

## ◆ 研究类 ◆

短叶茳茳 (*Cyperus malaccensis* subsp. *monophyllus*) 叶绿体全基因组特征及其系统发育分析\*陈 琨<sup>1,2</sup>, 潘良浩<sup>1,2</sup>, 廖 馨<sup>1,2\*\*</sup>

(1. 广西科学院, 广西海洋科学院(广西红树林研究中心), 广西红树林保护与利用重点实验室, 广西北海 536000; 2. 自然资源部北部湾滨海湿地生态系统野外科学观测研究站, 广西北海 536015)

**摘要:**短叶茳茳(*Cyperus malaccensis* subsp. *monophyllus*)为我国东南至南方沿海分布面积较大的原生滨海盐沼草本植物,在保滩护岸和改良土壤等领域发挥重要作用。本研究利用 Illumina 高通量测序技术获取短叶茳茳叶绿体全基因组信息特征,并探究其基因组特征及系统发育关系。结果显示,短叶茳茳叶绿体全基因组全长 186 314 bp,为典型的四分体结构;其中大单拷贝(LSC)区和小单拷贝(SSC)区分别为 101 162 和 10 310 bp,反向重复(IR)区为 37 421 bp,共注释叶绿体基因 141 个,包括 94 个蛋白编码基因、39 个转运 RNA 基因和 8 个核糖体 RNA 基因。共鉴定出 105 个简单重复序列(Simple Sequence Repeat, SSR)位点和 77 个长重复序列,挖掘出 183 个单核苷酸多态性(Single Nucleotide Polymorphism, SNP)位点。系统发育分析表明,在莎草属(*Cyperus*)中,短叶茳茳、茳茳(*C. malaccensis*)与香附子(*C. rotundus*)亲缘关系最近。本研究为进一步探究莎草科(Cyperaceae)植物的系统发育关系及近缘物种的区分提供了参考依据。

**关键词:**短叶茳茳;叶绿体基因组;高通量测序;特征分析;系统发育分析

中图分类号:Q943 文献标识码:A 文章编号:1002-7378(2024)03-0355-09

DOI:10.13657/j.cnki.gxkxyxb.20241108.013

短叶茳茳(*Cyperus malaccensis* subsp. *monophyllus*)属莎草科(Cyperaceae)莎草属(*Cyperus*),曾命名为 *C. malaccensis* var. *brevifolius*<sup>[1]</sup>,是我国东南至南方沿海分布面积较大的原生滨海盐沼草本植物,但其耐盐能力较差,常见于有淡水注入的潮间带淤泥质滩涂,常与红树林混生(图 1),因其植株密

集,根系稠密,在保滩护岸、促淤造陆、固碳、阻留悬浮物和改良土壤等方面具有较高的生态价值。与其原种茳茳(*C. malaccensis*)的主要形态区别在于茳茳的地上茎顶部的 3 片苞片极展开,长于花序,短叶茳茳则相反<sup>[1]</sup>。短叶茳茳主要分布在我国华南沿海和日本,我国集中分布区主要有福建闽江口<sup>[2]</sup>、福建泉州

收稿日期:2024-06-30

修回日期:2024-08-19

\* 广西科技人才专项(桂科 AD20159032),广西红树林保护与利用重点实验室基金项目(GKLMC-22A07)和广西红树林研究中心基本科研业务费项目(2023GMRC-02)资助。

## 【第一作者简介】

陈 琨(1995—),男,研究实习员,主要从事海洋分子生物学研究,E-mail:chenkun@mangrove.org.cn。

## 【\*\*通信作者简介】

廖 馨(1982—),女,副研究员,主要从事海洋生物学研究,E-mail:njuliaoxin@163.com。

## 【引用本文】

陈琨,潘良浩,廖馨.短叶茳茳(*Cyperus malaccensis* subsp. *monophyllus*)叶绿体全基因组特征及其系统发育分析[J].广西科学院学报,2024,40(3):355-363.

CHEN K, PAN L H, LIAO X. Complete Chloroplast Genome Characteristics and Phylogenetic Analysis of *Cyperus malaccensis* subsp. *monophyllus* [J]. Journal of Guangxi Academy of Sciences, 2024, 40(3): 355-363.

湾、广东珠江河口、广西廉州湾、广西钦州茅尾海滩涂等<sup>[3]</sup>。



图1 短叶茳芏的生境、花序及苞片

Fig. 1 Habitat, inflorescence and bract of *C. malaccensis* subsp. *monophyllus*

滨海盐沼、红树林和海草生态系统可高效固碳,是“蓝色碳汇”的重要组成部分,具有减缓气候变化的潜在作用。有研究指出,“盐沼草(以茳芏和短叶茳芏为主要种类)——红树林协同生态修复系统”可有效突破红树造林技术瓶颈<sup>[4,5]</sup>。国内学者开展了较全面的关于短叶茳芏的碳、氮循环过程,温室气体收支及重金属富集规律等方面的研究,也有学者对茳芏和短叶茳芏的群落分布动态变化开展了研究<sup>[6-8]</sup>。短叶茳芏及其原种茳芏,以及莎草属的其他物种在形态学上极为相近,较难区分,但尚未有学者利用分子生物学技术开展其物种鉴定、系统发生关系和遗传多样性的研究。叶绿体基因组具有重组率较低、后代遗传稳定等特点,相比于核基因组,是揭示物种系统发生关系的有效手段<sup>[9,10]</sup>。莎草科是一个庞大的单子叶植物类群,全球有5 600种,莎草属有约950种<sup>[11]</sup>,但在Genbank中仅上传了不到20个莎草属物种的叶绿体全基因组序列(截至2024年8月),其中不乏形态分类有争议的物种。为了厘清莎草属、莎草科的系统发生关系,亟需补充该属物种的叶绿体基因组信息。本研究以短叶茳芏的叶片为研究材料,采用Illumina高通量测序技术,对短叶茳芏叶绿体全基因组进行测序和组装,并利用生物信息技术对其结构特征进行分析,构建莎草科的系统发育进化树,为莎草属植物的进化和系统发育研究,以及后续开展物种鉴定和遗传多样性的研究提供分子基础信息。

## 1 材料与方法

### 1.1 材料

样品采自广西钦州市钦南区茅尾海

(108°34′43.80″E, 21°52′10.83″N),选取健康植株的新鲜嫩叶用于DNA提取,将带有花果的植株制成标本用于后期的形态鉴定。标本存放于广西海洋科学院(广西红树林研究中心)的标本室。

### 1.2 基因组DNA的提取及测序

从样品的新鲜健康叶片中提取总基因组DNA,送至上海凌恩生物科技有限公司,经检测合格后,采用Illumina NovaSeq 6000双末端测序策略进行建库测序。

### 1.3 基因组组装和注释

使用Trimmomatic v 0.38<sup>[12]</sup>对原始数据进行质量剪切,过滤质量比较低的数据。基因组序列通过NOVOPlasty v 4.2软件<sup>[13]</sup>进行组装。使用GeSeq软件<sup>[14]</sup>预测叶绿体基因组中的蛋白质编码(PCG)、转达RNA(tRNA)和核糖体RNA(rRNA)基因。最后,使用OGDRAW v 1.3.1软件<sup>[15]</sup>绘制短叶茳芏叶绿体基因组图谱。短叶茳芏基因组序列已上传GenBank(登录号:OR438648)。

### 1.4 SSR分析

通过MISA在线软件<sup>[16]</sup>进行短叶茳芏叶绿体基因组的简单重复序列(Simple Sequence Repeat, SSR)分析,参数设置<sup>[17]</sup>如下:(1)1-10,2-6,3-5,4-5,5-5和6-5,即1个碱基重复 $\geq 10$ 次;2个碱基重复 $\geq 6$ 次;3/4/5/6个碱基均重复 $\geq 5$ 次。(2)两个SSR的最小距离设置为100 bp,如果距离小于100 bp,两个SSR序列则组成一个复合微卫星。

### 1.5 重复序列分析

短叶茳芏叶绿体基因组的重复序列采用REPuter在线软件<sup>[18]</sup>(<http://bibiserv.techfak.uni-bielefeld.de/reputer>)进行分析,包括正向重复(Forward repeats)、反向重复(Reverse repeats)、互补重复(Complement repeats)和回文重复(Palindromic repeats)。参数设置:汉明距离(Hamming distance)设置为0,即重复序列完全一致,长度大于30 bp。

### 1.6 SNP检测及注释

利用MUMmer软件<sup>[19]</sup>(v3.23)分析潜在单核苷酸多态性(Single Nucleotide Polymorphism, SNP)位点;以茳芏的叶绿体基因组作为参考,将参考序列SNP位点两侧各100 bp的序列提取出来,然后利用组装软件BLAT v35在组装结果中进行比对,验证SNP位点。若比对长度小于101 bp,则认为该SNP为不可信SNP,被删除;若重复被比对上多次,则被认为是重复区域的SNP,也将被去除,最终获得可靠

的 SNP。

### 1.7 叶绿体基因组比较及系统发育分析

从 Genbank 下载莎草属 5 个物种菹荳 (*C. malaccensis*, OR238391)、香附子 (*C. rotundus*, MT473237)、高秆莎草 (*C. exaltatus*, MW123055)、水蜈蚣 (*C. aromaticus*, MT937178)、碎米莎草 (*C. iria*, MW123056) 的叶绿体全基因组序列并进行基因组序列和结构比较。从 GenBank 下载 12 个莎草科植物的叶绿体全基因组序列, 包括菹荳、香附子、高秆莎草、水蜈蚣、碎米莎草、细荳 (*Isolepis setacea*, MW041568)、*Eleocharis cellulosa*, MN985041、荸荠 (*E. dulcis*, MN967018)、翼果薹草 (*Carex neurocarpa*, KU238086)、团穗薹草 (*Carex agglomerata*, MT795185)、宽叶薹草 (*Carex siderosticta*, KP751906) 和割鸡荳 (*Hypolytrum nemorum*, KU207098), 并以禾本科植物高粱 (*Sorghum bicolor*, NC008602) 作为外群, 基于叶绿体 SNP 矩阵构建系统发育树, 并按照 1.6 节描述的方法获得 SNP 位点<sup>[20]</sup>。利用 PhyML v 3.0 软件 (<http://www.atgc-montpellier.fr/phyml/>) 构建系统进化树, 采用最大

似然法, 分支的可信度基于 Bootstrap 重复抽样 1 000 次, 其他参数选择默认设置。

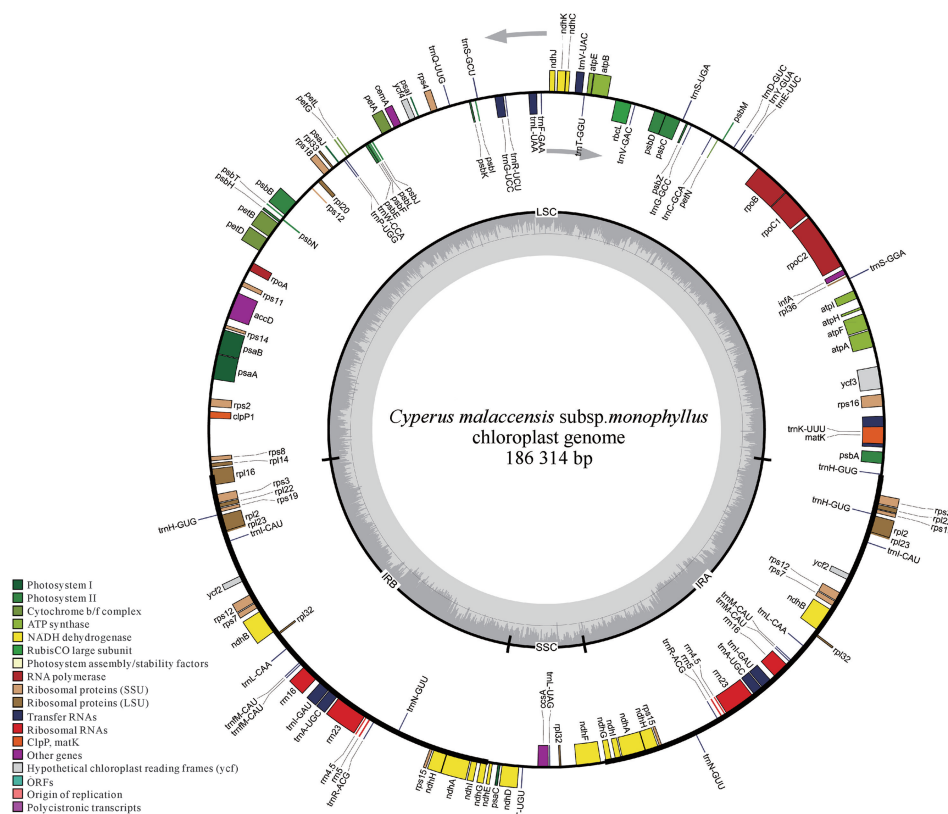
## 2 结果与分析

### 2.1 测序数据

利用高通量测序平台 Illumina NovaSeq 6000 测得 Raw data 3.44 G, GC 含量为 37.73%, Q20 为 94.01%, Q30 为 86.75%, 去除低质量 reads, 得到 Clean data 3.12 G, 有效数据的平均 GC 含量为 37.58%, 有效数据的 Q20 和 Q30 分别为 95.51% 和 89.00%, 测序数据质量合格。

### 2.2 短叶菹荳叶绿体基因大小

短叶菹荳叶绿体全基因组呈环状双链, 序列全长为 186 314 bp, GC 含量为 33.17%, 包含大单拷贝 (LSC) 区、小单拷贝 (SSC) 区及两个反向重复序列 (IRA, IRB) 区, 其中 LSC 起止位置为 1-101 162, 长度为 101 162 bp; SSC 起止位置为 138 584-148 893, 长度为 10 310 bp; IRA 和 IRB 长度均为 37 421 bp, 起止位置分别为 148 894-186 314 和 101 163-138 583 (图 2)。



Genes drawn on the outside of the circle are transcribed in a counterclockwise direction, while genes drawn on the inside of the circle are transcribed in a clockwise direction.

图 2 短叶菹荳叶绿体基因组图谱

Fig. 2 Genome map of *Cyperus malaccensis* subsp. *monophyllus* chloroplast

### 2.3 短叶茛苳叶绿体基因注释、基因功能及分类

短叶茛苳叶绿体基因组共注释了 141 个基因, 主要分为 3 大类, 即 94 个蛋白质编码 (PCG) 基因, 39 个转运 RNA (tRNA) 基因, 8 个核糖体 RNA (rRNA) 基因。27 个基因在 IR 区重复, 包括 15 个蛋白质编码基因 (*rps3*, *rpl22*, *rps19*, *rpl2*, *rpl23*, *ycf2*, *rps12*, *rps7*, *ndhB*, *rpl16*, *rps15*, *ndhH*, *ndhA*, *ndhI*, *ndhG*), 8 个 tRNA 基因 (*trnN*-GUU, *trnR*-ACG, *trnA*-UGC, *trnI*-GAU, *trnH*-GUG, *trnFM*-CAU, *trnL*-CAA, *trnI*-CAU), 4 个 rRNA 基因

(*rrn4.5*, *rrn5*, *rrn16*, *rrn23*)。此外, 根据功能进行分类, 又可分为光合作用相关基因、自我复制相关基因、其他蛋白编码基因和未知功能基因。其中, 与光合作用相关基因有 49 个, 可进一步细分为 6 个组, 包括光合系统 I、光合系统 II、NADH 脱氢酶、细胞色素复合物、ATP 合酶和二磷酸核酮糖羧化酶大亚基; 自我复制相关基因有 69 个, 可进一步细分为 5 个组, 分别是核糖体大亚基蛋白、核糖体小亚基蛋白、RNA 聚合酶亚基、rRNA 基因和 tRNA 基因; 其他蛋白编码基因有 6 个; 未知功能基因有 4 个 (表 1)。

表 1 短叶茛苳叶绿体基因组基因分类

Table 1 List of genes of *C. malaccensis* subsp. *monophyllus* chloroplast genome

基因分类 Category of genes	基因分组 Group of genes	基因名称 Name of genes
Photosynthesis genes	Photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i>
	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>
	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>
	Cytochrome b/f complex	<i>petA</i> , <i>petB</i> , <i>petD</i> , <i>petG</i> , <i>petL</i> , <i>petN</i>
	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>
	Self-replication genes	Large subunits of ribosome
Small subunits of ribosome		<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>
RNA polymerase		<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> , <i>rpoC2</i>
rRNA		<i>rrn4.5</i> , <i>rrn5</i> , <i>rrn16</i> , <i>rrn23</i>
tRNA		<i>trnA</i> -UGC, <i>trnC</i> -GCA, <i>trnD</i> -GUC, <i>trnE</i> -UUC, <i>trnF</i> -GAA, <i>trnG</i> -GCC, <i>trnG</i> -UCC, <i>trnH</i> -GUG, <i>trnI</i> -CAU, <i>trnI</i> -GAU, <i>trnK</i> -UUU, <i>trnL</i> -CAA, <i>trnL</i> -UAA, <i>trnL</i> -UAG, <i>trnM</i> -CAU, <i>trnN</i> -GUU, <i>trnP</i> -UGG, <i>trnQ</i> -UUG, <i>trnR</i> -ACG, <i>trnR</i> -UCU, <i>trnS</i> -GCU, <i>trnS</i> -GGA, <i>trnS</i> -UGA, <i>trnT</i> -GGU, <i>trnT</i> -UGU, <i>trnV</i> -GAC, <i>trnV</i> -UAC, <i>trnW</i> -CCA, <i>trnY</i> -GUA, <i>trnFM</i> -CAU
Other genes		Maturase
	Protease	<i>clpP1</i>
	Envelope membraneprotein	<i>cemA</i>
	Acetyl-CoA carboxylase	<i>accD</i>
	c-type cytochrome synthesis gene	<i>ccsA</i>
	Translation initiation factor	<i>infA</i>
	Genes of unknown function	ORF of unknown function

### 2.4 叶绿体基因组 SSRs 分析

通过对短叶茛苳叶绿体全基因组序列进行 MI-SA 分析, 共检测出 105 个符合条件的 SSR 位点, 均

由 A 或 T 组成, 表明 SSRs 的碱基组成偏向使用 A/T 碱基。其中包括 91 个单核苷酸重复基序、13 个二核苷酸重复基序、1 个三核苷酸重复基序, 未发现四



核苷酸/五核苷酸/六核苷酸重复基序(表2)。在所有SSRs中,最多的重复基序是A/T,然后依次是AT/AT和ATA,这些重复基序从碱基数量来看,具有A/T碱基偏好性,这与短叶荳荳叶绿体全基因组AT含量占比高一致(62.42%)。从分布区段上看,SSRs的主要分布在LSC区,有73个SSR位点;其次是IR区(IRA和IRB各有13个);SSC区最少,仅有6个SSR位于该区段。

表2 短叶荳荳叶绿体基因组中SSR位点类型及数量

Table 2 Type and number of SSR loci of *C. malaccensis* subsp. *monophyllus* chloroplast genome

SSR 重复基序 SSR repeat motifs	长度/bp Length/bp	SSR 数量/个 Number of SSRs	合计占比/% Percentage of total/%
Mononucleotide			
A/T	10	39	86.67
	11	21	
	12	15	
	13	3	
	14	5	
	15	1	
	16	3	
	17	2	
	19	1	
	21	1	
	Dinucleotide		
AT/AT	6	5	12.38
	7	4	
	8	3	
	17	1	
Trinucleotide			
AAT/ATT	5	1	0.95

## 2.5 长重复序列分析

利用REPuter在线软件对短叶荳荳的叶绿体基因组进行重复序列分析,共鉴定出77个长重复序列(图3),包括41个正向重复序列、4个反向重复序列、30个回文重复序列和2个互补重复序列。短叶荳荳叶绿体基因组重复序列长度主要为30-60 bp(图3),占全部重复序列的78%。这些重复序列可为开发种群进化标记提供基础。

## 2.6 SNP分析

以荳荳的叶绿体全基因组作为参照,对短叶荳荳

的叶绿体全基因组序列进行SNP分析,共鉴定到183个SNP位点,包括位于基因区的SNP位点37个和位于基因间区的SNP位点146个。其中,基因区包括17个同义突变、20个非同义突变(表3)。

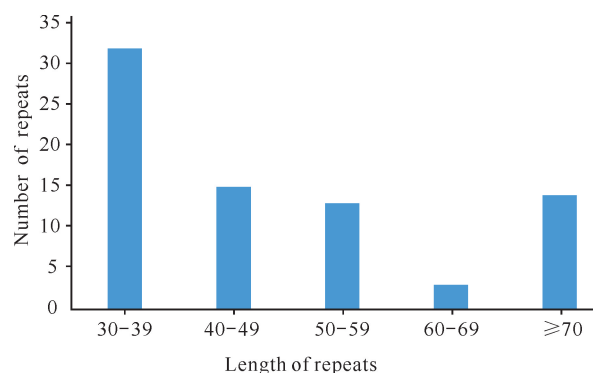


图3 长重复序列长度范围分布

Fig. 3 Distribution of length ranges of long repeats sequences

表3 短叶荳荳 SNP 分析

Table 3 SNP analysis of *C. malaccensis* subsp. *monophyllus*

项目 Items	SNP 数量 SNP number	占比/% Percentage/%	
CDS region	Synonymous	17	9.3
	Nonsynonymous	20	10.9
Intergenic region	146	79.8	
Total SNP	183	100.0	

## 2.7 叶绿体基因组比较分析

选取莎草属6个物种的叶绿体基因组进行比较分析<sup>[21,22]</sup>,发现短叶荳荳与其原种荳荳的叶绿体基因组结构基本相同。莎草属6个物种叶绿体基因组长度为182 986-187 401 bp,其中水蜈蚣叶绿体基因组长度最长,香附子最短;6个物种的基因组GC含量相差不大,为33.16%-33.26%;短叶荳荳和荳荳具有相同的基因组分数量和数量,而香附子的基因总数均少于另外5个同属植物(表4)。短叶荳荳和荳荳的基因总数和编码蛋白质基因的数量均高于其他4种,说明其叶绿体在生长过程中能行使更多的功能。

## 2.8 系统发育分析

为进一步研究短叶荳荳的遗传背景和分类学关系,将短叶荳荳叶绿体基因组序列与其他12个莎草科植物的叶绿体基因组序列进行多重比对后,以高粱作为外群,基于叶绿体SNP矩阵构建系统发育树(图4)。结果显示,莎草属(*Cyperus*)、荳荳属(*Eleocharis*)和藁草属(*Carex*)的单系性得到强烈支持(支持率均为100%),细荳属(*Isolepis*)与莎草属形成姐妹

群,并得到 100% 支持,共同与荸荠属构成姐妹群,支持率为 100%。莎草属中,短叶荳苳与其原种荳苳、香附子聚为一支,亲缘关系较近,支持率为 99%,从表 4 莎草属 6 个物种叶绿体基因组结构比较

Table 4 Comparison of chloroplast genome structure of 6 species of *Cyperus*

物种名称 Name of species	Genbank ID	基因组长度/bp Genome length/bp	GC 含量/% GC content/%	LSC 长度/bp LSC length/bp	SSC 长度/bp SSC length/bp	IR 长度/bp IR length/bp	基因总数 Total number of genes	蛋白质基因数量 Number of protein genes	rRNA 基因数量 Number of rRNA genes	tRNA 基因数量 Number of tRNA genes
<i>C. malaccensis</i> subsp. <i>monophyllus</i>	OR438648	186 314	33.17	101 162	10 310	37 421	141	94	8	39
<i>C. malaccensis</i>	OR238391	186 098	33.18	100 934	10 296	37 434	141	94	8	39
<i>C. iria</i>	MW123056	185 697	33.16	99 360	10 267	38 035	135	89	8	38
<i>C. exaltatus</i>	MW123055	187 139	33.18	100 780	10 271	38 044	135	90	8	41
<i>C. rotundus</i>	MT937176	182 986	33.26	100 733	10 315	35 969	133	84	8	41
<i>C. aromaticus</i>	MT937178	187 401	33.13	101 085	10 496	37 910	136	87	8	41

形态学上来看,这 3 种也更为相似。碎米莎草、高秆莎草和水蜈蚣聚为另一支,但水蜈蚣聚为一支的支持率仅为 83%。

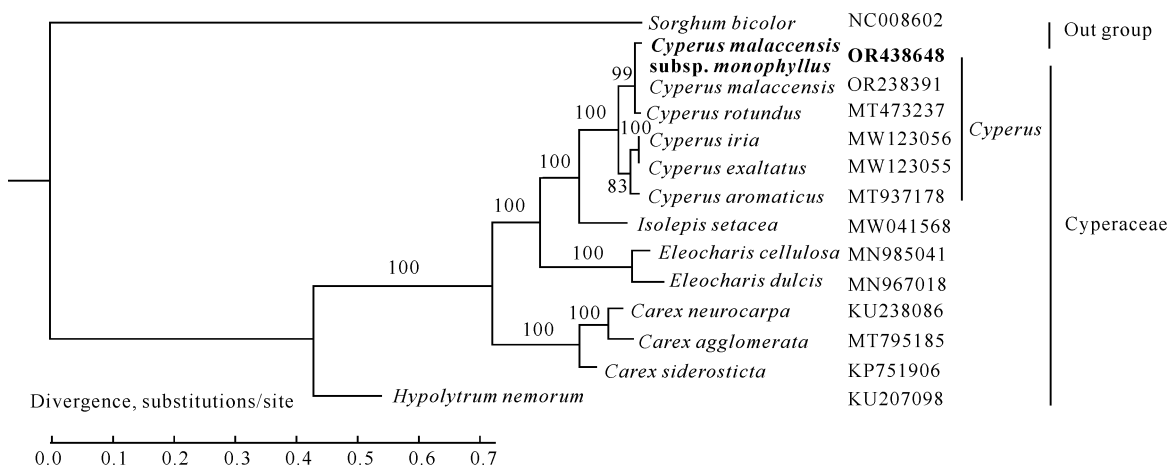


图 4 基于叶绿体全基因组序列最大似然法构建的 14 个物种的系统发育树

Fig. 4 Phylogenetic tree of 14 species constructed based on the maximum likelihood method of chloroplast genome sequence

### 3 讨论

短叶荳苳为荳苳的亚种,是海湾河口滩涂绿化的优良植物。二者是我国南方沿海分布面积最大的 2 种原生滨海盐沼草本植物,也是典型的盐沼植物,在保滩护岸、促淤造陆、阻留悬浮物和改良土壤等方面具有较高的生态价值。近年来,研究者们对短叶荳苳的研究主要集中在环境、生态、湿地、土壤等多个学科<sup>[6,23-25]</sup>。

叶绿体基因组序列是植物鉴定和建立物种间遗传关系的一种通用工具。目前,短叶荳苳的叶绿体基因组尚未有研究者报道,同属的也仅有荳苳<sup>[26]</sup>、碎米莎草<sup>[22]</sup>、油莎草(*C. esculentus*)<sup>[27]</sup>等被报道。本研究基于 Illumina 高通量测序技术获得莎草科莎草属

植物短叶荳苳的叶绿体基因组全序列,并首次通过生物信息学手段对其进行分析研究。结果显示,短叶荳苳叶绿体基因组全长为 186 314 bp,为典型的四分体结构,通过比较短叶荳苳与其他 5 个同属植物叶绿体基因组序列,发现上述植物的叶绿体基因组大小相近(182 986—187 401 bp),说明属间基因组大小具有一定的稳定性<sup>[28-30]</sup>。短叶荳苳叶绿体基因组 GC 含量为 33.17%,共注释了 141 个基因,所编码的基因类别、数量及排列顺序与另外 5 种同属植物差异不大,同时具有高度相似的 GC 含量(33.16%—33.26%)。此外,短叶荳苳的基因组总长度、LSC 和 SSC 长度均大于荳苳,但短叶荳苳的 IR 区要短于荳苳,说明短叶荳苳的 IR 区存在收缩现象。

叶绿体基因组中常用来鉴定物种分类的分子条

形码有 matK、rbcL 和 trnH-psbA, 但并不能区分短叶茛苳及其原种茛苳(相似性分别为 99.94%、99.93%、99.91%)。植物通用的分子条形码 SSR 是一种高效的分子标记, 具有数量多、重现性高、显性遗传、单亲遗传和相对保守等优点, 是在种群和个体水平上进行物种鉴定和遗传变异评估的最佳选择。Ren 等<sup>[27]</sup>对油莎草进行长重复序列分析, 获得 49 个长重复序列, 包括 15 个正向重复序列和 34 个回文重复序列。本研究的短叶茛苳叶绿体基因组中共检测到 77 个长重复序列(长度主要为 30—60 bp), 其中 41 个正向重复序列、4 个反向重复序列、30 个回文重复序列和 2 个互补重复序列。短叶茛苳叶绿体基因组的 SSR 位点共有 105 个, 出现频率最高的为单核苷酸重复基序, 占 SSR 总数的 86.67%, SSR 位点数量与其他同属植物差别较大, 远少于油莎草<sup>[27]</sup>(396 个), 推测可能是不同物种在鉴定 SSR 位点时软件设置参数不完全一致导致的, 如油莎草 SSR 鉴定设置的单核苷酸到六核苷酸的重复次数依次为 8、5、4、3、3、3, 而本研究设置的参数是 10、6、5、5、5、5。本研究填补了莎草属植物短叶茛苳 SSR 位点研究的空白, 为短叶茛苳分子标记的开发和莎草属物种鉴定提供了分子基础信息。

SNP 是指由单个核苷酸的变异所引起的 DNA 序列多态性。在基因组 DNA 中, 任何碱基都有可能发生突变, 因此 SNP 既可能位于编码基因上, 也有可能位于非编码序列上, 在编码区内的 SNP (coding SNP, cSNP) 因其可能影响个体的功能而备受关注。本研究通过对短叶茛苳的叶绿体基因组进行 SNP 挖掘分析, 共鉴定到 183 个 SNP 位点, 包括位于基因区的 SNP 位点 37 个和位于基因间区的 SNP 位点 146 个。其中, 基因区包括 17 个同义突变、20 个非同义突变。这些 SNP 位点可为短叶茛苳的物种鉴定、遗传多样性分析和物种溯源等研究提供技术基础资料。

为进一步研究短叶茛苳的遗传背景和分类学关系, 从 NCBI 数据库公开的植物中选取了包括短叶茛苳在内的 13 个莎草科植物进行聚类分析。结果显示短叶茛苳与香附子、茛苳聚为一支, 分支支持率达到 99%, 是亲缘关系最近的 3 个物种。水蜈蚣与另两种莎草属物种碎米莎草、高秆莎草聚为一支的支持率仅 83%。经进一步核实, 目前水蜈蚣在《中国植物志》中的拉丁名已修订为 *Kyllinga polyphylla*, 属莎草科水蜈蚣属。Fan 等<sup>[21]</sup>在 2024 年 5 月发表的论文中对短叶水蜈蚣(*Cyperus brevifolius*)叶绿体基因组进

行系统发育分析, 显示 *C. brevifolius* 属于莎草属, 但支持率为 74%。说明莎草属物种依据形态特征进行的分类学分析和基于基因序列的系统发育关系分析存在一定的矛盾, 甚至已发表的论文和已上传 Genbank 的数据也存在着诸多混淆之处。因此, 对形态相似的物种进行基于叶绿体基因组的测序, 开发筛选出有效的分子条形码, 对系统进化以及物种鉴定等领域的研究具有理论和实践意义。

本研究对短叶茛苳的叶绿体基因组进行了组装和注释, 分析了基因组特征并构建了系统发育树, 挖掘了一批 SSR 和 SNP 标记, 为莎草属植物的物种鉴定及亲缘关系分析提供了进一步的分子基础资料, 为莎草属植物物种后续的系统发育研究及植物资源开发提供理论基础。本研究仅采集了钦州茅尾海一地的短叶茛苳, 后续研究也将从沿海各地采集短叶茛苳及茛苳的样品, 对挖掘的 SSR 和 SNP 进行设计筛选和验证, 开发有效的分子标记。

#### 参考文献

- [1] 中国科学院中国植物志委员会. 中国植物志: 第十一卷 莎草科[M]. 北京: 科学出版社, 1961.
- [2] ZENG A, HU W, ZENG C, et al. Litter decomposition and nutrient dynamics of native species (*Cyperus malaccensis*) and alien invasive species (*Spartina alterniflora*) in a typical subtropical estuary (Min River) in China [J]. *Estuaries and Coasts*, 2020, 43 (7): 1873-1883.
- [3] 潘良浩, 史小芳, 曾聪, 等. 广西滨海盐沼生态系统研究现状及展望[J]. *广西科学*, 2017, 24(5): 453-461.
- [4] 何斌源, 赖廷和, 潘良浩, 等. 盐沼草: 白骨壤混种减轻污损动物危害的生物防治效果研究[J]. *广西植物*, 2014, 34(2): 203-211.
- [5] 何斌源, 赖廷和, 王欣, 等. 盐沼草对桐花树人工林污损动物危害的生物防治研究[J]. *广西科学*, 2013, 20(3): 185-192.
- [6] 林晓雪, 黄佳芳, 李慧, 等. 闽江河口芦苇沼泽和短叶茛苳沼泽生态系统含碳温室气体的年收支[J]. *生态学报*, 2022, 42(22): 9186-9198.
- [7] 牟晓杰, 刘兴土, 仝川, 等. 闽江河口短叶茛苳湿地 CH<sub>4</sub> 和 N<sub>2</sub>O 排放对氮输入的短期响应[J]. *环境科学*, 2012, 33(7): 2482-2489.
- [8] 宋萍, 封磊, 洪伟, 等. 闽江河口湿地土壤植物系统 Zn、Cd 的分布特征[J]. *福建林学院学报*, 2009, 29(1): 17-22.
- [9] 陆奇丰, 黄至欢, 骆文华. 极小种群濒危植物广西火桐、

- 丹霞梧桐的叶绿体基因组特征[J]. 生物多样性, 2021, 29(5):586-595.
- [10] BIRKY C W JR. Uniparental inheritance of mitochondrial and chloroplast genes; mechanisms and evolution [J]. Proceedings of the National Academy of Sciences, 1995, 92(25):11331-11338.
- [11] TAHERI Y, HERRERA-BRAVO J, HUALA L, et al. *Cyperus* spp. : a review on phytochemical composition, biological activity, and health-promoting effects [J]. Oxidative Medicine and Cellular Longevity, 2021, 2021(1):4014867.
- [12] BOLGER A M, LOHSE M, USADEL B. Trimmomatic: a flexible trimmer for illumina sequence data [J]. Bioinformatics, 2014, 30(15):2114-2120.
- [13] NICOLAS D, PATRICK M, GUILLAUME S. NOVO-Plasty: de novo assembly of organelle genomes from whole genome data [J]. Nucleic Acids Research, 2017, 45(4):e18.
- [14] TILLICH M, LEHWARK P, PELLIZZER T, et al. GeSeq- versatile and accurate annotation of organelle genomes [J]. Nucleic Acids Research, 2017, 45(W1):W6-W11.
- [15] GREINER S, LEHWARK P, BOCK R. Organellar Genome DRAW (OGDRAW) version 1. 3. 1: expanded toolkit for the graphical visualization of organellar genomes [J]. Nucleic Acids Research, 2019, 47(W1):W59-W64.
- [16] BEIER S, THIEL T, MÜNCH T, et al. MISA-web: a web server for microsatellite prediction [J]. Bioinformatics, 2017, 33(16):2583-2585.
- [17] 谢海坤, 焦健, 樊秀彩, 等. 基于高通量测序组装‘赤霞珠’叶绿体基因组及其特征分析[J]. 中国农业科学, 2017, 50(9):1655-1665.
- [18] KURTZ S, CHOUDHURI J V, OHLEBUSCH E, et al. REPuter: the manifold applications of repeat analysis on a genomic scale [J]. Nucleic Acids Research, 2001, 29(22):4633-4642.
- [19] KURTZ S, PHILLIPPY A, DELCHER A L, et al. Velvet and open software for comparing large genomes [J]. Genome Biology, 2004, 5(2):1-9.
- [20] LI D M, ZHAO C Y, LIU X F. Complete chloroplast genome sequences of *Kaempferia galanga* and *Kaempferia elegans*; molecular structures and comparative analysis [J]. Molecules, 2019, 24(3):474.
- [21] FAN X, YUAN K, ZHENG Y. The complete chloroplast genome and phylogenetic analysis of *Cyperus brevifolius* (Rottb.) Hassk. 1844 (Cyperaceae) [J]. Mitochondrial DNA Part B, 2024, 9(5):621-624.
- [22] YANG L L, NIU J Q, TANG W W. The complete chloroplast genome of pioneering plant *Cyperus iria* L. (Cyperaceae) in ecological restoration [J]. Mitochondrial DNA Part B, 2021, 6(4):1335-1336.
- [23] 姚钦予, 孙志高, 胡星云, 等. 氮负荷增强对闽江口短叶荳苳湿地植物-土壤系统氮累积与分配的影响[J]. 生态学报, 2024, 44(11):4831-4843.
- [24] 张鹏飞, 孙志高, 陈冰冰, 等. 闽江河口芦苇与短叶荳苳空间扩展对湿地系统磷赋存特征的影响[J]. 应用生态学报, 2020, 31(12):4258-4266.
- [25] 王华, 孙志高, 李家兵, 等. 闽江口芦苇与短叶荳苳湿地土壤无机硫形态分布特征及其影响因素[J]. 生态学报, 2019, 39(13):4921-4932.
- [26] PAN L, CHEN K, LIAO X. The complete chloroplast genome and phylogenetic analysis of *Cyperus malaccensis* Lam (Cyperaceae) [J]. Mitochondrial DNA Part B, 2024, 9(1):114-118.
- [27] REN W, GUO D, XING G, et al. Complete chloroplast genome sequence and comparative and phylogenetic analyses of the cultivated *Cyperus esculentus* [J]. Diversity, 2021, 13(9):405.
- [28] 段春燕, 张凯, 段义忠. 长柄扁桃叶绿体基因组比较[J]. 生物工程学报, 2020, 36(12):4921-4932.
- [29] 阳秀玫, 陈亮明, 禹霖, 等. 华中樱叶绿体基因组特征与分子标记[J]. 中南林业科技大学学报, 2024, 44(1):175-184.
- [30] 杨芳. 七里香蔷薇叶绿体基因组测序及结构分析[J]. 基因组学与应用生物学, 2019, 38(8):3586-3594.



# Complete Chloroplast Genome Characteristics and Phylogenetic Analysis of *Cyperus malaccensis* subsp. *monophyllus*

CHEN Kun<sup>1,2</sup>, PAN Lianghao<sup>1,2</sup>, LIAO Xin<sup>1,2\*\*</sup>

(1. Guangxi Key Laboratory of Mangrove Conservation and Utilization, Guangxi Academy of Marine Science (Guangxi Mangrove Research Center), Guangxi Academy of Sciences, Beihai, Guangxi, 536000, China; 2. Observation and Research Station of Coastal Wetland Ecosystem in Beibu Gulf, Ministry of Natural Resources, Beihai, Guangxi, 536015, China)

**Abstract:** *Cyperus malaccensis* subsp. *monophyllus* is a native coastal salt marsh herb with a large distribution area from the southeast to the south coast of China, which plays an important role in the fields of beach protection and soil improvement. In this study, Illumina high-throughput sequencing technology was used to obtain the whole genome information characteristics of the chloroplast of *C. malaccensis* subsp. *monophyllus*, and to explore its genomic characteristics and phylogenetic relationship. The results showed that the complete chloroplast genome of *C. malaccensis* subsp. *monophyllus* was 186 314 bp in length, which was a typical tetrad structure. The large single-copy region and the small single-copy region were 101 162 and 10 310 bp, respectively, and the inverted repeat region was 37 421 bp. A total of 141 chloroplast genes were annotated, including 94 protein-coding genes, 39 transport RNA genes and 8 ribosomal RNA genes. A total of 105 Simple Sequence Repeat (SSR) sites and 77 long repeat sequences were identified, and 183 Single Nucleotide Polymorphism (SNP) sites were mined. Phylogenetic analyses showed that *C. malaccensis* subsp. *monophyllus*, *C. malaccensis* had the closest relationship with *C. rotundus* in *Cyperus*. This study provides a reference for further exploring the phylogenetic relationship of Cyperaceae plants and the differentiation of related species.

**Key words:** *Cyperus malaccensis* subsp. *monophyllus*; chloroplast genome; high-throughput sequencing; characterization; phylogenetic analysis

责任编辑: 陆 雁



微信公众号投稿更便捷

联系电话: 0771-2503923

邮箱: gxxkxyxb@gxas.cn

投稿系统网址: <http://gxxk.ijournal.cn/gxxkxyxb/ch>