

# 新型水果空心泡叶绿体基因组特征及其系统发育分析

吴永飞<sup>1</sup>, 杨雪莲<sup>1\*</sup>, 王霞<sup>1</sup>, 颜丽<sup>1</sup>, 张万萍<sup>1,2</sup>

1 贵州大学农学院, 贵州 贵阳 550025

2 贵州大学 贵州省蔬菜研究院, 贵州 贵阳 550025

吴永飞, 杨雪莲, 王霞, 颜丽, 张万萍. 新型水果空心泡叶绿体基因组特征及其系统发育分析[J]. 生物工程学报, 2023, 39(7): 2939-2953.

WU Yongfei, YANG Xuelian, WANG Xia, YAN Li, ZHANG Wanping. Characteristics and phylogenetic analysis of chloroplast genome of a new type of fruit *Rubus rosaefolius*[J]. Chinese Journal of Biotechnology, 2023, 39(7): 2939-2953.

**摘要:** 为探究空心泡(*Rubus rosaefolius*)叶绿体基因组特征, 本研究以空心泡为试验材料, 采用 Illumina NovaSeq 平台进行高通量测序, 获得空心泡完整的叶绿体基因组序列, 并进行空心泡叶绿体基因组序列特征和系统发育分析。结果表明: 空心泡的完整叶绿体基因组总长度为 155 650 bp, 具有典型的四分体结构, 包括 2 个反向重复序列(各 25 748 bp)、1 个大拷贝区(85 443 bp)、1 个小拷贝区(18 711 bp)。空心泡叶绿体全基因组共鉴定出 131 个基因, 包括 86 个蛋白质编码基因、37 个 tRNA 基因和 8 个 rRNA 基因, 全基因组的 GC 含量为 36.9%。空心泡叶绿体基因组包含 47 个散在重复序列、72 个简单重复序列(simple sequence repeating, SSR)位点, 密码子偏好性为亮氨酸密码子, 偏好使用 A/U 结尾的密码子。系统发育分析表明, 空心泡与小叶悬钩子(*Rubus taiwanicola*)亲缘关系最近, 其次是能高悬钩子(*Rubus rubroangustifolius*)和腺萼悬钩子(*Rubus glandulosopunctatus*)。空心泡的叶绿体基因组特征及其系统发育分析, 为空心泡的遗传多样性研究和叶绿体开发利用提供理论依据。

**关键词:** 空心泡; 叶绿体基因组; 重复序列; 密码子偏好性; 系统发育

资助项目: 国家自然科学基金(31860225); 贵州省科技计划项目(黔科合平台人才[2018]5781 号, 黔科合基础[2019]1408 号)  
This work was supported by the National Natural Science Foundation of China (31860225) and by the Guizhou Provincial Science and Technology Plan Project (Qian Kehe Platform Personnel [2018] No. 5781, Qian Kehe Foundation [2019] No. 1408).

\*Corresponding author. E-mail: yxl4684221@126.com

Received: 2022-10-16; Accepted: 2023-01-07; Published online: 2023-01-10

# Characteristics and phylogenetic analysis of chloroplast genome of a new type of fruit *Rubus rosaefolius*

WU Yongfei<sup>1</sup>, YANG Xuelian<sup>1\*</sup>, WANG Xia<sup>1</sup>, YAN Li<sup>1</sup>, ZHANG Wanping<sup>1,2</sup>

<sup>1</sup> College of Agriculture, Guizhou University, Guiyang 550025, Guizhou, China

<sup>2</sup> Vegetable Research Academy, Guizhou University, Guiyang 550025, Guizhou, China

**Abstract:** The genomic DNA of *Rubus rosaefolius* was extracted and sequenced by Illumina NovaSeq platform to obtain the complete chloroplast genome sequence, and the sequence characteristics and phylogenetic analysis of chloroplast genes were carried out. The results showed that the complete chloroplast genome of the *R. rosaefolius* was 155 650 bp in length and had a typical tetrad structure, including two reverse repeats (25 748 bp each), a large copy region (85 443 bp) and a small copy region (18 711 bp). A total of 131 genes were identified in the whole genome of *R. rosaefolius* chloroplast, including 86 protein coding genes, 37 tRNA genes and 8 rRNA genes. The GC content of the whole genome was 36.9%. The genome of *R. rosaefolius* chloroplast contains 47 scattered repeats and 72 simple sequence repeating (SSR) loci. The codon preference is leucine codon, and the codon at the end of A/U is preferred. Phylogenetic analysis showed that *R. rosaefolius* had the closest relationship with *R. taiwanicola*, followed by *R. rubraangustifolius* and *R. glandulosopunctatus*. The chloroplast genome characteristics and phylogenetic analysis of *R. rosaefolius* provide a theoretical basis for its genetic diversity research and chloroplast development and utilization.

**Keywords:** *Rubus rosaefolius*; chloroplast genome; repetitive sequence; codon bias; phylogenetic analysis

叶绿体是绿色植物特有的细胞器, 作为光合作用的主要场所, 为植株生长发育提供能量, 在绿色植物的几个重要生化过程和光合作用中发挥着不可或缺的作用<sup>[1]</sup>。高等植物叶绿体基因组的结构非常保守, 大多数基因组为双链环状 DNA 分子, 其结构由 4 个部分组成, 分别是大单拷贝区 (large single-copy region, LSC)、小单拷贝区 (small single-copy region, SSC) 和一对反向重复区 (inverted repeat, IRs)<sup>[2]</sup>。叶绿体基因组是从母本遗传下来的, 在基因含量和基因组结构上高度保守, 大多数叶绿体基因组的长度在 120–160 kb 之间, 含有 110–130 个单拷贝的

基因<sup>[3]</sup>。叶绿体基因组的研究早期主要用于研究珍稀濒危植物的进化历史和保护, 目前已广泛应用于各物种遗传多样性、系统进化等的研究<sup>[4]</sup>。随着高通量测序技术的不断完善, 自首次获得烟草<sup>[5]</sup>叶绿体基因组序列信息, 目前已有大花君子兰<sup>[6]</sup>、苦瓜<sup>[7]</sup>、山核桃<sup>[8]</sup>、安息香猕猴桃<sup>[9]</sup>等多种植物的叶绿体基因组完成测序, 解析了叶绿体基因组特征和重复序列等, 确定了系统位置及其与近缘物种的亲缘关系。

空心泡 (*Rubus rosaefolius* Sm) 别名蔷薇莓, 属于蔷薇科悬钩子属, 在全球广泛分布。该物种原产于所罗门群岛、新喀里多尼亚、瓦努阿

图、毛里求斯、中南半岛和中国<sup>[10]</sup>。空心泡开白花和结红色果实,果实 4-5 月成熟,营养丰富,口感香甜,入药可治肺热咳嗽、盗汗、跌打损伤等<sup>[11]</sup>,在传统医学中用于治疗腹泻和胃部疾病<sup>[12]</sup>,具有镇痛、抗菌、降压等药理特性<sup>[10]</sup>。研究表明,空心泡还具有降血压<sup>[13]</sup>、抗氧化<sup>[14]</sup>、抗癌<sup>[15]</sup>等功能。此外,作为新型水果,空心泡富含红色素,该色素属水溶性花青苷类色素,有较好的光和热稳定性,适用 pH 值范围宽,适于用作天然食用着色剂,且可通过调节 pH 获得不同色泽<sup>[16]</sup>。因此,空心泡的遗传和基因组信息对于系统研究和野生资源保护具有重要意义。

空心泡作为第 3 代新型水果,是一种优良的野生果树资源,具有巨大潜在开发价值。然而目前关于空心泡的研究主要集中在营养成分的提取<sup>[11]</sup>和药理特性<sup>[10]</sup>,而关于空心泡叶绿体基因组的系统研究尚鲜见报道。本研究以空心泡为材料,通过高通量测序、组装和基因注释,获得空心泡叶绿体全基因组,并完成叶绿体基因组系统发育分析,以此提高对悬钩子属叶绿体基因组结构的了解,为今后悬钩子属植物的遗传多样性和系统发育分析提供参考。

## 1 材料与方法

### 1.1 试验材料

2021 年 10 月空心泡植物样本采样于贵州省植物园(26°37'20"N, 106°43'29"E)。取新鲜叶片用液氮处理,于贵州大学农学院园艺本科教学实验室-80 °C 超低温冰箱中保存待用。

### 1.2 方法

#### 1.2.1 空心泡 DNA 提取和测序

采用十六烷基三甲基溴化铵(cetyltrimethylammonium bromide, CTAB)法<sup>[17]</sup>

从 300 mg 新鲜叶片中提取基因组 DNA,提取的基因组 DNA 经检测合格后,用超声波法将 DNA 片段化;对片段化的 DNA 进行片段纯化、末端修复、3'端加 A、连接测序接头,再用琼脂糖凝胶电泳进行片段大小选择,进行聚合酶链式反应(polymerase chain reaction, PCR)扩增形成测序文库,并对文库进行质检,质检合格的文库用 Illumina NovaSeq 进行测序(深圳市惠通生物科技有限公司)。原始数据经过去接头污染,去除低质量序列后得到 3.82 G 的干净数据(clean reads)。

#### 1.2.2 空心泡叶绿体基因组的组装和注释

原始测序数据使用 NGS QC Tool Kit v2.3.3<sup>[18]</sup>进行质控,使用 de novo 组装软件 SPAdes v3.11.0,从 3.82 G 的高质量数据中组装出完整的叶绿体基因组<sup>[19]</sup>。使用 PGA 软件对完整的叶绿体基因组进行注释<sup>[20]</sup>。将最终得到的基因组注释信息以及基因组序列提交至 GenBank(登录号为 OL435124.1)。最后使用 OGDRAW 软件<sup>[21]</sup>制作空心泡叶绿体基因组完整图谱。

#### 1.2.3 空心泡叶绿体密码子偏好性和重复序列分析

运用 CodonW 软件<sup>[22]</sup>统计分析空心泡叶绿体基因组密码子的偏好性(relative synonymous codon usage, RSCU)。利用 Reputer 软件<sup>[23]</sup>预测空心泡草叶绿体基因组散在重复序列。即正向重复(forward, F)、反向重复(reverse, R)、互补重复(complement, C)、回文重复(palindromic, P),最小重复长度(minimal repeat size)设置为 30,最小排列值(minimum alignment score)为 50,最多碱基错配(hamming distance)为 3。利用 MISA 软件<sup>[24]</sup>检测空心泡叶绿体基因组的简单重复序列(simple sequence repeat, SSR)位点,单核苷酸(mononucleotide)、二核苷酸(dinucleotide)、三核

昔酸(trinucleotide)、四核苷酸(tetranucleotide)、五核苷酸(pentanucleotide)和六核苷酸(hexanucleotide)的最小重复值分别设置为 10、5、4、3、3 和 3。

#### 1.2.4 IR 区边界收缩和扩张分析

运用 IRscope 软件<sup>[25]</sup>对空心泡和同属的腺萼悬钩子(*Rubus glandulosopunctatus*, NC\_057624.1)、能高悬钩子(*Rubus rubroangustifolius*, NC\_057629.1)、小叶悬钩子(*Rubus taiwanicola*, NC\_056940.1)、光滑悬钩子(*Rubus tsangii*, NC\_056940.1) 4 个物种进行叶绿体基因组边界收缩与扩张比较分析, 绘制对比图。

#### 1.2.5 空心泡叶绿体基因组系统发育树的构建

从 NCBI 中下载蔷薇科 34 个物种的叶绿体基因组序列, 以刺蔷薇 [*Rosa acicularis* (MK714016.1)] 作为外群进行系统进化分析。利用 MAFFT 7.037 软件<sup>[26]</sup>对 35 种蔷薇科植物进行多重比对。随后, 使用 model-finder var 1.6 来选择 TVM+F+I+G4 模型<sup>[27]</sup>。最后, 利用 RAxML var 8.2.9 软件<sup>[28]</sup>构建系统发育树。自举检测值为 1 000。

## 2 结果与分析

### 2.1 空心泡叶绿体全基因组基本特征分析

空心泡叶绿体全基因组(GenBank 登录号: OL435124)由 155 650 bp 核苷酸组成(图 1), 具有典型的四分体结构特征, 即 1 个大单拷贝区(LSC)、1 个小单拷贝区(SSC)和 1 对反向重复序列(IRs)。

LSC 区长度为 85 443 bp, SSC 长度为 18 711 bp, 两个反向重复序列长度为 25 748 bp, 全基因组的 GC 含量为 36.9%。空心泡叶绿体基因组注释结果显示, 空心泡叶绿体基因组编码 131 个基因, 包括 86 个蛋白质编码基因、37 个

tRNA 基因和 8 个 rRNA 基因(表 1)。其中 5 个蛋白编码基因(protein coding genes, PCGs) (*rpl2*、*rpl23*、*ycf2*、*ndhB* 和 *rps7*), 4 个 rRNA 基因(*rrn16*、*rrn23*、*rrn4.5* 和 *rrn5*), 7 个 tRNA 基因 (*trnA-UGC*、*trnI-CAU*、*trnI-GAU*、*trnL-CAA*、*trnN-GUU*、*trnR-ACG* 和 *trnV-GAC*) 在 IR 区域复制, *ycf1* 基因位于两个 IR 区和 SSC 区交界处。

编码基因根据其产物功能的不同可分为光合作用相关基因(44)、自我复制相关基因(59)、其他(6)、未知功能基因(4)。其中, 大多数基因不具有内含子, 只有 *petB*、*ndhA*、*rpl2*、*rps16* 等 14 个基因含有 1 个内含子, *clpP* 和 *ycf3* 含有 2 个内含子。此外, 基因 *rps12* 存在反式剪接(表 2)。

### 2.2 空心泡叶绿体基因组重复序列分析

空心泡叶绿体基因组中共鉴定出 47 个散在重复序列, 其中有 25 个正向重复(53.19%), 3 个反向重复(6.38%), 2 个互补重复(4.26%)和 17 个回文重复(36.17%)。除了一个回文重复序列较长外(25 748 bp), 其余重复序列均在 30–56 bp。各重复序列在基因编码区(*ndhA*、*clpP*、*rpl16* 等)、基因间隔区(*trnS-GCU-trnG-UCC*、*trnV-GAC-rps7*、*rrn4.5-rrn5* 等)、tRNA (*trnG-UCC*、*trnS-GGA*、*trnG-GCC*)和 *ycf3*、*clpP* 基因内含子区域均有分布(表 3)。

### 2.3 空心泡叶绿体基因组 SSR 分析

空心泡叶绿体基因组鉴定出 6 种不同序列的 72 个 SSR 位点, 包含单核苷酸、二核苷酸和三核苷酸, 没有出现四核苷酸、五核苷酸和六核苷酸(表 4)。其中, 单核苷酸重复数量最多, 共 58 个, 占 80.56%, 主要以 A/T 重复序列为主, 共 55 个, C/G 重复类型仅为 3 个; 二核苷酸重复数量为 11 个, 包括 AG/CT (1SSR)和 AT/AT (10SSR)两种重复序列, 占

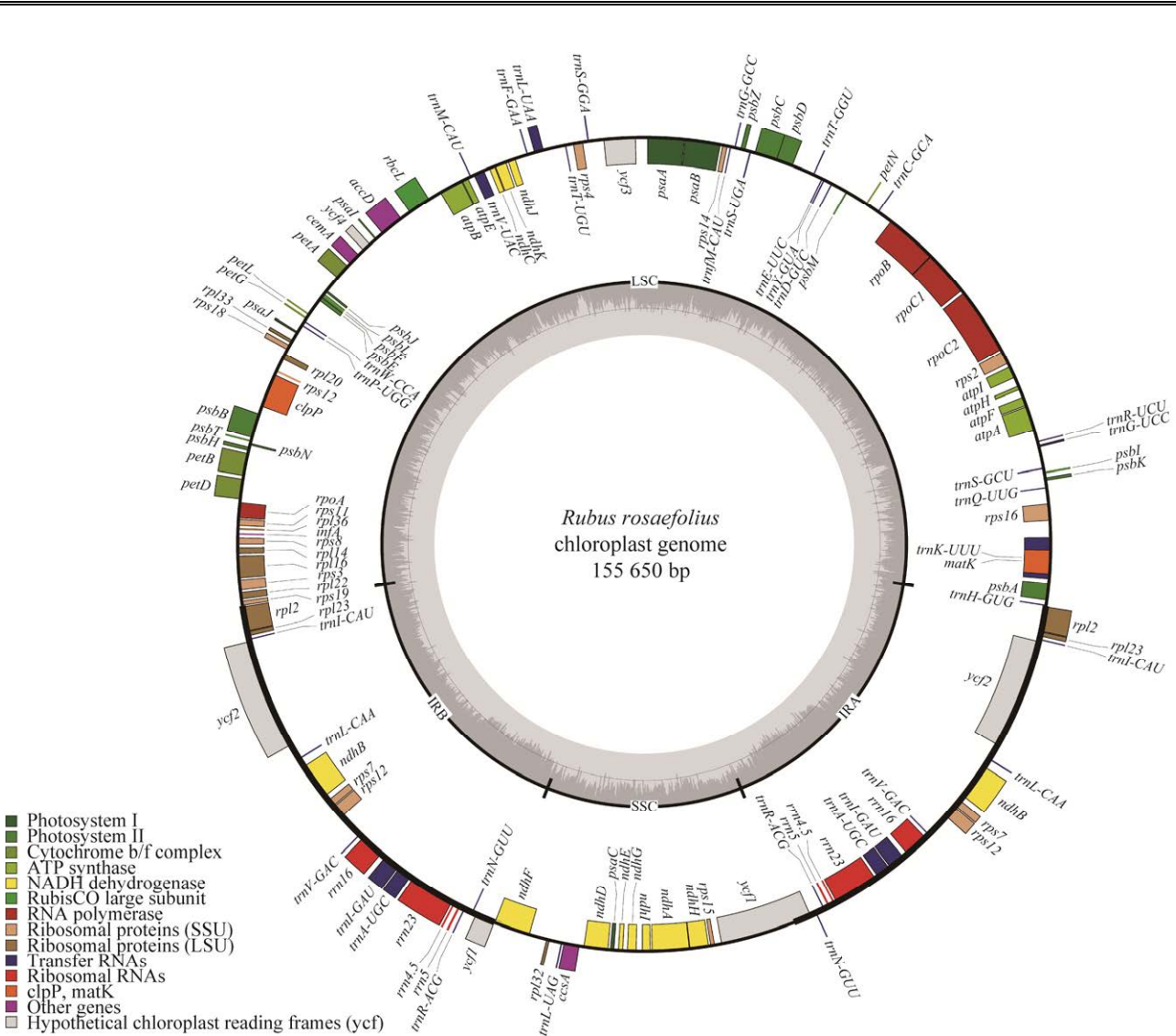


图 1 空心泡叶绿体基因组图谱  
Figure 1 Complete chloroplast genome map of *Rubus rosaefolius*.

表 1 叶绿体基因组基因基本特征  
Table 1 Basic characteristics of chloroplast genome genes

|                     | <i>Rubus rosaefolius</i> |
|---------------------|--------------------------|
| Total length (bp)   | 155 650                  |
| LSC (bp)            | 85 443                   |
| SSC (bp)            | 18 711                   |
| IR (bp)             | 25 748                   |
| Total gene          | 131                      |
| Protein-coding gene | 86                       |
| tRNA gene           | 37                       |
| rRNA gene           | 8                        |
| GC (%)              | 36.9                     |

15.28%；三核苷酸重复数量仅为 3 个，具有 AAG/CTT (1SSR)和 AAT/ATT (2SSR)两种重复方式，占 4.17%。

2.4 空心泡密码子偏好性分析

密码子使用情况显示，空心泡叶绿体基因组除终止密码子外，共有 48 957 个密码子，共有 22 种氨基酸被这些密码子编码，各种氨基酸的使用频率在 1.30%–10.10%之间(表 5)。空心泡叶绿体基因组密码子数量最多的氨基酸是亮氨酸(Leu)，具有 5 239 个密码子，占 10.10%，

其次是异亮氨酸(Ile)和丝氨酸(Ser), 分别具有 4 501 个和 3 732 个密码子, 占总密码子数的 8.68%和 7.19%, 数量最少的氨基酸是色氨酸(Trp), 具有 677 个密码子, 占 1.30%。同义密

码子使用频率(relative synonymous codon usage, RSCU)>1 的密码子有 32 个, 其中 28 个密码子以 A/U 结尾, 只有 4 个以 C/G 结尾, 说明空心泡叶绿体基因组偏好使用 A/U 结尾的密码子。

表 2 空心泡叶绿体基因组注释基因信息

Table 2 Genes present in the chloroplast genome of *Rubus rosaefolius*

| Function of gene              | Group of gene                           | Gene name   |
|-------------------------------|---|---|
| Genes for photosynthesis (44) | ATP synthase                            | <i>atpA, atpB, atpE, atpF, atpH, atpI</i>   |
|                               | Cytochrome b/f complex                  | <i>petA, petB*, petD*, petG, petL, petN</i>   |
|                               | NADH dehydrogenase                      | <i>ndhA*, ndhB*<sup>1</sup>, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>   |
|                               | Photosystem I                           | <i>psaA, psaB, psaC, psaI, psaJ</i>   |
|                               | Photosystem II                          | <i>psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ</i>   |
|                               | Subunit of rubisco                      | <i>rbcL</i>   |
|                               | Ribosomal proteins (SSU)                | <i>rps2, rps3, rps4, rps7<sup>1</sup>, rps8, rps11, rps12<sup>#</sup>, rps14, rps15, rps16*, rps18, rps19</i>   |
|                               |   | <i>rpl2*<sup>1</sup>, rpl14, rpl16*, rpl20, rpl22, rpl23<sup>1</sup>, rpl32, rpl33, rpl36</i>   |
|                               | Ribosomal RNAs                          | <i>rrn4.5<sup>1</sup>, rrn5<sup>1</sup>, rrn16<sup>1</sup>, rrn23<sup>1</sup></i>   |
|                               | RNA polymerase                          | <i>rpoA, rpoB, rpoC1*, rpoC2</i>  |
| Self replication (59)         | Transfer RNAs                           | <i>trnA-UGC*<sup>1</sup>, trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnJ<sup>#</sup>-CAU, trnG-GCC, trnG-UCC*, trnH-GUG, trnI-CAU<sup>1</sup>, trnI-GAU*<sup>1</sup>, trnK-UUU*, trnL-CAA<sup>1</sup>, trnL-UAA*, trnL-UAG, trnM-CAU, trnN-GUU<sup>1</sup>, trnP-UGG, trnQ-UUG, trnR-ACG<sup>1</sup>, trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC<sup>1</sup>, trnV-UAC*, trnW-CCA, trnY-GUA</i> |
|                               | Protease                                | <i>clpP**</i>   |
|                               | Maturase                                | <i>matK</i>   |
|                               | Envelop membrane protein                | <i>cemA</i>   |
|                               | Subunits of acetyl-CoA-carboxylase      | <i>accD</i>   |
|                               | Translation initiation factor           | <i>infA</i>   |
|                               | C-type cytochrome synthesis             | <i>ccsA</i>   |
| Unknown function (4)          | Hypothetical chloroplast reading frames | <i>ycf1<sup>1</sup>, ycf2<sup>1</sup>, ycf3**, ycf4</i>   |

\*: Genes with one intron; \*\*: Genes with two introns; #: Trans-spliced genes; <sup>1</sup>: Genes with two copies in IR region.

表 3 空心泡叶绿体基因组的重复序列

Table 3 Repetitive sequences of chloroplast genome of *Rubus rosaefolius*

| Number | Size   | Position1 | Type | Position2 | Location 1               | Location 2               |
|--------|--------|-----------|------|-----------|--------------------------|--------------------------|
| 1      | 25 748 | 85 443    | P    | 129 902   | <i>rps19-rpl2</i>        | <i>ycf1</i>              |
| 2      | 56     | 10 263    | P    | 10 263    | <i>trnS-GCU-trnG-UCC</i> | <i>trnS-GCU-trnG-UCC</i> |
| 3      | 48     | 67 316    | F    | 67 337    | <i>trnW-CCA-trnP-UGG</i> | <i>trnW-CCA-trnP-UGG</i> |
| 4      | 38     | 53 778    | P    | 53 778    | <i>trnM-CAU-atpE</i>     | <i>trnM-CAU-atpE</i>     |
| 5      | 40     | 99 849    | F    | 121 587   | <i>rps7-trnV-GAC</i>     | <i>ndhA</i>              |

(待续)

(续表 3)

| Number | Size | Position1 | Type | Position2 | Location 1               | Location 2               |
|--------|------|-----------|------|-----------|--------------------------|--------------------------|
| 6      | 40   | 121 587   | P    | 141 204   | <i>ndhA</i>              | <i>trnV-GAC-rps7</i>     |
| 7      | 39   | 44 392    | F    | 99 851    | <i>ycf3</i>              | <i>rps7-trnV-GAC</i>     |
| 8      | 39   | 44 392    | P    | 141 203   | <i>ycf3</i>              | <i>trnV-GAC-rps7</i>     |
| 9      | 34   | 52 613    | P    | 52 613    | <i>ndhC-trnV-UAC</i>     | <i>ndhC-trnV-UAC</i>     |
| 10     | 33   | 10 492    | F    | 10 547    | <i>trnS-GCU-trnG-UCC</i> | <i>trnS-GCU-trnG-UCC</i> |
| 11     | 33   | 71 406    | F    | 71 436    | <i>clpP</i>              | <i>clpP</i>              |
| 12     | 32   | 29 407    | F    | 29 439    | <i>petN-psbM</i>         | <i>petN-psbM</i>         |
| 13     | 38   | 44 392    | F    | 121 589   | <i>ycf3</i>              | <i>ndhA</i>              |
| 14     | 38   | 75 352    | P    | 75 352    | <i>psbT-psbN</i>         | <i>psbT-psbN</i>         |
| 15     | 37   | 10 710    | F    | 10 728    | <i>trnG-UCC</i>          | <i>trnG-UCC</i>          |
| 16     | 30   | 8 446     | P    | 46 157    | <i>trnQ-UUG-psbK</i>     | <i>trnS-GGA</i>          |
| 17     | 32   | 47 646    | F    | 47 674    | <i>trnT-UGU-trnL-UAA</i> | <i>trnT-UGU-trnL-UAA</i> |
| 18     | 37   | 10 433    | R    | 52 678    | <i>trnS-GCU-trnG-UCC</i> | <i>ndhC-trnV-UAC</i>     |
| 19     | 31   | 10 403    | F    | 10 448    | <i>trnS-GCU-trnG-UCC</i> | <i>trnS-GCU-trnG-UCC</i> |
| 20     | 31   | 13 025    | F    | 13 042    | <i>atpF-atpH</i>         | <i>atpF-atpH</i>         |
| 21     | 36   | 82 892    | F    | 82 913    | <i>rpl16</i>             | <i>rpl16</i>             |
| 22     | 30   | 44 404    | F    | 99 863    | <i>ycf3</i>              | <i>rps7-trnV-GAC</i>     |
| 23     | 30   | 44 404    | P    | 141 200   | <i>ycf3</i>              | <i>trnV-GAC-rps7</i>     |
| 24     | 30   | 67 699    | R    | 67 702    | <i>trnP-UGG-psaI</i>     | <i>trnP-UGG-psaI</i>     |
| 25     | 34   | 58 510    | F    | 58 541    | <i>rbcL-accD</i>         | <i>rbcL-accD</i>         |
| 26     | 31   | 115 296   | R    | 115 296   | <i>rpl32-trnL-UAG</i>    | <i>rpl32-trnL-UAG</i>    |
| 27     | 30   | 10 448    | C    | 52 686    | <i>trnS-GCU-trnG-UCC</i> | <i>ndhC-trnV-UAC</i>     |
| 28     | 30   | 57 916    | F    | 57 941    | <i>rbcL</i>              | <i>rbcL</i>              |
| 29     | 30   | 60 287    | P    | 60 287    | <i>accD-psaI</i>         | <i>accD-psaI</i>         |
| 30     | 30   | 67 338    | F    | 67 359    | <i>trnW-CCA-trnP-UGG</i> | <i>trnW-CCA-trnP-UGG</i> |
| 31     | 30   | 108 593   | F    | 108 625   | <i>rrn4.5-rrn5</i>       | <i>rrn4.5-rrn5</i>       |
| 32     | 30   | 108 593   | P    | 132 438   | <i>rrn4.5-rrn5</i>       | <i>rrn5-rrn4.5</i>       |
| 33     | 30   | 108 625   | P    | 132 470   | <i>rrn4.5-rrn5</i>       | <i>rrn5-rrn4.5</i>       |
| 34     | 30   | 132 438   | F    | 132 470   | <i>rrn5-rrn4.5</i>       | <i>rrn5-rrn4.5</i>       |
| 35     | 32   | 8 444     | F    | 36 121    | <i>trnQ-UUG-psbK</i>     | <i>psbC-trnS-UGA</i>     |
| 36     | 32   | 10 409    | P    | 10 424    | <i>trnS-GCU-trnG-UCC</i> | <i>trnS-GCU-trnG-UCC</i> |
| 37     | 31   | 6 546     | P    | 6 546     | <i>rps16-trnQ-UUG</i>    | <i>rps16-trnQ-UUG</i>    |
| 38     | 31   | 60 402    | F    | 60 420    | <i>accD-psaI</i>         | <i>accD-psaI</i>         |
| 39     | 31   | 97 211    | P    | 97 211    | <i>ndhB</i>              | <i>ndhB</i>              |
| 40     | 31   | 97 211    | F    | 143 851   | <i>ndhB</i>              | <i>ndhB</i>              |
| 41     | 31   | 143 851   | P    | 143 851   | <i>ndhB</i>              | <i>ndhB</i>              |
| 42     | 30   | 9 893     | F    | 37 155    | <i>trnS-GCU-trnG-UCC</i> | <i>trnG-GCC</i>          |
| 43     | 30   | 10 403    | C    | 52 686    | <i>trnS-GCU-trnG-UCC</i> | <i>ndhC-trnV-UAC</i>     |
| 44     | 30   | 10 720    | F    | 10 738    | <i>trnG-UCC</i>          | <i>trnG-UCC</i>          |
| 45     | 30   | 36 123    | P    | 46 157    | <i>psbC-trnS-UGA</i>     | <i>trnS-GGA</i>          |
| 46     | 30   | 39 399    | F    | 41 623    | <i>psaB</i>              | <i>psaA</i>              |
| 47     | 30   | 67 317    | F    | 67 359    | <i>trnW-CCA-trnP-UGG</i> | <i>trnW-CCA-trnP-UGG</i> |

F: Forward repeat; R: Reverse repeat; C: Complement repeat; P: Palindromic repeat.

表 4 空心泡叶绿体 SSR 信息

Table 4 SSRs in chloroplast complete genomes of *Rubus rosaefolius*

| Type (number)       | Sequence | Number of repeats |   |   |   |   |   |    |    |    |    |    |    |       | Total (%) |
|---------------------|----------|-------------------|---|---|---|---|---|----|----|----|----|----|----|-------|-----------|
|                     |          | 4                 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |       |           |
| Mononucleotide (58) | A/T      | –                 | – | – | – | – | – | 27 | 9  | 10 | 5  | 3  | 1  | 80.56 |           |
|                     | C/G      | –                 | – | – | – | – | – | 1  | 2  |    |    |    |    |       |           |
| Dinucleotide (11)   | AG/CT    | –                 | 1 |   |   |   |   |    |    |    |    |    |    | 15.28 |           |
|                     | AT/AT    | –                 | 6 | 3 | 1 |   |   |    |    |    |    |    |    |       |           |
| Trinucleotide (3)   | AAG/CTT  | 1                 |   |   |   |   |   |    |    |    |    |    |    | 4.17  |           |
|                     | AAT/ATT  | 2                 |   |   |   |   |   |    |    |    |    |    |    |       |           |

–: No repeat.

表 5 空心泡各氨基酸相对同义密码子使用度

Table 5 Relative synonymous codon usage (RSCU) in the chloroplast genome of *Rubus rosaefolius*

| Amino acid | Codon | Number | RSCU | Ratio (%) | Amino acid | Codon | Number | RSCU | Ratio (%) |
|------------|-------|--------|------|-----------|------------|-------|--------|------|-----------|
| Phe        | UUU   | 2 262  | 1.25 | 6.96      | Tyr        | UAU   | 1 491  | 1.37 | 4.20      |
|            | UUC   | 1 348  | 0.75 |           |            | UAC   | 686    | 0.63 |           |
| Leu        | UUA   | 1 299  | 1.49 | 10.10     | TER*       | UAA   | 1 238  | 1.27 | 3.72      |
|            | UUG   | 1 070  | 1.23 |           |            | UAG   | 694    | 0.71 |           |
|            | CUU   | 997    | 1.14 |           | His        | CAU   | 821    | 1.37 | 2.31      |
|            | CUC   | 609    | 0.70 |           |            | CAC   | 378    | 0.63 |           |
|            | CUA   | 801    | 0.92 |           | Gln        | CAA   | 969    | 1.39 | 2.69      |
|            | CUG   | 463    | 0.53 |           |            | CAG   | 428    | 0.61 |           |
| Ile        | AUU   | 1 782  | 1.19 | 8.68      | Asn        | AAU   | 1 858  | 1.41 | 5.09      |
|            | AUC   | 1 052  | 0.70 |           |            | AAC   | 785    | 0.59 |           |
|            | AUA   | 1 667  | 1.11 |           | Lys        | AAA   | 2 166  | 1.39 | 6.00      |
| Met        | AUG   | 883    | 1.00 | 1.70      |            | AAG   | 945    | 0.61 |           |
| Val        | GUU   | 797    | 1.35 | 4.55      | Asp        | GAU   | 1 014  | 1.44 | 2.71      |
|            | GUC   | 397    | 0.67 |           |            | GAC   | 394    | 0.56 |           |
|            | GUA   | 758    | 1.29 |           | Glu        | GAA   | 1238   | 1.37 | 3.47      |
|            | GUG   | 407    | 0.69 |           |            | GAG   | 563    | 0.63 |           |
| Ser        | UCU   | 1 167  | 1.41 | 7.19      | Cys        | UGU   | 729    | 1.25 | 2.24      |
|            | UCC   | 950    | 1.15 |           |            | UGC   | 434    | 0.75 |           |
|            | UCA   | 1 014  | 1.23 |           | TER*       | UGA   | 994    | 1.02 | 1.92      |
|            | UCG   | 601    | 0.73 |           | Trp        | UGG   | 677    | 1.00 |           |
| Pro        | CCU   | 659    | 1.08 | 4.71      | Arg        | CGU   | 370    | 0.69 | 3.10      |
|            | CCC   | 639    | 1.05 |           |            | CGC   | 266    | 0.49 |           |
|            | CCA   | 714    | 1.17 |           |            | CGA   | 582    | 1.08 |           |
|            | CCG   | 433    | 0.71 |           |            | CGG   | 390    | 0.72 |           |
| Thr        | ACU   | 725    | 1.22 | 4.60      | Ser        | AGU   | 705    | 0.85 | 2.35      |
|            | ACC   | 577    | 0.97 |           |            | AGC   | 514    | 0.62 |           |
|            | ACA   | 678    | 1.14 |           | Arg        | ACA   | 1 044  | 1.94 | 3.13      |
|            | ACG   | 406    | 0.68 |           |            | AGG   | 582    | 1.08 |           |
| Ala        | GCU   | 497    | 1.32 | 2.91      | Gly        | GGU   | 566    | 1.00 | 4.36      |
|            | GCC   | 344    | 0.91 |           |            | GGC   | 374    | 0.66 |           |
|            | GCA   | 440    | 1.17 |           |            | GGA   | 764    | 1.35 |           |
|            | GCG   | 229    | 0.61 |           |            | GGG   | 559    | 0.99 |           |

\*: Represents the stop codon.



## 2.5 IR区收缩与扩张分析

将空心泡叶绿体基因组的 IR-LSC 和 IR-SSC 边界与已报道的 4 个悬钩子属物种进行比较, 图 2 显示, 5 个物种的 LSC/IRb 边界均位于 *rps19* 和 *rpl2* 之间, 空心泡 *rps19* 基因距离 LSC/IRb 边界 25 bp, 光滑悬钩子距离边界 16 bp, 其余 3 个物种均距离边界 24 bp。除了光滑悬钩子外, 其余 4 个物种的 IRb/SSC 边界均存在完整的 *ycf1* 基因; 其中腺萼悬钩子 *ycf1* 基因位于 IRb 区部分的长度为 1 092 bp, 而小叶悬钩子、能高悬钩子和空心泡 *ycf1* 基因位于 IRb 区部分的长度均为 1 215 bp; *ycf1* 基因进入 SSC 区的长度从 14 bp 到 26 bp, 其中小叶悬钩子和能高悬钩子 *ycf1* 扩张最大。5 个物种在 IRb/SSC 边界的 SSC 一侧均存在完整的 *ndhF* 基因, 且均距离 IRb/SSC 边界 16 bp 的长度。5 个物种在 SSC/IRa 边界存在完整的 *ycf1* 基因, 光滑悬钩子的 *ycf1* 基因长度为 5 036 bp, 空心泡 *ycf1* 基因长度为 5 705 bp, 其余 3 个物种的 *ycf1* 基因长度均为 5 819 bp。5 个物种在 IRa 区均存在 *trnN* 基因。5 个物种在 IRa/LSC 边界的 LSC

一侧均存在 *trnH* 基因和 *psbA* 基因, 光滑悬钩子的 *trnN* 基因刚好与边界齐平, 空心泡 *trnN* 基因距离边界 2 bp, 其余 3 个物种均距离边界 1 bp。

## 2.6 系统进化分析

以空心泡、33 个悬钩子属植株和刺蔷薇(作为外群)的叶绿体基因组构建最大似然(maximum likelihood, ML)系统发育树, 以获取空心泡与它们的关系。结果显示, 能高悬钩子(*Rubus rubroangustifolius*)和腺萼悬钩子(*Rubus glandulosopunctatus*)先聚为一支, 再与小叶悬钩子(*Rubus taiwanicola*)聚为一支, 这 3 个物种组成的分支与空心泡聚为一支, 自展支持率大于 90%, 结果可信度高, 说明空心泡与小叶悬钩子亲缘关系最近, 其次是能高悬钩子和腺萼悬钩子(图 3)。

## 3 讨论

悬钩子属植物具有高度可变的形态, 复杂的无融合生殖、多倍体和杂交等情况, 是开花植物中最难分类的类群之一<sup>[29]</sup>。徐玉芬<sup>[30]</sup>在悬钩子属植物叶绿体基因组的研究中发现, 悬钩子属

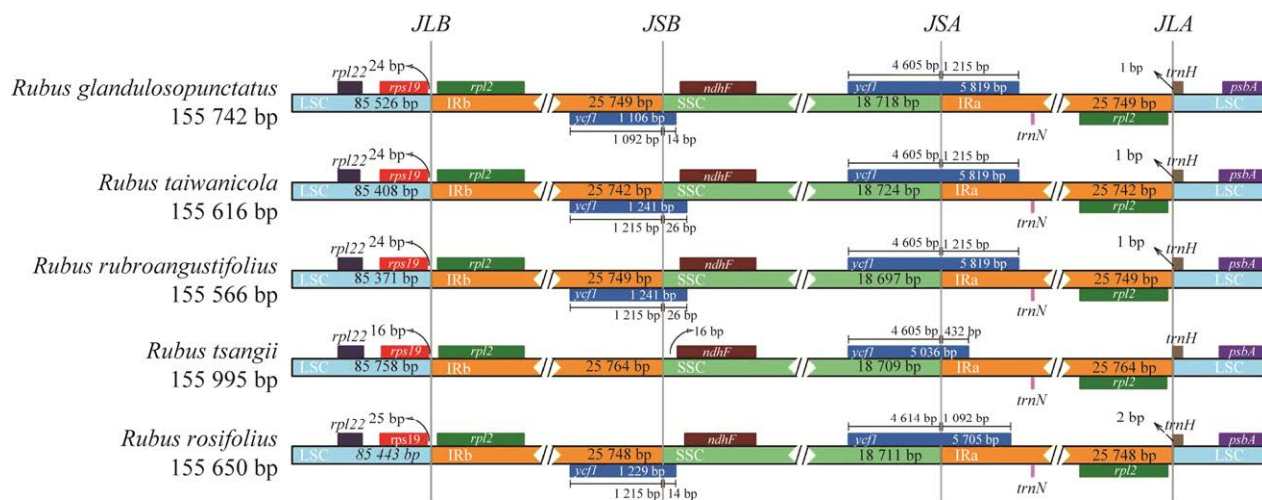


图 2 悬钩子属 5 个物种的叶绿体基因组 IR 区边界分析

Figure 2 Boundary analysis of IR regions of chloroplast genomes from five *Rubus* plant.

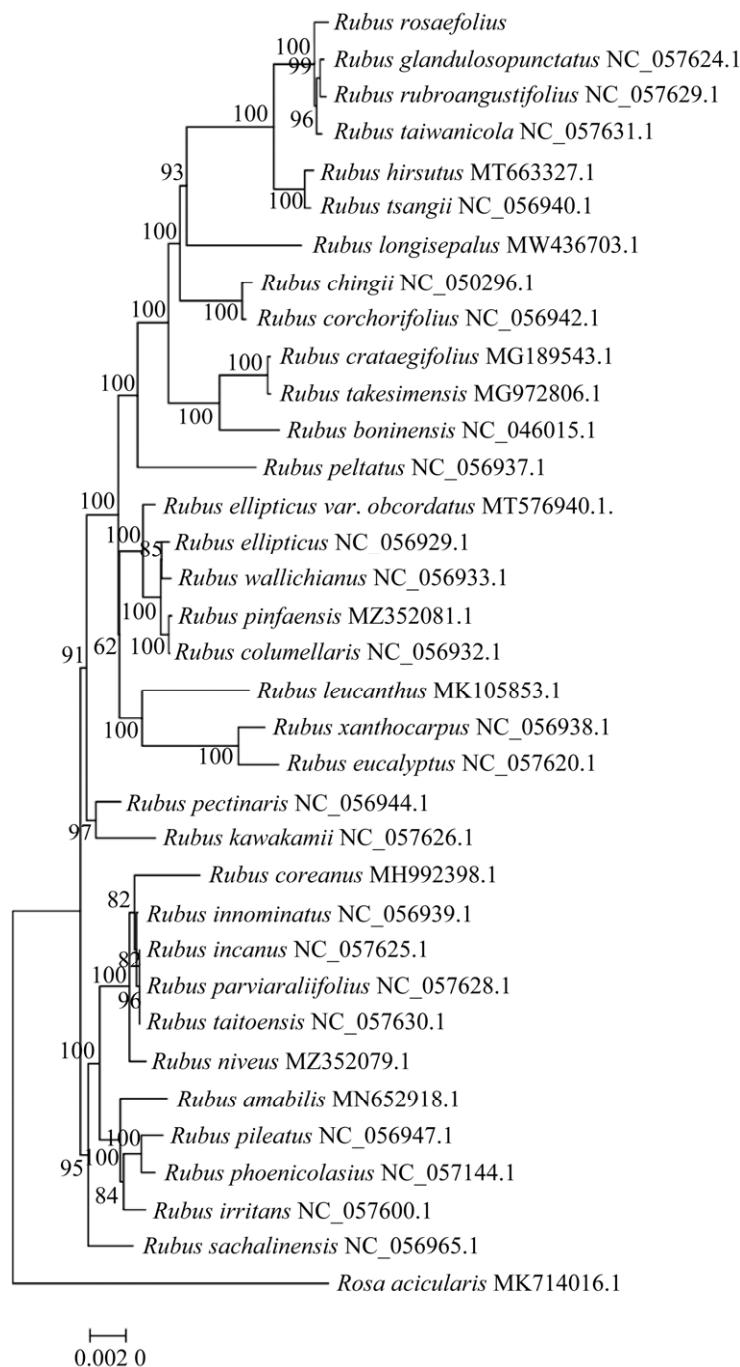


图3 基于35个物种完整叶绿体基因组的ML系统发育树

Figure 3 The ML phylogenetic tree based on 35 species of complete chloroplast genomes. The number on the right side of the node represents the bootstrap value of 1 000 iterations.

植物叶绿体基因组大小为 155 464–156 506 bp，能编码 126–135 个基因，总 GC 含量为 36.9%–37.3%。本研究组装的空心泡叶绿体基因组全序

列长度为 155 650 bp，注释了 131 个基因，GC 含量为 36.9%。这与高粱泡<sup>[31]</sup>、牛叠肚<sup>[32]</sup>、多腺悬钩子<sup>[33]</sup>等悬钩子属植物的叶绿体基因组结

构特征类似,表明悬钩子属植物叶绿体基因组大小、基因组成等高度保守。在亲缘关系较近的蓬蘽<sup>[34]</sup>中发现,其叶绿体基因组(156 380 bp)和总 GC 含量(37.0%)均大于空心泡,再结合形态(空心泡叶子较蓬蘽细,且多以 5 到 7 叶为主,蓬蘽 3 叶到 5 叶为主)、地域分布(空心泡主要分布于江西、湖南、四川、贵州等地,蓬蘽主要分布于河南、江西、浙江、江苏等地),将二者初步区分,但由于悬钩子属近源物种之间易杂交产生新品种,使其分类及系统发育关系的厘清工作十分困难。

IR 区、LSC 区和 SSC 区的扩张和收缩会引起叶绿体基因组大小出现差异,是物种进化过程中常见的现象<sup>[35]</sup>。本次研究通过比较 5 个悬钩子属植物 IR 区边界附近的基因位置和长度(*rps19*、*rpl2*、*ndhF*、*ycf1* 和 *trnH*),来表现基因的收缩与扩张。结果显示,5 个悬钩子属植物叶绿体基因组的 LSC/IR 和 SSC/IR 边界高度相似,5 个物种 LSC/IRb 边界均位于 *rps19* 和 *rpl2* 之间,在 IRb/SSC 边界的 SSC 一侧均存在完整的 *ndhF* 基因,且距离 IRb/SSC 边界的长度相同,在 SSC/IRa 边界均存在完整的 *ycf1* 基因,这与 Yu 等<sup>[29]</sup>的研究结果一致,说明悬钩子属物种 IR 区与 LSC 和 SSC 区的边界保守性较高。此外,空心泡在 LSC/IRb 边界的 *rps19* 基因及 SSC/IRa 边界的 *ycf1* 基因均存在收缩现象,在 IRa/LSC 边界的 *trnH* 基因存在扩张现象,其基因收缩和扩张的原因可能是 IR 边界变化引起<sup>[30]</sup>。因此,尽管悬钩子物种的叶绿体基因组结构和基因排序较为保守,但 IR 边界仍存在的收缩扩张变化,表明悬钩子属叶绿体基因组中基因与其排序的保守性和 IR 边界收缩扩张的多样性并存。

重复序列在叶绿体基因组中发挥着重要的进化作用,能促进叶绿体基因组重排,影响基

因组结构的变化,并且可以增加种群遗传多样性<sup>[36]</sup>。空心泡叶绿体基因组中共含有 47 个散在重复序列,其中正向重复和回文重复为主要重复序列,占比接近 90%,大多数重复序列分布在基因间隔区和内含子序列中。重复序列分布现象与豆瓣菜叶绿体基因组<sup>[37]</sup>的研究结果一致。简单重复序列是叶绿体基因组中一段长度为 1–6 bp 的重复序列,具有共显性、多等位基因、高度可变、多态性等特点,是研究遗传变异性的最佳标记<sup>[38–39]</sup>。本研究结果显示,空心泡叶绿体基因组共检测出 72 个 SSR 位点,占比最大的是单核苷酸(80.56%),其中含有丰富的 AT 碱基,而多聚的 A 或 T 为最多的类型,这种现象在桃属<sup>[40]</sup>、山楂属<sup>[41]</sup>等中也有报道,进一步证实了被子植物的 SSR 主要是由短 poly A 和 poly T 构成。与同属的插田泡(44SSR)、光滑高粱泡(54SSR)和黑莓‘阿拉好’(40SSR)叶绿体基因组相比,本研究发现空心泡所含 SSR 位点大于这 3 个物种,且碱基类型更为丰富,空心泡包含单核苷酸、二核苷酸、三核苷酸三种类型,而插田泡、光滑高粱泡和黑莓‘阿拉好’均只含 2 种碱基类型<sup>[42]</sup>。空心泡重复序列的获取,为悬钩子属植物的进一步的演化研究提供参考。

密码子偏好性是普遍存在于生物体内基因组中一个重要的进化特征,对于基因表达和生物体进化具有重要意义<sup>[43]</sup>。同义密码子使用频率(RSCU)能够直观反映出密码子使用的偏好性,当密码子的 RSCU>1 时,表明该密码子使用频率较高,当密码子的 RSCU=1 时,表明该密码子没有偏好性,RSCU<1 表明该密码子使用频率比其他同义密码子低<sup>[44]</sup>。本研究发现,高频密码子有 32 个,其中 87.5%的密码子第 3 位碱基为 A/U,说明空心泡叶绿体基因组具有偏好性,偏好使用 A/U 结尾的密码子。在同属的高粱

泡<sup>[31]</sup>研究中发现, RSCU>1 的密码子有 33 个, 其中, 28 个密码子的第 3 位碱基为 A/U, 4 个以 C/G 结尾。本研究中 RSCU>1 的密码子虽比高粱泡少, 但其密码子偏好性与高粱泡一致。

叶绿体相关基因的表达虽受到核基因的调控, 但其基因组含有大量的遗传信息, 且叶绿体基因组一般缺乏重组、核苷酸替换率低以及通常为单亲遗传, 因此在物种鉴定和系统发育等研究中发挥着重要的作用<sup>[45-46]</sup>。系统发育分析表明, 空心泡与小叶悬钩子关系最为密切, 其次是能高悬钩子和腺萼悬钩子。

## 4 结论

本研究采用高通量测序技术, 对空心泡叶绿体进行测序, 获得完整的叶绿体基因组序列。空心泡的完整叶绿体基因组总长度为 155 650 bp, 具有典型的四分体结构, 包括 2 个反向重复序列(IRs 各 25 748 bp)、1 个大拷贝区(85 443 bp)、1 个小拷贝区(18 711 bp); 共鉴定出 131 个基因, 包括 86 个蛋白质编码基因、37 个 tRNA 基因和 8 个 rRNA 基因, 全基因组的 GC 含量为 36.9%; 共鉴定了 47 个散在重复序列, 以正向重复和回文重复为主, 主要分布在基因间隔区和内含子序列; 包含 72 个 SSR 位点, 其中单核苷酸 58 个, 二核苷酸 11 个、三核苷酸 3 个; 密码子使用率最高的是亮氨酸, 使用率最低的是色氨酸, 偏好使用 A/U 结尾的密码子。空心泡与同属植物中小叶悬钩子、能高悬钩子和腺萼悬钩子亲缘关系最近。空心泡叶绿体基因组的解析, 确定了空心泡在悬钩子属的系统位置和系统发育关系, 为今后悬钩子属植物系统发育分析提供理论依据。

## REFERENCES

- [1] LIU HZ, YE H, ZHANG NY, MA JY, WANG JT, HU GJ, LI MD, ZHAO P. Comparative analyses of chloroplast genomes provide comprehensive insights into the adaptive evolution of *Paphiopedilum* (Orchidaceae)[J]. *Horticulturae*, 2022, 8(5): 391.
- [2] 热伊汉古丽·图尔迪, 慕丽红, 田新民. 扁果草叶绿体基因组特征分析[J]. *生物工程学报*, 2022, 38(8): 2999-3013.
- [3] TURDI R, MU LH, TIAN XM. Characteristics of the chloroplast genome of *Isopyrum anemonoides*[J]. *Chinese Journal of Biotechnology*, 2022, 38(8): 2999-3013 (in Chinese).
- [4] DU XY, ZENG T, FENG Q, HU LJ, LUO X, WENG QB, HE JF, ZHU B. The complete chloroplast genome sequence of yellow mustard (*Sinapis alba* L.) and its phylogenetic relationship to other Brassicaceae species[J]. *Gene*, 2020, 731: 144340.
- [5] 欧金梅, 杨亚滢, 钱程程, 吴瑞, 王瑞, 黄璐琦. 基于叶绿体基因的药用梅群体遗传学研究[J]. *中草药*, 2022, 53(17): 5469-5475.
- [6] OU JM, YANG YT, QIAN CC, WU R, WANG R, HUANG LQ. Analysis of population genetics of medicinal plum based on chloroplast gene[J]. *Chinese Traditional and Herbal Drugs*, 2022, 53(17): 5469-5475 (in Chinese).
- [7] SHINOZAKI K, OHME M, TANAKA M, WAKASUGI T, HAYASHIDA N, MATSUBAYASHI T, ZAITA N, CHUNWONGSE J, OBOKATA J, YAMAGUCHI-SHINOZAKI K, OHTO C, TORAZAWA K, MENG BY, SUGITA M, DENO H, KAMOGASHIRA T, YAMADA K, KUSUDA J, TAKAIWA F, KATO A, et al. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression[J]. *The EMBO Journal*, 1986, 5(9): 2043-2049.
- [8] 郑祎, 张卉, 王钦美, 高悦, 张志宏, 孙玉新. 大花君子兰叶绿体基因组及其特征[J]. *园艺学报*, 2020, 47(12): 2439-2450.
- [9] ZHENG Y, ZHANG H, WANG QM, GAO Y, ZHANG ZH, SUN YX. Complete chloroplast genome sequence of *Clivia miniata* and its characteristics[J]. *Acta Horticulturae Sinica*, 2020, 47(12): 2439-2450 (in Chinese).
- [10] RASHEED S, ZAIDI S, AZIM MK. The chloroplast genome sequence of *Momordica charantia* L. (bitter gourd)[J]. *Gene Reports*, 2020, 21: 100963.
- [11] MO ZH, LOU WR, CHEN YQ, JIA XD, ZHAI M,

- GUO ZR, XUAN JP. The chloroplast genome of *Carya illinoensis*: genome structure, adaptive evolution, and phylogenetic analysis[J]. *Forests*, 2020, 11(2): 207.
- [9] YANG AH, LIU SJ, LIU TY, HU M, ZHONG YD, LIU LP, YU FX. The complete chloroplast genome sequence of *Actinidia styracifolia* C. F. Liang[J]. *Mitochondrial DNA Part B*, 2020, 5(1): 90-91.
- [10] QUADROS APO, ALMEIDA LM, PETREANU M, NIERO R, ROSA PCP, SAWAYA ACHF, MANTOVANI MS, GAIVÃO IOM, MAISTRO EL. Risk assessment via genotoxicity, metabolism, apoptosis, and cell growth effects in a HepG2/C3A cell line upon treatment with *Rubus rosifolius* (Rosaceae) leaves extract[J]. *Journal of Toxicology and Environmental Health Part A*, 2020, 83(13/14): 495-508.
- [11] 刘世彪, 吕江明, 孙蓓育, 蒋孝波, 张世鑫, 彭小列. 三种野生莓类果实的营养成分及蛇莓急性毒性实验[J]. *营养学报*, 2009, 31(3): 307-309.
- LIU SB, LV JM, SUN BY, JIANG XB, ZHANG SX, PENG XL. The nutritional components of wild fruits *Rubus corchorifolius*, *R. rosaefolius*, *Duchesnea indica* and acute toxicity of *D. indica* fruit in mice[J]. *Acta Nutrimenta Sinica*, 2009, 31(3): 307-309 (in Chinese).
- [12] BAGATTOLI PCD, CIPRIANI DC, MARIANO LNB, CORREA M, WAGNER TM, NOLDIN VF, FILHO VC, NIERO R. Phytochemical, antioxidant and anticancer activities of extracts of seven fruits found in the southern Brazilian flora[J]. *Indian Journal of Pharmaceutical Sciences*, 2016, 78(1): 34-40.
- [13] de SOUZA P, BOEING T, SOMENSI LB, CECHINEL-ZANCHETT CC, BASTOS JK, PETREANU M, NIERO R, CECHINEL-FILHO V, da SILVA LM, de ANDRADE SF. Diuretic effect of extracts, fractions and two compounds 2 $\alpha$ ,3 $\beta$ ,19 $\alpha$ -trihydroxy-urs-12-en-28-oic acid and 5-hydroxy-3,6,7,8,4'-pentamethoxyflavone from *Rubus rosaefolius* Sm. (Rosaceae) leaves in rats[J]. *Naunyn-Schmiedeberg's Archives of Pharmacology*, 2017, 390(4): 351-360.
- [14] OLIVEIRA BD, RODRIGUES AC, CARDOSO BMI, RAMOS ALCC, BERTOLDI MC, TAYLOR JG, da CUNHA LR, PINTO UM. Antioxidant, antimicrobial and anti-quorum sensing activities of *Rubus rosaefolius* phenolic extract[J]. *Industrial Crops and Products*, 2016, 84: 59-66.
- [15] PETREANU M, FERREIRA EK, SAGAZ APM, VENDRAMINI-COSTA DB, RUIZ ALTG, de CARVALHO JE, CAMPOS A, CECHINEL FILHO V, DELLE MONACHE F, NIERO R. Uncommon trimethoxylated flavonol obtained from *Rubus rosaefolius* leaves and its antiproliferative activity[J]. *Evidence-Based Complementary and Alternative Medicine: ECAM*, 2015, 2015: 341216.
- [16] 陈炳华, 刘剑秋, 黄惠玲, 许春琼. 空心泡红色素的性质及其稳定性研究[J]. *福建师范大学学报(自然科学版)*, 2002, 18(3): 77-83.
- CHEN BH, LIU JQ, HUANG HL, XU CQ. Basic properties of *Rubus rosaefolius* red pigment and its stability[J]. *Journal of Fujian Teachers University (Natural Science)*, 2002, 18(3): 77-83 (in Chinese).
- [17] POREBSKI S, BAILEY LG, BAUM BR. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components[J]. *Plant Molecular Biology Reporter*, 1997, 15(1): 8-15.
- [18] PATEL RK, JAIN M. NGS QC toolkit: a toolkit for quality control of next generation sequencing data[J]. *PLoS One*, 2012, 7(2): e30619.
- [19] BANKEVICH A, NURK S, ANTIPOV D, GUREVICH AA, DVORKIN M, KULIKOV AS, LESIN VM, NIKOLENKO SI, PHAM S, PRJIBELSKI AD, PYSHKIN AV, SIROTKIN AV, VYAHHI N, TESLER G, ALEKSEYEV MA, PEVZNER PA. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing[J]. *Journal of Computational Biology: a Journal of Computational Molecular Cell Biology*, 2012, 19(5): 455-477.
- [20] QU XJ, MOORE MJ, LI DZ, YI TS. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes[J]. *Plant Methods*, 2019, 15: 50.
- [21] GREINER S, LEHWARK P, BOCK R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes[J]. *Nucleic Acids Research*, 2019, 47(W1): W59-W64.
- [22] SHIELDS DC, SHARP PM. Synonymous codon usage in *Bacillus subtilis* reflects both translational selection and mutational biases[J]. *Nucleic Acids Research*, 1987, 15(19): 8023-8040.
- [23] LOHSE M, DRECHSEL O, KAHLAU S, BOCK R. OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial

- genomes and visualizing expression data sets[J]. Nucleic Acids Research, 2013, 41(W1): W575-W581.
- [24] THIEL T, MICHALEK W, VARSHNEY R, GRANER A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.)[J]. Theoretical and Applied Genetics, 2003, 106(3): 411-422.
- [25] AMIRYOUSEFI A, HYVÖNEN J, POCZAI P. IRscope: an online program to visualize the junction sites of chloroplast genomes[J]. Bioinformatics, 2018, 34(17): 3030-3031.
- [26] KATOH K, STANDLEY DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability[J]. Molecular Biology and Evolution, 2013, 30(4): 772-780.
- [27] KALYAANAMOORTHY S, MINH BQ, WONG TKF, von HAESELER A, JERMIIN LS. ModelFinder: fast model selection for accurate phylogenetic estimates[J]. Nature Methods, 2017, 14(6): 587-589.
- [28] STAMATAKIS A. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies[J]. Bioinformatics, 2014, 30(9): 1312-1313.
- [29] YU JJ, FU J, FANG YP, XIANG J, DONG HJ. Complete chloroplast genomes of *Rubus* species (Rosaceae) and comparative analysis within the genus[J]. BMC Genomics, 2022, 23(1): 32.
- [30] 徐玉芬. 中国悬钩子属叶绿体系统发育基因组学研究[D]. 广州: 仲恺农业工程学院硕士学位论文, 2020.
- XU YF. Chloroplast phylogenomics of *Rubus* in China[D]. Guangzhou: Master's Thesis of Zhongkai University of Agriculture and Engineering, 2020 (in Chinese).
- [31] 赵文植, 董章宏, 辛静, 沈伟祥, 王飞, 徐建, 常晓勇, 辛培尧. 高粱泡叶绿体基因组特征分析[J]. 云南农业大学学报(自然科学), 2022, 37(3): 435-446.
- ZHAO WZ, DONG ZH, XIN J, SHEN WX, WANG F, XU J, CHANG XY, XIN PY. Analysis on the chloroplast genome characteristics of *Rubus lambertianus* Ser[J]. Journal of Yunnan Agricultural University (Natural Science), 2022, 37(3): 435-446 (in Chinese).
- [32] YANG JY, PAK JH, KIM SC. The complete chloroplast genome sequence of Korean raspberry *Rubus crataegifolius* (Rosaceae)[J]. Mitochondrial DNA Part B, 2017, 2(2): 793-794.
- [33] ZHANG GY, LIU YR, HAI P. The complete chloroplast genome of Tibetan medicinal plant *Rubus phoenicolasius* Maxim[J]. Mitochondrial DNA Part B, 2021, 6(3): 886-887.
- [34] WANG QR, HUANG ZR, GAO CS, GE YQ, CHENG RB. The complete chloroplast genome sequence of *Rubus hirsutus* Thunb. and a comparative analysis within *Rubus* species[J]. Genetica, 2021, 149(5): 299-311.
- [35] WANG WC, CHEN SY, ZHANG XZ. Whole-genome comparison reveals divergent IR borders and mutation hotspots in chloroplast genomes of herbaceous bamboos (Bambusoideae: Olyreae)[J]. Molecules (Basel, Switzerland), 2018, 23(7): 1537.
- [36] PARK I, YANG S, CHOI G, KIM WJ, MOON BC. The complete chloroplast genome sequences of *Aconitum pseudolaevae* and *Aconitum longecassidatum*, and development of molecular markers for distinguishing species in the *Aconitum* subgenus lycopodium[J]. Molecules (Basel, Switzerland), 2017, 22(11): 2012.
- [37] YAN C, DU JC, GAO L, LI Y, HOU XL. The complete chloroplast genome sequence of watercress (*Nasturtium officinale* R. Br.): genome organization, adaptive evolution and phylogenetic relationships in Cardamineae[J]. Gene, 2019, 699: 24-36.
- [38] CUSARO CM, GRAZIOLI C, ZAMBUTO F, CAPELLI E, BRUSONI M. An improved method for assessing simple sequence repeat (SSR) variation in *Echinochloa crus-galli* (L.) P. Beauv (barnyardgrass)[J]. Diversity, 2021, 14(1): 3.
- [39] 张雨, 苏旭, 刘玉萍, 刘涛, 郑长远, 苏丹丹, 王亚男, 吕婷. 喜马红景天叶绿体基因组特征及其系统发育分析[J]. 植物研究, 2022, 42(4): 602-612.
- ZHANG Y, SU X, LIU YP, LIU T, ZHENG CY, SU DD, WANG YN, LÜ T. Characteristics of complete chloroplast genome and phylogenetic analysis of *Rhodiola himalensis* (Crassulaceae)[J]. Bulletin of Botanical Research, 2022, 42(4): 602-612 (in Chinese).
- [40] DU ZY, LU K, ZHANG K, HE YM, WANG HT, CHAI GQ, SHI JG, DUAN YZ. The chloroplast genome of *Amygdalus* L. (Rosaceae) reveals the phylogenetic relationship and divergence time[J]. BMC Genomics, 2021, 22(1): 645.
- [41] WU LW, CUI YX, WANG Q, XU ZC, WANG Y, LIN YL, SONG JY, YAO H. Identification and

- phylogenetic analysis of five *Crataegus* species (Rosaceae) based on complete chloroplast genomes[J]. *Planta*, 2021, 254(1): 14.
- [42] 杨芮. 三个树莓类资源叶绿体基因组比较分析[D]. 雅安: 四川农业大学硕士学位论文, 2019.
- YANG R. Complete chloroplast genome and comparative analyses of three bramble resources (*Rubus*)[D]. Yaan: Master's Thesis of Sichuan Agricultural University, 2019 (in Chinese).
- [43] LIU SX, XUE DY, CHENG R, HAN HX. The complete mitogenome of *Apocheima cinerarius* (Lepidoptera: Geometridae: Ennominae) and comparison with that of other lepidopteran insects[J]. *Gene*, 2014, 547(1): 136-144.
- [44] 王婧, 王天翼, 王罗云, 张建国, 曾艳飞. 沙枣叶绿体全基因组序列及其使用密码子偏性分析[J]. 西北植物学报, 2019, 39(9): 1559-1572.
- WANG J, WANG TY, WANG LY, ZHANG JG, ZENG YF. Assembling and analysis of the whole chloroplast genome sequence of *Elaeagnus angustifolia* and its codon usage bias[J]. *Acta Botanica Boreali-Occidentalia Sinica*, 2019, 39(9): 1559-1572 (in Chinese).
- [45] 唐钰莹, 刘阳轩, 潘婷, 魏明, 马欣荣, 汪松虎. 叶绿体蛋白 CV 调节质体逆向信号的研究[J]. 植物生理学报, 2020, 56(7): 1553-1560.
- TANG YY, LIU YX, PAN T, WEI M, MA XR, WANG SH. Study on chloroplast protein CV regulating plastid retrograde signaling[J]. *Plant Physiology Journal*, 2020, 56(7): 1553-1560 (in Chinese).
- [46] FENG Y, GAO XF, ZHANG JY, JIANG LS, LI X, DENG HN, LIAO M, XU B. Complete chloroplast genomes provide insights into evolution and phylogeny of *Campylotropis* (Fabaceae)[J]. *Frontiers in Plant Science*, 2022, 13: 895543.

(本文责编 陈宏宇)