

引用格式:

吴民华, 陈依虹, 谭靖怡, 邹山扬, 叶晓霞, 黄琼林. 三桠苦叶绿体基因组结构和序列特征分析[J]. 湖南农业大学学报(自然科学版), 2024, 50(2): 42–49.

WU M H, CHEN Y H, TAN J Y, ZOU S Y, YE X X, HUANG Q L. Analysis on structure and sequence feature of chloroplast genome from *Melicope pteleifolia*[J]. Journal of Hunan Agricultural University(Natural Sciences), 2024, 50(2): 42–49.

投稿网址: <http://xb.hunau.edu.cn>



三桠苦叶绿体基因组结构和序列特征分析

吴民华¹, 陈依虹², 谭靖怡³, 邹山扬², 叶晓霞¹, 黄琼林^{1*}

(1.广东医科大学基础医学院, 广东 湛江 524023; 2.广东医科大学第一临床医学院, 广东 湛江 524023; 3.广东医科大学药学院, 广东 湛江 524023)

摘要: 为了探明岭南地区特色中药三桠苦的叶绿体基因组结构和序列特征, 以三桠苦叶片为材料, 采用试剂盒法提取 DNA 和构建测序文库, 利用 Novaseq 平台进行高通量测序, 再结合生物信息学手段进行序列拼接、注释和特征分析。结果显示: 三桠苦叶绿体基因组为 158 992 bp 的环状四分体分子, 含有 134 个基因; 在三桠苦叶绿体基因组内找到 86 个简单重复序列, 主要是 A/T 单核苷酸重复; 检测到 26 907 个密码子, 以 A/T 结尾的密码子有着更高的使用频次; 序列比较分析表明, 三桠苦与同科植物在非编码区存在更明显的序列变异; 系统进化树揭示不同来源的三桠苦亲缘关系最近, 但在基因组成和碱基序列上呈现出多样性。

关键词: 三桠苦; 叶绿体基因组; 结构特征; 序列比对

中图分类号: R282

文献标志码: A

文章编号: 1007–1032(2024)02–0042–08

Analysis on structure and sequence feature of chloroplast genome from *Melicope pteleifolia*

WU Minhua¹, CHEN Yihong², TAN Jingyi³, ZOU Shanyang², YE Xiaoxia¹, HUANG Qionglin^{1*}

(1.School of Basic Medicine, Guangdong Medical University, Zhanjiang, Guangdong 524023, China; 2.The First Clinic Medical College, Guangdong Medical University, Zhanjiang, Guangdong 524023, China; 3.School of Pharmacy, Guangdong Medical University, Zhanjiang, Guangdong 524023, China)

Abstract: To illustrate the structure and sequence characteristics of chloroplast genome from *Melicope pteleifolia*, the leaf of *M. pteleifolia* was used for DNA isolation and library construction. The high-throughput sequencing was fulfilled by Novaseq platform, and then acquired sequence was assembled, annotated and analyzed using bioinformatics means. The results revealed that the chloroplast genome of *M. pteleifolia* was a circular quadripartite molecule with a size of 158 992 bp and encoded 134 genes. Altogether 86 simple sequence repeat was found in the plastome and the A/T mononucleotide was predominate type. Total 26 907 codon was screened out and those ended with A/T were frequently used. Sequence comparison demonstrated more divergence and occurred in non-coding region between *M. pteleifolia* and other Rutaceae species. Phylogenetic tree indicated various chloroplast genome sequences of *M. pteleifolia* shared close relationship and exhibited diversity in gene composition and base sequence.

Keywords: *Melicope pteleifolia*; chloroplast genome; structure feature; sequence alignment

三桠苦(*Melicope pteleifolia*)是芸香科蜜茛萸属 乔木植物, 又称三丫苦、三叉苦、三岔叶、三支枪、

收稿日期: 2023–07–28

修回日期: 2024–04–01

基金项目: 广东省基础与应用基础研究基金项目(2018A030310116); 广东省大学生创新创业训练计划项目(S202310571113); 湛江市科技计划项目(2019A01008); 广东医科大学大学生创新创业训练计划项目(GDMU2022113)

作者简介: 吴民华(1981—), 男, 广东湛江人, 博士, 副教授, 主要从事中药分子药理研究, wugdmczp@gdmu.edu.cn; *通信作者, 黄琼林, 博士, 副教授, 主要从事医学生物化学研究, perfecthql@163.com

小黄散等,主产于广东、广西、海南,在湖南、福建、江西等省亦有较广分布^[1]。三桠苦的根、茎、枝、叶皆可入药,具有清热解毒、祛风除湿、消肿止痛等功效^[2]。三桠苦历来以野生资源使用,随着近些年的大量采伐,其野生资源蕴藏量锐减,无法保障药用和食用需求,且其生态环境、品种质量均受到不同程度的破坏^[3];因此,加强资源保护和栽培育种等基础研究已成为三桠苦资源可持续利用的当务之急。

陆生植物的叶绿体基因组具有保守的环状双链四段式结构,由一大一小的单拷贝区(SC)以及贯穿它们之间的一对反向互补区(IR)构成。叶绿体基因一般为 130 个左右,主要是与光合作用和叶绿体自身复制相关的基因^[4-5]。尽管叶绿体基因组的结构和基因编码区序列比较稳定,但其基因间隔区和非编码区具有较快的进化速率,在不同物种之间有较明显的碱基差异^[6]。此外,在植物进化过程中,IR 区通常发生扩张和收缩,会引起叶绿体基因组及其 4 个组成部分在长度上的变异^[7]。叶绿体基因组分析已广泛应用于植物的物种鉴别^[8]、遗传多样性^[9]、系统进化^[10]等方面,是开展分子生物学研究的重要手段之一。

本研究中,拟对三桠苦的叶绿体基因组进行序列解读,并将其与相关物种的叶绿体基因进行比较分析,旨在阐明其结构特点,从基因组层面为三桠苦的品种鉴定、居群进化、品种栽培等研究提供参考。

1 材料和方法

1.1 植物样品的采集和处理

三桠苦新鲜叶片采集于广西壮族自治区百色市平果市新安镇,清洗干净后自然晾干,备用。采用北京康为世纪植物 DNA 磁珠法提取试剂盒(货号 CW2528)提取 DNA;采用美国因美纳 DNA 文库制备试剂盒(货号:FC-131-1024)制备测序文库。高通量测序由深圳惠通生物科技有限公司运用美国因美纳 Novaseq 技术平台完成。

1.2 叶绿体基因组的序列组装和基因注释

测序所得的三桠苦叶绿体基因组原始读序经质检和过滤处理后分别采用 SPAdes 3.11.0 和 PLANN 进行序列拼接和基因注释;采用 Bankit 将完成拼接和注释的三桠苦叶绿体基因组序列上传

至 GenBank,登记号为 OQ784937。通过 OGDRAW 进行三桠苦叶绿体基因组的图谱绘制。

1.3 叶绿体基因组的特点分析

采用 Codon W 计算三桠苦叶绿体基因组中所有密码子的相对同义密码子使用情况(RSCU),以评价三桠苦叶绿体基因组的密码子偏好性。按照黄琼林^[11]的方法,采用 MISA v2.1 网络工具搜索三桠苦叶绿体基因组中存在的简单重复序列(SSR),分析其组成类型和分布区域。

1.4 叶绿体基因组的序列比较

运用 IRscope 网络软件比较三桠苦、花椒(*Zanthoxylum bungeanum*)、臭节草(*Boenninghausenia albiflora*)、白鲜(*Dictamnus dasycarpus*)、臭檀吴萸(*Tetradium daniellii*)叶绿体基因组的 IR/SC 边界差异,再使用脚本将上述植物的叶绿体基因组序列转换并导入 mVISTA 软件进行序列多重比对。

1.5 系统进化分析

在 GenBank 下载 24 种(表 1)已发布的三桠苦及

表 1 在 GenBank 下载的叶绿体基因组序列信息

Table 1 Information of chloroplast genome sequences downloaded from GenBank

科	属	种	序列号
芸香科	蜜茛萸属	三桠苦 <i>Melicope pteleifolia</i>	MW046256
		三桠苦 <i>Melicope pteleifolia</i>	MW263046
		三桠苦 <i>Melicope pteleifolia</i>	MT784750
石椒草属	臭节草	<i>Boenninghausenia albiflora</i>	MZ569021
		白鲜属	白鲜 <i>Dictamnus dasycarpus</i>
	白鲜属	白鲜 <i>Dictamnus dasycarpus</i>	MZ569023
		花椒属	花椒 <i>Zanthoxylum bungeanum</i>
		两面针 <i>Zanthoxylum nitidum</i>	MK613864
		吴茛萸属	臭檀吴萸 <i>Tetradium daniellii</i>
	柑橘属	柚 <i>Citrus maxima</i>	MN782007
		甜橙 <i>Citrus sinensis</i>	DQ864733
	九里香属	九里香 <i>Murraya exotica</i>	MW722359
		千里香 <i>Murraya paniculata</i>	MT744742
	黄檗属	川黄檗 <i>Phellodendron chinense</i>	MT916287
		黄檗 <i>Phellodendron amurense</i>	KY707335
	山小橘属	山小橘 <i>Glycosmis pentaphylla</i>	KU949005
		小花山小橘 <i>Glycosmis parviflora</i>	MW714375
	黄皮属	黄皮 <i>Clausena lansium</i>	OL944012
		假黄皮树 <i>Clausena excavata</i>	KU949003
楝科	香椿属	香椿 <i>Toona sinensis</i>	MK954108
		红椿 <i>Toona ciliata</i>	OK572964
橄榄科	嘉榄属	白头树 <i>Garuga forrestii</i>	OM009254
		羽叶白头树 <i>Garuga pinnata</i>	ON361004
姜科	山姜属	高良姜 <i>Alpinia officinarum</i>	MT254526

其相关物种的叶绿体基因组序列,以单子叶植物高良姜的叶绿体基因组序列作为外群参照,采用 R 语言比较上述叶绿体基因组序列的蛋白质编码基因(PCG),分析它们的组成差异。随后选用所有序列共有的单拷贝 PCG,利用 MAFFT v7.429 软件进行比对,使用 Gblock 软件对模糊的区域进行修整,串联起来后构建系统进化树。系统进化树运用 IQtree v1.6.12 软件完成,所用方法为最大似然法,自展检验值为 1 000。

2 结果与分析

2.1 三桠苦叶绿体基因组的结构和基因组成

如图 1 所示,三桠苦叶绿体基因组为 158 992 bp 的环状双链四分体分子,由 85 101 bp 的大单拷贝区(LSC)、27 640 bp 的反向互补区 A(IRA)、18 611 bp 的小单拷贝区(SSC)和 27 640 bp 的反向互补区 B(IRB)依次排列而成,GC 平均含量为 38.3%。

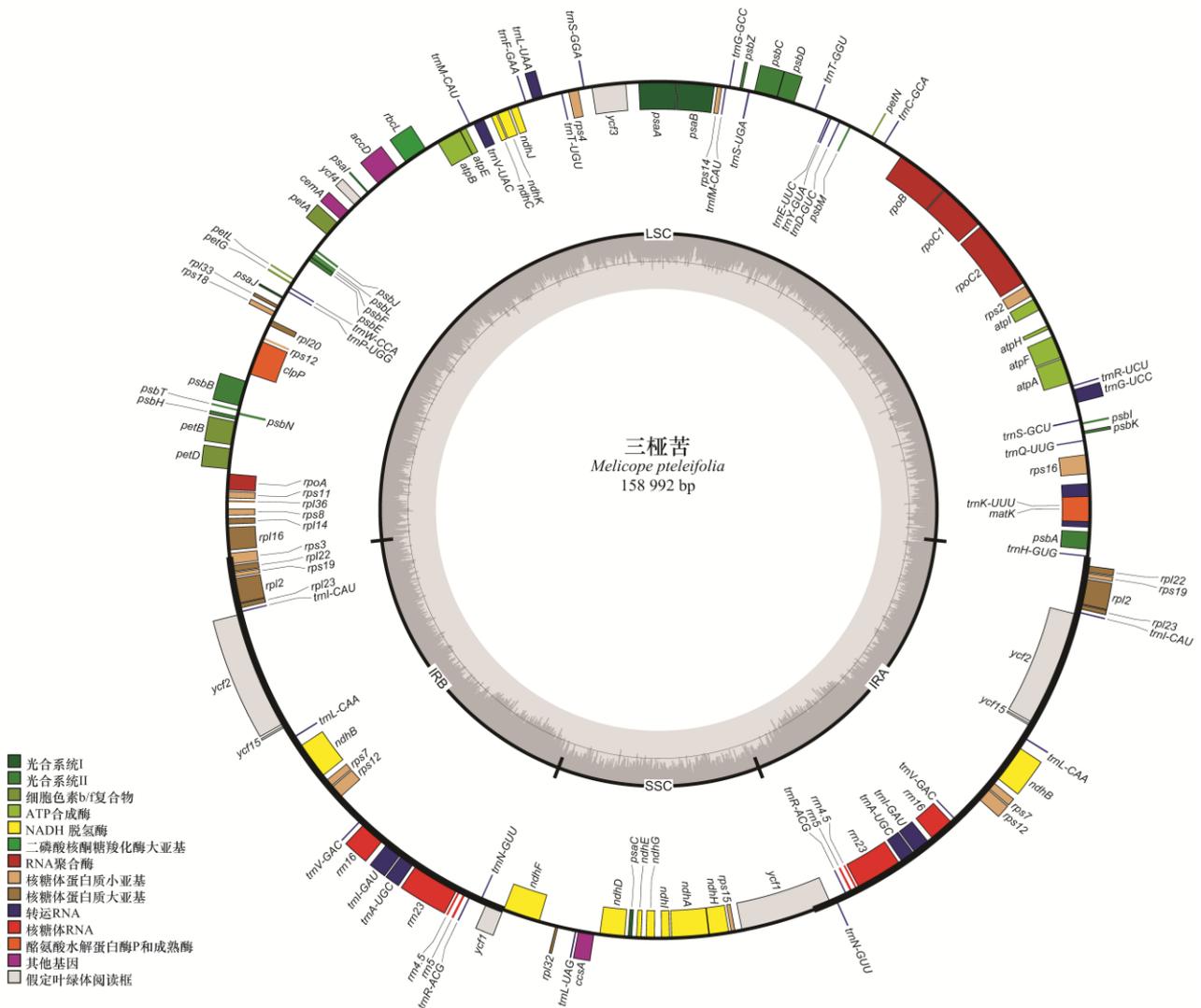


图 1 三桠苦叶绿体基因组图谱

Fig.1 Chloroplast genome map of *Melicope pteleifolia*

三桠苦叶绿体基因组共含有 134 个基因,包括 89 个蛋白质编码基因(PCG)、8 个核糖体 RNA(rRNA)以及 37 个转运 RNA(tRNA)。三桠苦叶绿体基因(表

2)主要是与植物光合作用以及叶绿体 DNA 自身复制相关的基因,还有少数基因如假定叶绿体阅读框未能获得功能注释。

表 2 三桠苦叶绿体基因组的基因信息

Table 2 Gene information in chloroplast genome of *Melicope pteleifolia*

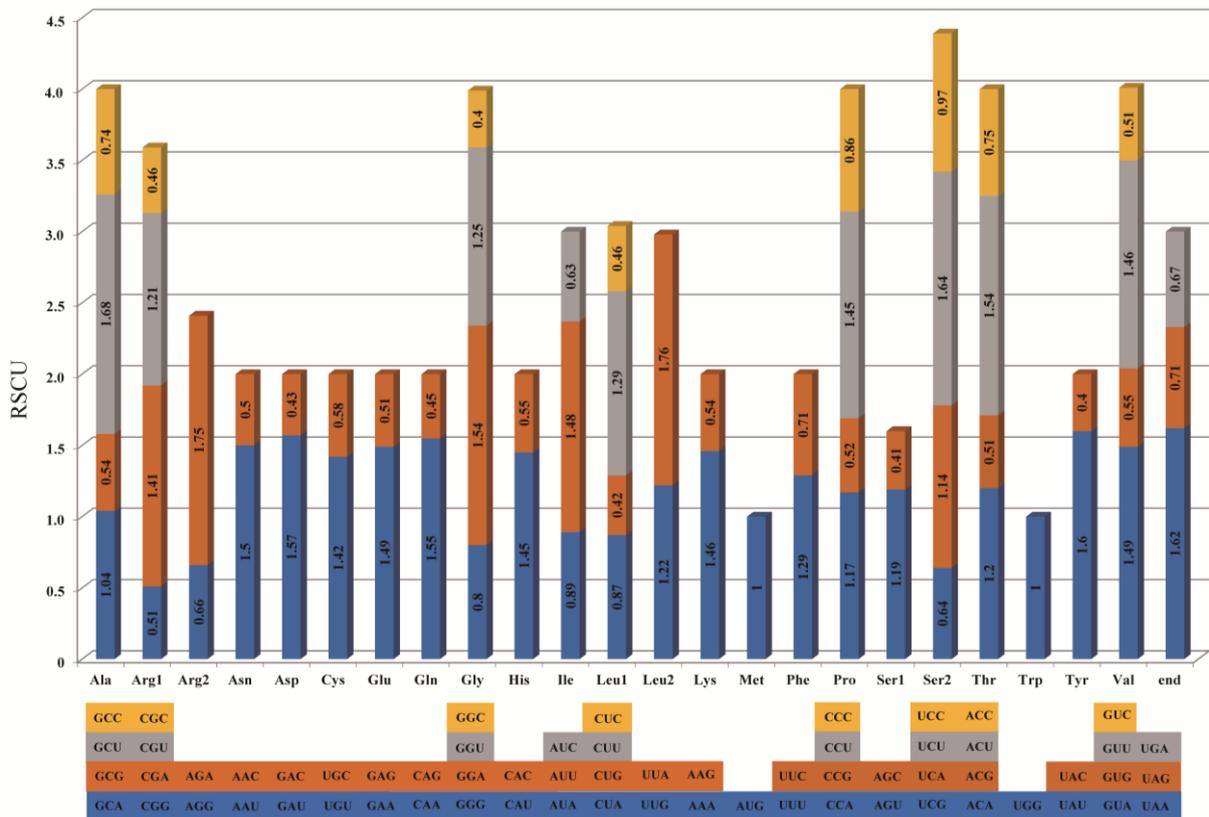
基因类别	基因功能	基因名称
光合作用	ATP 合成酶	<i>atpA, atpB, atpE, atpF¹, atpH, atpI</i>
	细胞色素 b/f 复合物	<i>petA, petB¹, petD¹, petG, petL, petN</i>
	NADH 脱氢酶	<i>ndhA¹, ndhB^{1*}, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>
	二磷酸核酮糖羧化酶大亚基	<i>rbcL</i>
	光合系统 I	<i>psaA, psaB, psaC, psaI, psaJ</i>
	光合系统 II	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ</i>
自我复制	核糖体蛋白质小亚基	<i>rps2, rps3, rps4, rps7*, rps8, rps11, rps12*, rps14, rps15, rps16¹, rps18, rps19*</i>
	核糖体蛋白质大亚基	<i>rpl2^{1*}, rpl14, rpl16¹, rpl20, rpl22*, rpl23*, rpl32, rpl33, rpl36</i>
	RNA 聚合酶	<i>rpoA, rpoB, rpoC1¹, rpoC2</i>
	核糖体 RNA	<i>rrn4.5*, rrn5*, rrn16*, rrn23*,</i>
	转运 RNA	<i>trnH-GUG, trnK-UUU¹, trnM-CAU, trnI-CAU*, trnV-UAC¹, trnF-GAA, trnL-UAA¹, trnT-UGU, trnS-GGA, trnJ-M-CAU, trnG-GCC, trnS-UGA, trnT-GGU, trnE-UUC, trnY-GUA, trnD-GUC, trnC-GCA, trnR-UCU, trnG-UCC¹, trnS-GCU, trnQ-UUG, trnW-CCA, trnP-UGG, trnL-CAA*, trnV-GAC*, trnI-GAU^{1*}, trnA-UGC^{1*}, trnR-ACG*, trnN-GUU*, trnL-UAG</i>
未知基因	假定叶绿体阅读框	<i>ycf1*, ycf2*, ycf3², ycf4, ycf15*</i>
其他基因		<i>matK, cemA, accD, ccsA, clpP²</i>

右上标 1、2 分别表示含有 1、2 个内含子的基因；#表示反式剪接基因；*表示双拷贝基因。

2.2 密码子偏好性

在三桠苦叶绿体基因组的 PCG 中共检测到 26 907 个密码子，包括 26 818 个氨基酸编码密码子和 89

个终止密码子。三桠苦叶绿体基因组密码子编码的氨基酸及其 RSCU 值见图 2，甲硫氨酸(Met)和色氨酸(Trp)只有 1 种密码子编码，其他氨基酸则由 2~



Ala 丙氨酸; Arg 精氨酸; Asn 天冬酰胺; Asp 天冬氨酸; Cys 半胱氨酸; Glu 谷氨酸; Gln 谷氨酰胺; Gly 甘氨酸; His 组氨酸; Ile 异亮氨酸; Leu 亮氨酸; Lys 赖氨酸; Met 甲硫氨酸; Phe 苯丙氨酸; Pro 脯氨酸; Ser 丝氨酸; Thr 苏氨酸; Trp 色氨酸; Tyr 酪氨酸; Val 缬氨酸; end 终止密码子。

图 2 三桠苦叶绿体基因组的相对同义密码子使用情况

Fig.2 Relative synonymous codon usage in chloroplast genome of *Melicope pteleifolia*

6种密码子编码,这些具有编码同一氨基酸或终止氨基酸翻译功能的不同密码子为同义密码子。在所有密码子中,有30种密码子的RSCU值大于1,说明这些密码子在编码对应氨基酸或终止氨基酸翻译时的使用频次高于其他同义密码子。29种密码子以A/T结尾,只有1种(编码亮氨酸的UUG)以C/G结尾,说明三桠苦叶绿体基因组倾向使用以A/T结尾的密码子。

2.3 SSR特点

依照预设的参数,三桠苦叶绿体基因组SSR的搜索结果如表3所示,三桠苦叶绿体基因组共有86个SSR,由单、二、三、四、五核苷酸组成,无六

表3 三桠苦叶绿体基因组SSR信息

Table 3 SSR loci in <i>Melicope pteleifolia</i> chloroplast genome			
类型	重复基序	数量	比例/%
单核苷酸	A/T	62	95.4
	C/G	3	4.6
二核苷酸	AT/AT	5	100.0
三核苷酸	AAT/ATT	7	70.0
	AAG/CTT	3	30.0
四核苷酸	AAAT/ATTT	3	60.0
	AACT/AGTT	1	20.0
	ACAT/ATGT	1	20.0
五核苷酸	AAATC/ATTTG	1	100.0
总计		86	

核苷酸。其中单核苷酸SSR数量最多,占全部SSR的75.6%(65/86);其次是三核苷酸SSR,占比11.6%(10/86);最少是五核苷酸SSR,仅有1个。在各类型SSR中,分别以A/T、AT/AT、AAT/ATT和AAAT/ATTT为主要重复基序,可见三桠苦叶绿体基因组SSR主要是由A或T碱基及其组合构成。

2.4 序列比较分析

在植物叶绿体基因组中,IR与相邻的SC形成边界,即IR边界。在LSC与IRB、IRB与SSC、SSC与IRA、IRA与LSC两两之间可依次形成4个IR边界,分别称JLB、LSB、JSA和JLA。三桠苦等5种芸香科植物叶绿体基因组的IR边界比较如图3所示。5种植物的JLB均落在基因内,其中三桠苦的JLB位于*rps3*基因中,而其他4种植物的JLB均处于*rpl22*基因里。5种植物的JSB存在较大的差异,三桠苦的JSB在*ycf1*基因内,花椒的JSB在*ycf1*和*ndhF*基因的重叠区,臭节草的JSB在*trnN*和*ndhF*基因的间隔区,而白鲜和臭檀吴萸的JSB则在*ndhF*基因内。5种植物的JSA都位于*ycf1*基因内,分别在距离该基因5'端的1069、1704、1093、1092、1406 bp处。三桠苦的JLA位于*rpl22*和*trnH*基因之间的基因间隔区,而其他4种植物的JLA则处在*rps19*和*trnH*基因之间的基因间隔区。

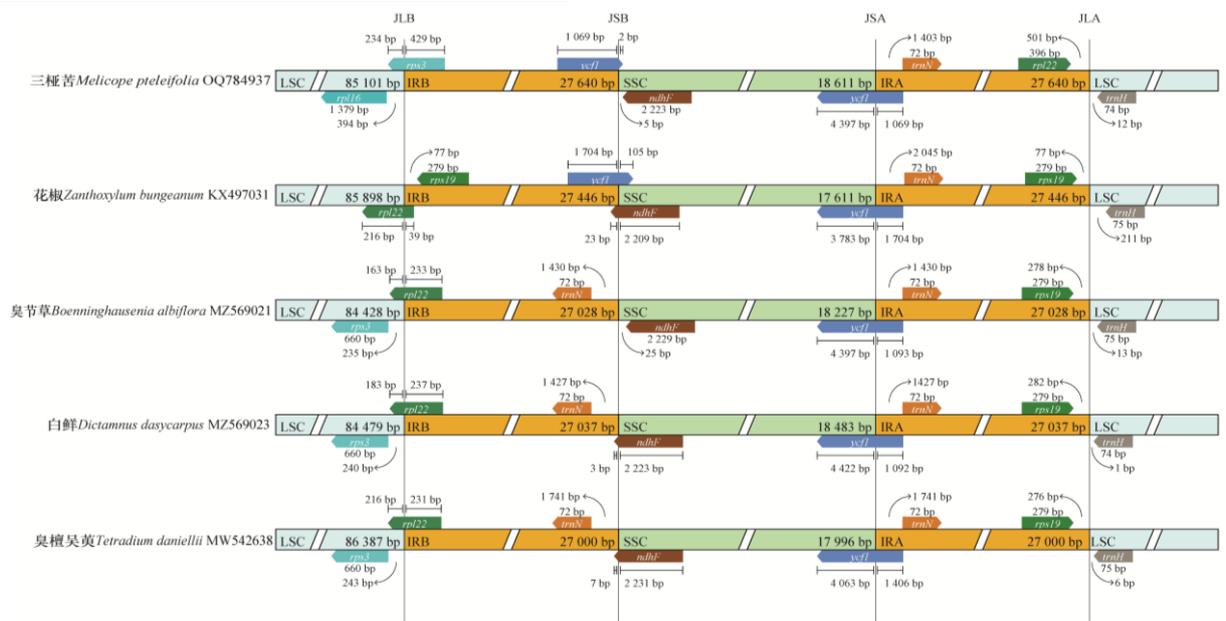
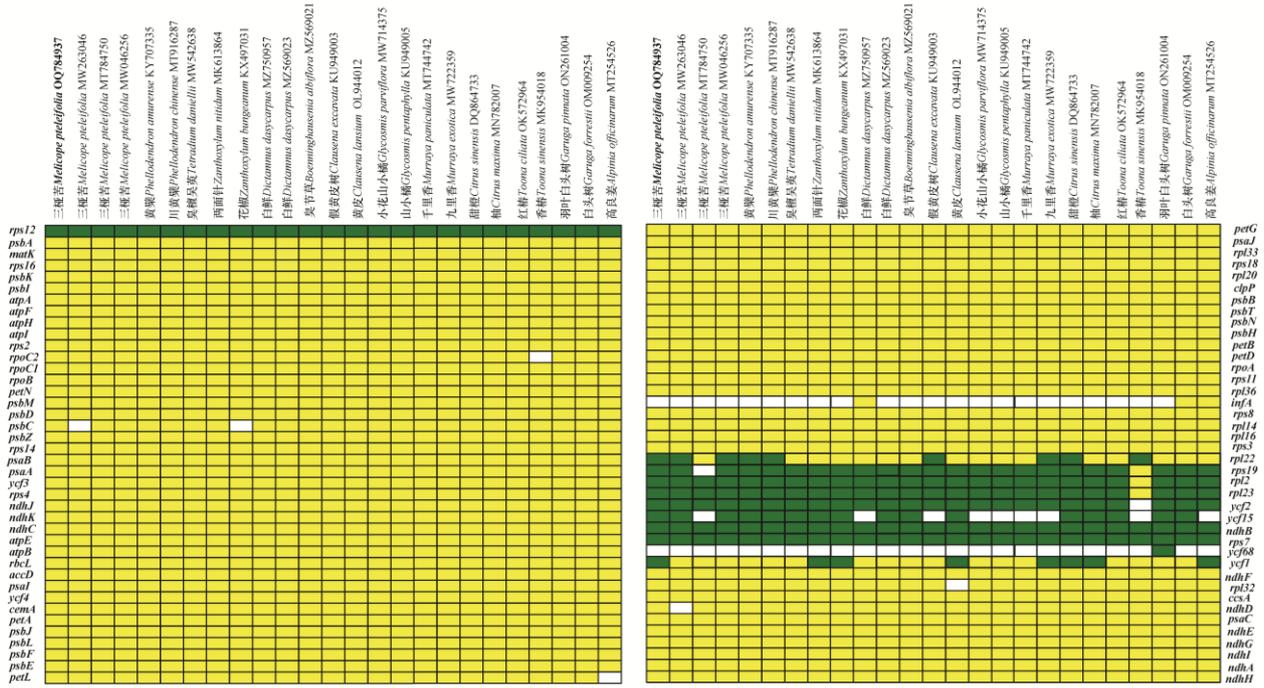


图3 三桠苦等5种植物的叶绿体基因组IR边界的比较

Fig.3 Comparison on IR boundaries in chloroplast genomes of five plants including *Melicope pteleifolia*



绿色表示该基因为双拷贝基因；黄色表示该基因为单拷贝基因；白色表示该基因缺失。

图5 三桠苦与花椒等叶绿体基因组 PCG 组成比较

Fig.5 Comparison on PCG composition of chloroplast genomes including *Melicope pteleifolia* and *Zanthoxylum bungeanum*

利用上述 25 条叶绿体基因组序列共有的 71 个单拷贝 PCG 构建聚类树，结果如图 6 所示。来自三桠苦的 4 条叶绿体基因组序列最先聚集在一起，形成蜜茱萸属分支，说明叶绿体基因组在三桠苦内存在较小的种内差异。随后蜜茱萸属再先后与吴茱萸

属、黄檗属、花椒属等共同聚成芸香科分支。楝科、橄榄科植物也分别聚成相应的分支。系统进化树聚类效果良好，直观地展现了三桠苦在芸香科内的进化位置。

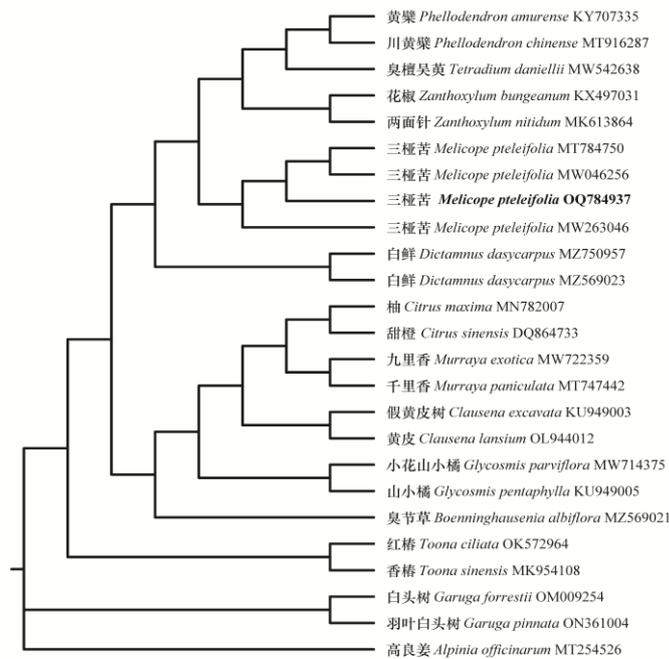


图6 基于叶绿体基因组建成的三桠苦系统进化树

Fig.6 Phylogenetic tree of *Melicope pteleifolia* constructed using chloroplast genome

3 结论与讨论

本研究结果表明,三桠苦叶绿体基因组具有植物典型的环状双链四分体结构,全长 158 992 bp,由 85 101 bp 的 LSC、18 611 bp 的 SSC 以及穿插在它们之间的 27 640 bp 的 IRA、IRB 构成,GC 含量为 38.3%,与前人^[12-14]报道的基本一致。本研究还发现三桠苦叶绿体基因组含有 89 个 PCG、8 个 rRNA 和 37 个 tRNA,共 134 个基因,而且主要是与光合作用和叶绿体自身复制相关的基因,与 LIU 等^[14]报道的三桠苦叶绿体基因组完全一致。上述结果表明,叶绿体基因组结构和基因组在不同来源的三桠苦样品中具有高度稳定性。

植物在进化过程中会倾向使用特定的一种或几种密码子来编码氨基酸^[15]。本研究发现,在三桠苦叶绿体基因组的 64 种密码子中,有 30 种密码子的 RSCU 值高于 1,且其中 29 种的末位碱基是 A/T,表明三桠苦在演化进程中具有较明显的密码子使用偏好性,倾向使用以 A/T 结尾的密码子。相似的情况也存在于高良姜^[11]、金丝李^[16]等多数植物中。

SSR 是植物基因组广泛存在的一种高频重复的 DNA 短序列,是研究品种鉴定、进化分析和亲缘关系的常用 DNA 手段^[17]。本研究从三桠苦的叶绿体基因组检测到 86 个 SSR,主要分布在 LSC 和基因间隔区,而且以 A/T 单核苷酸重复居多,这可能与三桠苦叶绿体基因组的低 GC 含量相关。

序列比较分析表明,三桠苦与同科植物在非编码区存在更多的碱基突变,符合植物叶绿体基因组非编码区具有更快进化速率的特点。系统进化分析发现,不同来源的 4 条三桠苦叶绿体基因组序列在 PCG 组成上存在基因缺失等变异,从而在系统进化树中出现先后聚集的现象,说明不同产地或居群的三桠苦在叶绿体基因组组成和碱基序列上可以呈现多态性,提示叶绿体基因组是探讨三桠苦种内多样性的有效途径。

本研究较系统地解析了三桠苦叶绿体基因组结构组成和序列特点,并基于叶绿体基因组序列探讨了三桠苦与相关植物的序列差异和亲缘关系,这些结果可为今后三桠苦的品种栽培、遗传进化、分子育种等资源保护和利用研究提供基因组材料。

参考文献:

[1] 中国植物志编辑委员会. 中国植物志: 第 41 卷[M]. 北京: 科学出版社, 1995.

- [2] 广东省食品药品监督管理局. 广东省中药材标准: 第一册[M]. 广州: 广东科技出版社, 2004.
- [3] 罗登花. 三叉苦种质资源 ISSR 遗传多样性分析及其品质评价[D]. 广州: 广州中医药大学, 2018.
- [4] DANIELL H, LIN C S, YU M, et al. Chloroplast genomes: diversity, evolution, and applications in genetic engineering[J]. *Genome Biology*, 2016, 17(1): 134.
- [5] 宋菊, 龙月红, 林丽梅, 等. 五加科植物叶绿体基因组结构与进化分析[J]. *中草药*, 2017, 48(24): 5070-5075.
- [6] 倪梁红, 赵志礼, 米玛. 药用植物叶绿体基因组研究进展[J]. *中药材*, 2015, 38(9): 1990-1994.
- [7] GUO S, LIAO X J, CHEN S Y, et al. A comparative analysis of the chloroplast genomes of four *Polygonum* medicinal plants[J]. *Frontiers in Genetics*, 2022, 13: 764534.
- [8] LI F, XIE X N, HUANG R, et al. Chloroplast genome sequencing based on genome skimming for identification of *Eriobotryae Folium*[J]. *BMC Biotechnology*, 2021, 21(1): 69.
- [9] LI W Q, LIU Y L, YANG Y, et al. Interspecific chloroplast genome sequence diversity and genomic resources in *Diospyros*[J]. *BMC Plant Biology*, 2018, 18(1): 210.
- [10] WANG Y H, WANG S, LIU Y L, et al. Chloroplast genome variation and phylogenetic relationships of *Atractylodes* species[J]. *BMC Genomics*, 2021, 22(1): 103.
- [11] 黄琼林. 高良姜叶绿体基因组测序与特征分析[J]. *热带作物学报*, 2021, 42(1): 1-6.
- [12] WU Z J, WU L W, XING J Y, et al. The complete chloroplast genome of *Melicope pteleifolia* (Champ. Ex Benth.) T. G. Hartley (Rutaceae)[J]. *Mitochondrial DNA Part B:Resources*, 2021, 6(4): 1339-1340.
- [13] YU J W, CHEN L X, MAO J W, et al. The complete chloroplast genome of *Melicope pteleifolia*(Rutaceae), a traditional medicinal plant in Southeast China[J]. *Mitochondrial DNA Part B:Resources*, 2021, 6(1): 60-61.
- [14] LIU H H, CHEN Z Y, ZHAO Y Y, et al. The third complete chloroplast genome of *Melicope pteleifolia* (Rutaceae): a widely used folk medicinal herb[J]. *Mitochondrial DNA Part B:Resources*, 2021, 6(3): 1204-1206.
- [15] MA L N, CUI P, ZHU J, et al. Translational selection in human: more pronounced in housekeeping genes[J]. *Biology Direct*, 2014, 9: 17.
- [16] 梁喜献, 庞洁, 谭长强. 金丝李叶绿体基因组序列分析[J]. *分子植物育种*, 2023, 21(10): 3214-3220.
- [17] 黄琼林. 高良姜转录组简单重复序列位点分析[J]. *时珍国医国药*, 2022, 33(2): 386-389.

责任编辑: 毛友纯
英文编辑: 柳正