术对菜头肾叶绿体基因组进行测序, 分析组装后的叶绿体基因组结构、SSR (Simple Sequence Repeat, 简单重复序列) 位点、IR (inverted repeat, 反向重复序列) 区间特征比较及近缘物种亲缘关系, 为菜头肾的鉴定、系统发育研究奠定基础。

1 材料与方法

1.1 材料

菜头肾 *Championella sarcorrhiza* C. Ling 采自温州市瓯海区大罗山 (27°51'N, 120°40'E), 该地属亚热带海洋性季风气候区, 年均气温 17 °C, 无霜期 216~317 d, 年均降水量 1611 mm, 年均相对湿度 79%。经温州医科大学中药学教研室生药专家鲍康德副教授鉴定所有样品均来源于爵床科马蓝亚族黄猄草属菜头肾, 各样品均采集新鲜幼嫩的叶片, 液氮速冻下研磨, -80 °C 冰箱保存备用。

1.2 方法

1.2.1 DNA 基因提取 全基因组 DNA 采用植物基因组 DNA 提取试剂盒 (TIAN-GEN, DP305) 提取样本的总 DNA。

1.2.2 测序 总 DNA 样品经南京集思慧远生物科技有限公司进行测序, 得到测序 reads、质控。采用 Illumina HiSeq PE150 双末端测序策略进行建库测序, 建库类型为 350 bp DNA 小片段文库, 测序深

度为 10 倍。

1.2.3 叶绿体基因组组装 通过 Illumina HiSeq PE150 测序平台对样品进行测序, 用 MITObim 1.8^[18]组装软件进行组装。由于 NCBI 数据库中暂未收录菜头肾同属植物参考序列, 因此选用同为马蓝亚族的板蓝 *Strobilanthes cusia* (Nees) Kuntze (NC_037485.1) 为参考序列。然后用 Gap Closer^[19]1.12 软件 (<http://soap.genomics.org.cn/soapdenovo.html>) 对组装结果进行内洞修补, 最后去除冗余的短序列得到最后的组装结果。

1.2.4 基因组特征分析 先用 GeneMarkS (Version 4.17)^[20]软件对菜头肾叶绿体基因组进行编码基因预测, tRNAscan-SE 软件^[21] (Version 1.3.1)、rRNAmmer 软件^[22] (Version 1.2) 分别对 tRNA 和 rRNA 基因进行预测, 用 OGDRAW 软件^[23]呈现叶绿体基因组序列图, 再用 MISAL.0^[24]软件分析菜头肾叶绿体全基因组 SSR 位点, Find repeats 工具获得菜头肾、穿心莲等 6 种基因组的 IR 序列、大单拷贝区 (LSC) 和小单拷贝区 (SSC), MAFFT v7.394^[25] 软件进行序列比对, 最后使用 PhyML v3.0^[26]软件基于最大似然法构建系统发育树。

2 结果与分析

2.1 全基因组测序与叶绿体基因组组装

基于高通量测序, 发现菜头肾全基因组总长为 144 713 bp, 碱基为 5 196 698 700 bp, 具一个典型的 4 分区域结构, 包含 4 个区段: LSC 为 92 026 bp、SSC 为 17 825 bp、2 段反向互补重复的 IR 区 (IRA 和 IRB) 17 431 bp。GC 含量为 38.35%, 有效数据 Q₂₀ 值、Q₃₀ 值分别 96.58% 和 90.91%。IR、LSC 和 SSC 区域的 GC 值存在一定的差异, IR 区域的 GC 值最高, 为 45.47%; LSC 次之, 36.49%; SSC 区域的 GC 值最小, 仅 32.34% (表 1)。值得注意的是, 部分 *ycf1* 基因在 SSC 和 IRB 交界处重复 (图 1)。

2.2 叶绿体基因组结构基本特征

菜头肾叶绿体基因组共注释 129 个基因, 包括

表 1 菜头肾叶绿体基因组碱基组成

Table 1 Base composition of chloroplast genome in *C. sarcorrhiza*

区域	A/%	T/%	C/%	G/%	长度/bp	GC/%
LSC	31.1	32.4	18.6	17.8	92 026	36.4
SSC	33.8	33.7	17.0	15.2	17 825	32.3
IRA/IRB	27.3	27.1	21.3	24.0	17 431	45.4
总体	30.5	31.3	19.4	18.7	14 473	38.1

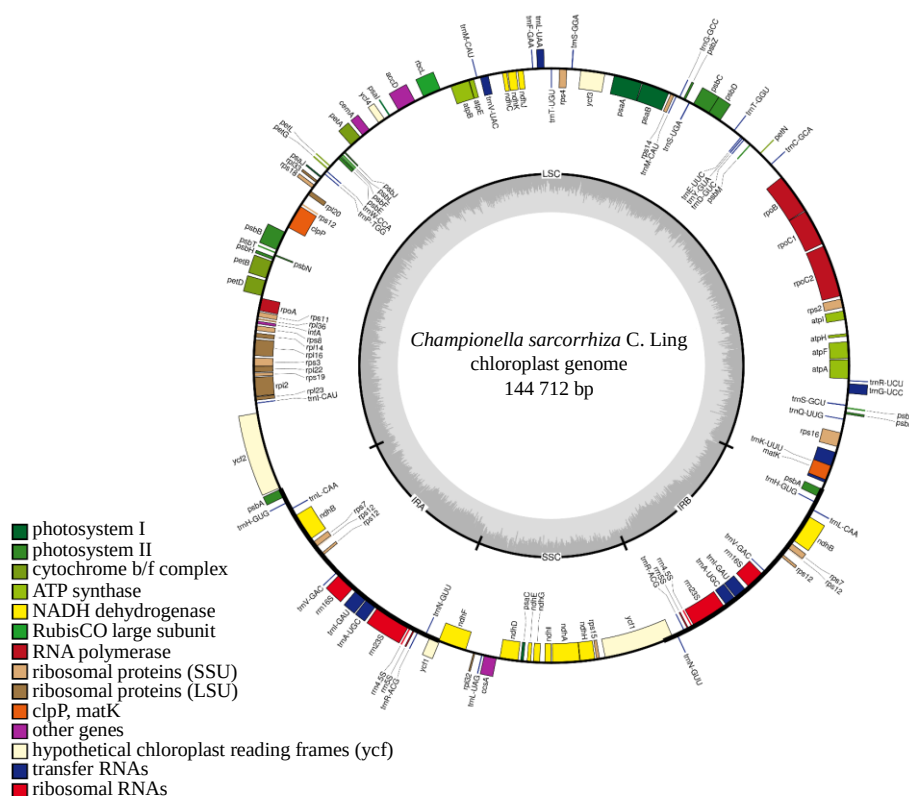


图 1 菜头肾叶绿体基因组

Fig. 1 Map of *C. sarcorrhiza* chloroplast genome

84 个蛋白编码基因, 37 个 tRNA 基因和 8 个 rRNA 基因。根据其功能可以把它们分为 4 大类: 分别是光合作用相关基因、表达相关基因、未知功能基因, 以及成熟酶基因 (*matK*)、囊膜蛋白基因 (*cemA*) 等其他基因。在这些基因中含有 24 个内含子和 17 个双拷贝基因(表 2), 包括 1 个光系统亚基(*psbA*)、4 个 rRNA 基因(*rrn16S*、*rrn23S*、*rrn4.5S*、*rrn5S*)、8 个 tRNA 基因(*trnA-UGC*、*trnH-GUG*、*trnI-GAU*、*trnL-CAA*、*trnM-CAU*、*trnN-GUU*、*trnR-ACG*、*trnV-GAC*)、1 个 NADH 脱氢酶亚基基因(*ndhB*)、2 个核糖体蛋白大小亚基编码基因(*rps7*、*rps12*)及 1 个未知功能蛋白基因(*ycf1*)。值得注意的是, 在菜头肾叶绿体基因组中存在一些未知功能的基因, 如 *ycf1*、*ycf2*、*ycf3*、*ycf4*, 还需要做进一步研究才能确定其功能。在 SSC 区共有 12 个 CDS 和 1 个 tRNA 基因(*trnL-UAG*)存在。然而绝大多数的基因存在于 LSC 区, 有 59 个 CDS 和 22 个 tRNA 基因(表 2), 其中 *trnK-UUU* 内含子最长, 其片段大小为 2450 bp, 包含 *matK* 基因。

2.3 SSR 分析及爵床科 IR 区边界比较分析

在菜头肾叶绿体全基因组中, 共搜索 184 个符

合条件的 SSR 位点(表 3), 多分布于 LSC 区, 其中包括 92 个单核苷酸重复单元, 10 个二核苷酸重复单元, 67 个三核苷酸重复单元, 9 个四核苷酸重复单元, 4 个五核苷酸重复单元和 2 个六核苷酸重复单元。在所有 SSR 中, 最多重复单元是 A/T, 然后依次是 AT/TA、TTG/TTA 和 ATA/ATA。由此可以看出, A/T 单元占优势, 具有碱基偏好性。这与菜头肾叶绿体全基因组富含 AT(含量为 61.65%)一致, 这种偏好性可能是由于 AT 比 GC 更容易解链有关。

通过与爵床亚科板蓝、穿心莲 *Andrographis paniculata* (Burm. F.) Nees、黄花恋岩花 *Echinacanthus lofouensis* (Lévl.) J. R. I. Wood、恋岩花 *Echinacanthus attenuates* Nees、鳄嘴花 *Clinacanthus nutans* (Burm. f.) Lindau 4 个属 5 个物种的叶绿体基因组 IR 区边界进行比较, 发现这 6 种植物叶绿体基因组长度 144 133~152 672 bp, 基因具有相对保守性, 从大到小依次为黄花恋岩花、恋岩花、穿心莲、鳄嘴花、菜头肾、板蓝。其中板蓝与菜头肾基因结构与种类最相似, *psbA* 与 *ycf1* 基因都有 2 份拷贝, *psbA* 基因绝大部分在

表 2 菜头肾叶绿体基因组编码的基因

Table 2 Genes present in chloroplast genome of *C. sarcorrhiza*

基因分类	基因分组	基因名称
光合作用相关基因	subunits of photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i>
	subunits of photosystem 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>
	subunits of NADH dehydrogenase	<i>ndhA</i> [*] , <i>ndhB</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>
	subunits of cytochrome b/f complex	<i>petA</i> , <i>petB</i> [*] , <i>petD</i> [*] , <i>petG</i> , <i>petL</i> , <i>petN</i>
	subunits of ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF</i> [*] , <i>atpH</i> , <i>atpI</i>
	large subunit of rubisco	<i>rbcL</i>
	subunits photochlorophyllide reductase	—
	proteins of large ribosomal subunit	<i>rpl14</i> , <i>rpl16</i> [*] , <i>rpl2</i> [*] , <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> , <i>rpl32</i> , <i>rpl33</i> , <i>rpl36</i>
表达相关基因	proteins of small ribosomal subunit	<i>rps11</i> , <i>rps12</i> ^{**△} , <i>rps14</i> , <i>rps15</i> , <i>rps16</i> [*] , <i>rps18</i> , <i>rps19</i> , <i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> [△] , <i>rps8</i>
	subunits of RNA polymerase	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> [*] , <i>rpoC2</i>
	ribosomal RNAs	<i>rrn16S</i> [△] , <i>rrn23S</i> [△] , <i>rrn4.5S</i> [△] , <i>trnC-GCA</i> , <i>rrn5S</i> [△]
	transfer RNAs	<i>trnA-UGC</i> ^{*△} , <i>trnR-ACG</i> [△] , <i>trnD-GUC</i> , <i>trnL-CAA</i> [△] , <i>trnE-UUC</i> , <i>trnN-GUU</i> [△] , <i>trnI-GAU</i> ^{*△} , <i>trnF-GAA</i> , <i>trnM-CAU</i> [△] , <i>trnG-GCC</i> , <i>trnG-UCC</i> ^{**} , <i>trnH-GUG</i> [△] , <i>trnI-CAU</i> , <i>trnK-UUU</i> [*] , <i>trnL-UAA</i> [*] , <i>trnL-UAG</i> , <i>trnP-TGG</i> , <i>trnQ-UUG</i> , <i>trnR-UCU</i> , <i>trnS-GCU</i> , <i>trnS-GGA</i> , <i>trnS-UGA</i> , <i>trnT-GGU</i> , <i>trnW-CCA</i> , <i>trnT-UGU</i> , <i>trnV-GAC</i> [△] , <i>trnV-UAC</i> [*] , <i>trnY-GUA</i>
	transfer RNAs	
其他基因	maturase	<i>matK</i>
	protease	<i>clpP</i> ^{**}
	envelope membrane protein	<i>cemA</i>
	acetyl-CoA carboxylase	<i>accD</i>
	c-type cytochrome synthesis gene	<i>ccsA</i>
	translation initiation factor	<i>infA</i>
未知功能基因	conserved hypothetical chloroplast ORF	<i>ycf1</i> [△] , <i>ycf1</i> [#] , <i>ycf2</i> , <i>ycf3</i> ^{**} , <i>ycf4</i>

*具有 1 个内含子的基因 **具有 2 个内含子的基因 △多拷贝基因 #假基因

*indicate one intron in the gene **indicate two introns in the gene △multiple copy gene #pseudogene.

表 3 菜头肾叶绿体基因组中 SSR 位点类型及数量

Table 3 Type and number of SSR loci of *C. sarcorrhiza* chloroplast genome

SSR 重复单元	长度/bp	SSR 数量/个	占比/%	SSR 重复单元	长度/bp	SSR 数量/个	占比/%
单核苷酸				GTT/TAA	19	6	8.96
A/T	95	86	93.48	TAT/TCT	10	6	8.96
C/G	34	6	6.52	TGA/TGC	6	2	2.99
二核苷酸				TTC/TTC	7	6	8.96
AT/TA	22	10	100.00	TTG/TTA	6	8	11.94
三核苷酸				CTG/GTA	6	3	4.48
AAC/AAG	6	3	4.48	四核苷酸			
AAT/AAT	7	5	7.46	AATA/AATT	6	4	44.45
ATG/ACT	6	2	2.99	ATAG/CAAT	6	2	22.22
AGA/AGC	6	4	5.97	TTAG/TTTC	6	3	33.33
ATA/ATA	7	7	10.45	五核苷酸			
ATT/CAA	10	5	7.46	AATGG/CCATT	6	2	50.00
CAG/CAT	6	2	2.99	TATAG/TTATT	13	2	50.00
CTT/GAA	6	6	8.96	六核苷酸			
GAT/GCA	6	2	2.99	AACTTA/TTTTAG	6	2	100.00

IRA 与 IRB 区内, 板蓝少部分基因 (约 379 bp) 位于 LSC/IRA 边界, 菜头肾 *psbA* 基因全部位于 IRA 区内; *ycfI* 基因分别位于 IRB/SSC 和 SSC/IRA 边界, 这可能与 IR/SSC 边界的扩张有关 (图 2), 位于 SSC/IRA 边界的 *ycfI* 基因长度为 5496 bp 和

5514 bp, 为正常基因, 但位于 IRB/SSC 边界的 *ycf1* 基因全长仅为 840 bp 和 834 bp, 它的 3' 端发生缺失, 该 *ycf1* 为一个假基因。穿心莲、鳄嘴花等其他物种没有 *psbA* 基因, 鳄嘴花 IR 区也没有 *ycf1* 基因, 基因差异最大。

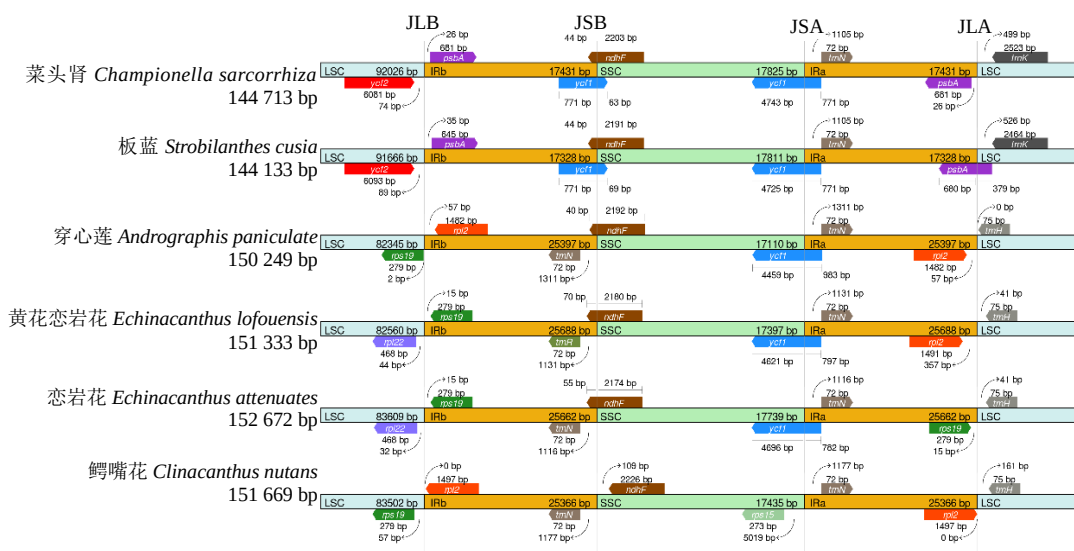


图 2 爵床科 6 个物种叶绿体全基因组 IR 边界的比较示意图

Fig. 2 Comparison schematic diagram of chloroplast genome-wide IR boundaries of six species of Acanthaceae

2.4 基于叶绿体基因组全序列的聚类分析

叶绿体基因组对系统发育研究具有重要意义^[27]。为了鉴定菜头肾在爵床科中的进化位置, 将测得的菜头肾叶绿体基因组序列与 NCBI 网站下载的其他 10 个物种叶绿体全基因组序列构建系统发育树。由于目前暂无黄猢草属植物叶绿体基因组 GenBank, 且爵床科植物叶绿体基因组研究有限, 在此选取了在分类上与菜头肾同一亚族的板蓝, 同一族的黄花恋岩花、黄花假杜鹃 *Barleria prionitis* L. 等 4 个物种, 同一亚科的鳄嘴花、穿心莲、孩儿草 *Rungia pectinata* (L.) Nees 3 个物种, 同一科的老鼠筋 *Acanthus ilicifolius* L. Sp. 和小花老鼠筋 *Acanthus ebracteatus* Vahl, 及不同科的日本紫茎 *Stewartia monadelphica* Sieb. et Zucc.。结果显示 (图 3), 不同科物种与爵床科距离最远, 说明亲缘关系较远, 此与《中国植物志》植物分类结果相一致^[28]; 包括菜头肾在内的 11 个爵床科物种集为一支, 属间关系比较清晰, 支持率为 100%; 菜头肾与板蓝组成的小分支是姐妹关系, 说明其亲缘关系密切, 此与《中国植物志》植物分类结果一致, 暗示它们在进化上具有共线性。

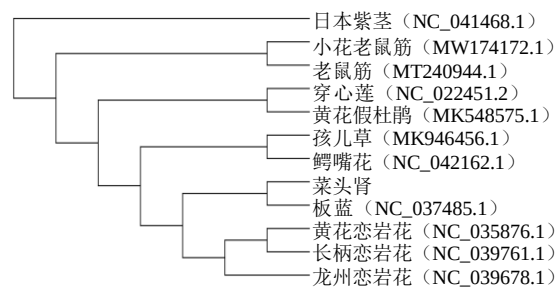


图3 基于叶绿体全基因组的12个物种的最大似然法 (ML) 聚类结果

Table 3 Cluster analysis of 12 species using complete chloroplast genome sequence by maximum likelihood (ML) method

3 讨论

菜头肾是温州民间特色中草药，主要作为验方“七肾汤”组方之一，根据临床病情的不同，可调整药物剂量，当肝炎迁延期、慢性肾炎腰痛时，重用菜头肾^[29-30]。此外，菜头肾曾被《中国物种红色名录》列为极危（CR）等级^[31]，因而具有重要的保护、研究和开发利用价值。本研究利用第2代测序技术，对爵床科植物菜头肾进行了叶绿体基因组测序，获得其全长 144 713 bp，注释得到 129 个基因；鉴定 184 个 SSR 位点，大多数位点具有 A/T 碱基偏好性。系统

进化分析表明菜头肾与已报道的同一亚族植物板蓝的亲缘关系最近,暗示它们在进化上具有共线性。

黄猗草属 *Championella* Bremek. 植物全世界共有 10 种,我国有 7~9 种。目前,黄猗草属植物遗传背景工作研究有限,本实验首次报道该属物种叶绿体基因组序列,不仅可以丰富叶绿体基因组序列的数量,还可以为研究菜头肾的鉴定、系统发育奠定基础。然而,与爵床科其它属所具有的物种多样性水平相比,基础研究有待加强。

通过对比分析,菜头肾与已发表的爵床科植物叶绿体基因组 GC 含量相似,基因构成和次序一致,体现了叶绿体基因组结构稳定,总体进化速率较低^[32-34]。因为叶绿体基因组在总体上的保守性,在利用叶绿体基因组信息时,通常会筛选其中的高变区如 *matK*、*trnH-psbA*、*rbcL* 等常用的叶绿体片段。值得注意的是,在该基因组中存在一些未知功能的基因,如 *ycf1*、*ycf2*、*ycf3* 和 *ycf4*,还需要做进一步研究才能确定其功能。

SSR 广泛存在于原核生物和真核生物基因中,具有共显性:本研究经分析鉴定出 184 个 SSR,多分布于 LSC 区,主要由碱基 A 和 T 组成,这与爵床属叶绿体基因 SSR 分析结果类似^[35-36]。SSR 分布还具有物种特异性:二至六核苷酸重复类型 SSR 位点中,菜头肾转录组以三核苷酸重复为主,其优势重复基元类型是 TTG/TTA 和 ATA/ATA,共占比 22.39%;不同属植物穿心莲和板蓝以二核苷酸重复为主,其优势重复基元类型是 AT/AT 和 AG/CT,穿心莲占比 33.60%,板蓝占比 90.70%^[37-38]。SSR 丰富的变异位点,为今后开发 SSR 标记及进一步揭示菜头肾的系统进化提供了可靠的分子手段。

《中国植物志》记载:菜头肾暂仍保留在黄猗草属,然而其具根茎,侧根肉质增厚,这在马蓝亚族 (*Strobilantheae*) 实不多见,或许为一新属,有待进一步深入研究。通过对比分析,发现菜头肾与板蓝叶绿体基因长度相近(相差 580 bp),mRNA、rRNA、tRNA 分别均为 84、8、37 个,IR 区间结构相近,基于叶绿体全基因组序列构建的系统发育树中,爵床科内分支支持率全为 100%,且菜头肾与同一亚族的板蓝亲缘关系最近,构成姐妹关系。但因受黄猗草属物种叶绿体全基因组序列信息不全的影响,无法利用叶绿体基因组对菜头肾进行属内鉴别。因此,对菜头肾的鉴别和系统发育研究,后期还需从形态学、遗传学等进一步研究、验证。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 中国科学院中国植物志编辑委员会. 中国植物志 [M]. 北京: 科学出版社, 2002: 93.
- [2] 甘慈尧. 浙南本草新编 [M]. 北京: 中国中医药出版社, 2016: 324-325.
- [3] 林泉. 浙南药用植物中的一新种菜头肾 [J]. 植物分类学报, 1975, 13(2): 93-94.
- [4] 郑毅, 周建武, 陈小玲, 等. 珍稀药用植物菜头肾的人工栽培技术 [J]. 农业科技通讯, 2016(9): 252-253.
- [5] 姚振生, 陈京, 尤志勉, 等. 浙江省永嘉县道志地区菜头肾资源调查 [J]. 江西科学, 2007, 25(4): 397-401.
- [6] 张森尧, 潘亚琴, 徐攀, 等. 中药菜头肾中化学成分初步分析 [J]. 世界中西医结合杂志, 2008, 3(4): 204-205.
- [7] 屈玲霞, 王扬, 刘永巧, 等. 菜头肾化学成分研究 [J]. 中国现代中药, 2019, 21(4): 438-440.
- [8] Shinozaki K, Ohme M, Tanaka M, et al. The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression [J]. *EMBO J*, 1986, 5(9): 2043-2049.
- [9] Ohyama K, Fukuzawa H, Kohchi T, et al. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA [J]. *Nature*, 1986, 322(6079): 572-574.
- [10] Arévalo-Gallegos S, Varela-Rodríguez H, Lugo-Aguilar H, et al. Transient expression of a green fluorescent protein in tobacco and maize chloroplast [J]. *Electron J Biotechnol*, 2020, 45: 1-9.
- [11] Badro H, Furtado A, Henry R. Relationships between Iraqi rice varieties at the nuclear and plastid genome levels [J]. *Plants* (Basel), 2019, 8(11): E481.
- [12] Jin L, Wan Z P, Wei D, et al. Characterization of the complete chloroplast genome of black soybeans, *Glycine max* (L.) Merr. (legume) [J]. *Mitochondrial DNA Part B*, 2019, 4(2): 3309-3311.
- [13] 李双双, 薛龙飞, 苏爱国, 等. 高等植物线粒体基因组测序和序列分析 [J]. 中国农业大学学报, 2011, 16(2): 22-27.
- [14] 张韵洁, 李德铎. 叶绿体系统发育基因组学的研究进展 [J]. 植物分类与资源学报, 2011, 33(4): 365-375.
- [15] 崔柳青, 李一帆, 潘卫东. 叶绿体基因工程研究进展 [J]. 生物技术通报, 2012(6): 1-6.
- [16] Bock R. Plastid biotechnology: Prospects for herbicide and insect resistance, metabolic engineering and molecular farming [J]. *Curr Opin Biotechnol*, 2007, 18(2): 100-106.
- [17] 范李强, 胡欢, 郑洪蕾, 等. 响叶杨(杨属)叶绿体基因

- 组测序与比较分析 [J]. 四川大学学报: 自然科学版, 2018, 55(1): 165-171.
- [18] Freyer R, Hoch B, Neckermann K, *et al.* RNA editing in maize chloroplasts is a processing step independent of splicing and cleavage to monocistronic mRNAs [J]. *Plant J*, 1993, 4(4): 621-629.
- [19] Luo R B, Liu B H, Xie Y L, *et al.* Erratum: SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler [J]. *Gigascience*, 2015, 4: 30.
- [20] Besemer J, Lomsadze A, Borodovsky M. GeneMarkS: A self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions [J]. *Nucleic Acids Res*, 2001, 29(12): 2607-2618.
- [21] Lowe T M, Eddy S R. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence [J]. *Nucleic Acids Res*, 1997, 25(5): 955-964.
- [22] Lagesen K, Hallin P, Rødland E A, *et al.* RNAmmer: Consistent and rapid annotation of ribosomal RNA genes [J]. *Nucleic Acids Res*, 2007, 35(9): 3100-3108.
- [23] Lohse M, Drechsel O, Bock R. OrganellarGenomeDRAW (OGDRAW): A tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes [J]. *Curr Genet*, 2007, 52(5/6): 267-274.
- [24] Biggeri A, Baccini M, Bellini P, *et al.* Meta-analysis of the Italian studies of short-term effects of air pollution (MISA), 1990-1999 [J]. *Int J Occup Environ Health*, 2005, 11(1): 107-122.
- [25] Katoh K, Standley D M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability [J]. *Mol Biol Evol*, 2013, 30(4): 772-780.
- [26] Guindon S, Dufayard J F, Lefort V, *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0 [J]. *Syst Biol*, 2010, 59(3): 307-321.
- [27] Hu Y H, Woeste K E, Zhao P. Completion of the chloroplast genomes of five Chinese *Juglans* and their contribution to chloroplast phylogeny [J]. *Front Plant Sci*, 2017, 7: 1955.
- [28] Ding P, Shao Y H, Li Q, *et al.* The complete chloroplast genome sequence of the medicinal plant *Andrographis paniculata* [J]. *Mitochondrial DNA A DNA Mapp Seq Anal*, 2016, 27(4): 2347-2348.
- [29] 胡欣欣. 浅议温州民间验方七肾汤的应用 [J]. 浙江中医杂志, 2015, 50(2): 131.
- [30] 张蕾蓉. 温州民间补肾草药和验方“七肾汤” [J]. 中草药, 2008, 39(6): 954-955.
- [31] 汪松, 解焱. 中国物种红色名录: 第 1 卷 [M]. 北京: 高等教育出版社, 2004: 393-395.
- [32] Jansen R K, Raubeson L A, Boore J L, *et al.* Methods for obtaining and analyzing whole chloroplast genome sequences [J]. *Methods Enzymol*, 2005, 395: 348-384.
- [33] Yaradua S S, Alzahrani D A, Albokhary E J, *et al.* Complete chloroplast genome sequence of *Justicia flava*: Genome comparative analysis and phylogenetic relationships among Acanthaceae [J]. *Biomed Res Int*, 2019, 2019: 4370258.
- [34] Yaradua S S, Alzahrani D A, Albokhary E J, *et al.* Corrigendum to “complete chloroplast genome sequence of *Justicia flava*: Genome comparative analysis and phylogenetic relationships among Acanthaceae” [J]. *Biomed Res Int*, 2020, 2020: 2049759.
- [35] Yaradua S S, Alzahrani D A, Albokhary E J, *et al.* Complete chloroplast genome sequence of *Justicia flava*: Genome comparative analysis and phylogenetic relationships among Acanthaceae [J]. *Biomed Res Int*, 2019, 2019: 4370258.
- [36] 钮峥洋. 亚洲广义爵床属叶绿体系统发育基因组学研究 [D]. 南京: 南京林业大学, 2020.
- [37] 李俊仁, 陈秀珍, 汤小婷, 等. 穿心莲转录组 SSR 位点信息分析 [J]. 中国中药杂志, 2018, 43(12): 2503-2508.
- [38] Chen H M, Shao J J, Zhang H, *et al.* Sequencing and analysis of *Strobilanthes cusia* (Nees) Kuntze chloroplast genome revealed the rare simultaneous contraction and expansion of the inverted repeat region in angiosperm [J]. *Front Plant Sci*, 2018, 9: 324.

[责任编辑 时圣明]