# 射干叶绿体基因组结构、序列特征与系统发育分析

蒋　明[1]，王军峰[2]，朱　晏[1]，黄雯馨[1]，应梦豪[1]，马佳莹[1]，戴晨宇[1]

1. 台州学院生命科学学院，浙江 台州　318000
2. 华东药用植物园科研管理中心，浙江 丽水　323000

**摘　要：目的**　以射干 *Belamcanda chinensis* 为材料，在测序、组装获得叶绿体基因组的基础上，明确其结构、序列特征及系统发育关系。**方法**　利用 PE150 双末端策略进行建库测序，用 NOVOPlasty 组装完整的叶绿体基因组，经 PCR 验证边界，借助生物信息学工具进行序列分析和系统发育研究。**结果**　射干的叶绿体基因组全长为 153 816 bp，大单拷贝区、反向重复区和小单拷贝区的长度分别为 83 143、26 214、18 245 bp。射干叶绿体基因组共有 133 个基因，编码基因、tRNA 和 rRNA 的数量分别为 92、38 和 8；*ycf1* 有 2 个拷贝，其中一个为假基因。系统发育分析结果表明，7 种植物叶绿体基因组在发育树上可分为 4 组，射干与同为鸢尾科的溪荪聚为一组，支持率达 100%。**结论**　射干叶绿体基因组的组装、序列分析和系统发育分析，为该药用植物的遗传结构和遗传多样性研究奠定了基础。

**关键词：**射干；叶绿体基因组；结构；序列分析；鸢尾科

**中图分类号：**R282.12　　**文献标志码：**A　　**文章编号：**0253 - 2670(2021)13 - 4039 - 08

**DOI：**10.7501/j.issn.0253-2670.2021.13.027

# Structure, sequence characteristics, and phylogenetic evolution analysis of *Belamcanda chinensis* chloroplast genome

JIANG Ming[1], WANG Jun-feng[2], ZHU Yan[1], HUANG Wen-xin[1], YING Meng-hao[1], MA Jia-ying[1], DAI Chen-yu[1]

1. College of Life Science, Taizhou University, Taizhou 318000, China
2. Scientific Research Management Center, East China Medicinal Botanical Garden, Lishui 323000, China

**Abstract: Objective**　To confirm the genome structure, sequence characteristics, and phylogenetic relationship by sequencing and assembling the chloroplast genome of a medicinal plant *Belamcanda chinensis*. **Methods**　A PE150 strategy was applied to construct library. The complete chloroplast genome was generated using NOVOPlasty, followed by PCR confirmation of the borders, and sequence analysis, as well as phylogenetic study was conducted by bioinformatic tools. **Results**　The full-length chloroplast genome was 153 816 bp in length, with a large single copy of 83 143 bp, an inverted repeat of 26 214 bp, and a small single copy of 18 245 bp. The *B. chinensis* chloroplast genome consisted of 133 genes, including 92 protein-coding genes, 38 tRNAs, and 8 rRNAs, respectively. There were two *ycf1* genes, one of which was a pseudogene. Phylogenetic evolution analysis results indicated that the seven chloroplast genomes can be divided into four groups, *B. chinensis* and *Iris sanguine* from Iridaceae were found to cluster in the same clade, with a support rate of 100%. **Conclusion**　Assembly, sequence analysis and phylogenetic evolution of *B. chinensis* chloroplast genome provides an insight into studies on both genetic structure and genetic diversity.

**Key words:** *Belamcanda chinensis* (L.) Redouté; chloroplast genome; structure; sequence analysis; Iridaceae

射干 *Belamcanda chinensis* (L.) Redouté 为鸢尾科（Iridaceae）射干属多年生草本植物，具结节状的根状茎，叶片剑形，花橙红色，花瓣上散生紫褐色斑点，分布于我国的吉林、山东、安徽、江苏、浙江和福建等 20 多个省份[1]。射干分布广泛，是一种十分常见的药用植物，有着悠久的栽培和药用历史；射干的主要药用部位为其根状茎，它富含酚类化合物，尤其是类黄酮和异黄酮类物质，如射干

**收稿日期：**2020-12-03

**基金项目：**台州市 211 人才工程经费资助（2012 年度）

**作者简介：**蒋　明（1973—），男，浙江嵊州人，博士，教授，硕士生导师，研究方向为植物基因组学、植物逆境生物学及其分子调控。E-mail: jiangming1973@139.com

苷、白射干素、鸢尾黄素和野鸢尾苷等[2]。射干具有清热解毒、散结消炎和利咽消肿等功效，可用于治疗扁桃腺炎、喉痹咽痛和腰痛等[3-5]。射干叶形优美、花色亮丽、花期较长，具有较好的观赏价值，常用于盆栽、庭院美化或药草园片植等，是一种观赏和药用兼用植物[6]。近年来，有关射干的研究主要集中在栽培技术、组培快繁、病虫害防治、化学成分、药理作用和转录组等方面[7-11]。

叶绿体是绿色植物进行光合作用和能量转换的细胞器，它将太阳能转换为化学能，用于物质积累和植物的生长发育[12]。叶绿体在植物逆境防御中也起着十分重要的作用，在不良环境下，产生大量的活性氧（reactive oxygen species，ROS），ROS 作为一种调控基因表达水平的信号分子，促使植物产生新的适应[13-14]。叶绿体基因组以双链环状形式存在于叶绿体中，被子植物的叶绿体基因组大小为 120～170 kb，测序和组装比核基因组容易[15]。随着测序技术的快速发展及测序费用的降低，越来越多的叶绿体基因组得以注释，目前已有 1000 多种植物完成叶绿体基因组的测序，它们在遗传多样性、基因组进化、基因水平转移、系统发育分析和群体遗传学等方面得到了应用[16-18]。目前，有关射干叶绿体基因组的测序、组装和注释等研究未见报道。本研究以射干叶片 DNA 为材料，在利用高通量测序的基础上拼接叶绿体基因组，明确其序列特征、基因组结构及与近缘特种间的系统发育关系，为后续开展遗传多样性和群体遗传学研究奠定基础。

## 1 材料与仪器

### 1.1 材料

射干叶片采自浙江舟山东福山岛，伴生植物有山菅 *Dianella ensifolia* (L.) DC.、小茄 *Lysimachia japonica* Thunb.、蔍状景天 *Sedum polytrichoides* Hemsl.和滨海前胡 *Peucedanum japonicum* Thunb.等。采集健康叶片，装入样品袋后带回实验室备用。

### 1.2 仪器

艾本德 Eppendorf 移液枪；超净工作台（苏州安泰空气技术有限公司）；ThinkPad P52 移动工作站；伯乐 C1000 型 PCR 仪（Bio-Rad 公司，美国）；Covaris 超声波 DNA 破碎仪（Chromatin Shearing，美国）；Illumina HiSeq X Ten 测序仪；北京六一 DYY-12 型电泳仪及电泳槽（北京市六一仪器厂）；伯乐 Gel Doc XR+凝胶成像系统（Bio-Rad 公司，美国）。

## 2 方法

### 2.1 DNA 的提取和文库构建

在无菌研砵中加入适量液氮，用研棒将叶片磨成细粉末，再利用十六烷基三甲基溴化铵（Cetyltrimethylammonium bromide，CTAB）方法提取基因组 DNA，经电泳检测后用于构建文库。用超声波破碎仪将基因组 DNA 片段化，经末端修复、添加 A 尾、两端加测序接头、产物纯化及 PCR 扩增等过程，完成文库的构建。

### 2.2 高通量测序

采用双端（paired-end，PE）策略，用 Illumina HiSeq X Ten 高通量测序仪进行测序，读长为 2×150 bp，共获得 3.55 G 原始数据。利用 NGS QC Toolkit v2.3.3 对原始数据进行过滤，去除接头和低质量的序列[19]。最终共得到 11 814 172 条 clean reads，$Q_{20}$ 值达 97.53%，$Q_{30}$ 为 92.73%，序列质量较高，可用于后续的拼接和注释。

### 2.3 叶绿体基因组的拼接和注释

叶绿体基因组的拼接在 ThinkPad P52 移动工作站上进行，拼接采用 NOVOPlasty 程序[20]。利用在线工具 DOGMA（dual organellar genoMe annotator）对序列进行基因注释，网址为 http://dogma.ccbb.utexas.edu/，起始和终止密码子通过手工方式进行调整[21]。tRNA 用 tRNAscan-SE（http:// www.lowelab.ucsc.edu/tRNAscan-SE/）和 ARAGORN 预测[22-23]。OGDRAW（organellar genome draw）用于生成叶绿体基因组圈图，网址为 http://www. ogdraw. mpimp-golm.mpg.de[24]。

### 2.4 边界序列的 PCR 克隆和测序验证

叶绿体基因组的 4 个边界采用 PCR 方法进行鉴定，根据拼接的草图，共设计了 4 对引物，分别是 YGUP1：5’-GGGCGAACCAAAAAGAATGATG-3’、YGDN1：5’-CTTTTGTAGCCAATCATTTATCGGG-3’；YGUP2：5’-GGTTATGGAAGAAGGAACCGAGAA-3’、YGDN2：5’-GCTATTTCCTCTGCTTGTATTGGT-3’；YGUP3：5’-CTATTTTACGTCTTTGCGCGC-3’、YGDN3：5’-CCGAGCTCGGGTTATGGAAG-3’；YGUP4：5’-CTGTAGACCCACGGAAAAATGT-3’、YGDN4：5’-GGTAGAGCCGGATCGAAGT-3’。

Eppendorf 管中，依次加入 15 μL $ddH_2O$、2.0 μL 10×缓冲液、0.4 μL dNTPs（10 mmol/L）、0.3 μL 的上游引物（20 μmol/L）、0.3 μL 下游引物

（20 μmol/L）、1.0 μL DNA 模板（50 ng/μL）和 0.5 μL *Taq* DNA 聚合酶（2 U/μL）。PCR 反应在伯乐 C1000 型 PCR 仪上进行，程序为：94 ℃变性 5 min；32 个循环中，94 ℃、30 s，54.3 ℃、45 s，72 ℃、100 s，共 32 个循环。PCR 产物经电泳、割胶、回收和纯化后，将其与 p-GEM T-easy 载体（Promega）连接，置于 4 ℃过夜。把连接产物导入大肠杆菌 DH5α 感受态细胞，经 PCR 检测后，各取 3 份菌液测序。

### 2.5 系统发育分析

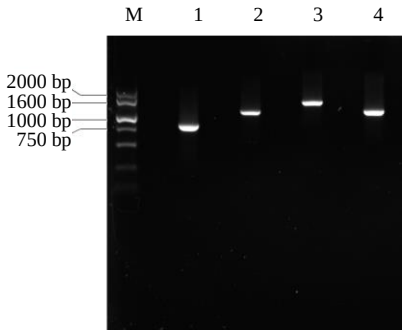
从 NCBI 数据库中下载 6 条叶绿体基因组的序列，它们分别来自溪荪 *Iris sanguinea* Donn ex Horn.（KT626943）、龙须菜 *Asparagus schoberioides* Kunth（KX790361）、蔥第百子莲 *Agapanthus coddii* F. M. Leight（KX790363）、惠普尔丝兰 *Hesperoyucca whipple* (Torr.) Trel.、鹿葱 *Lycoris squamigera* Maxim.（NC_040164）和马褂木 *Liriodendron chinense* (Hemsl.) Sargent.（KU170538）。其中的溪荪为鸢尾科植物；蔥第百子莲和鹿葱属石蒜科（Amaryllidaceae）；龙须菜和惠普尔丝兰为百合科（Liliaceae）植物；而外类群马褂木隶属木兰科（Magnoliaceae）。

利用 Geneious 11.1.5 内置的 Find repeats 工具预测叶绿体基因组的反向重复（inverted repeat，IR）序列、大单拷贝区（large single copy，LSC）和小单拷贝区（small single copy，SSC），并用 EXCEL 计算各自的 GC 值。Geneious 11.1.5 软件中的 MAFFT 7.388 程序用于叶绿体基因组的多重比对，导出的比对序列用 PhyML-SMS（Smart Model Selection in PhyML）在线工具获得最佳替代模型，最后生成最大似然（maximum likelihood，ML）树[25-26]。利用 Mega X 软件构建最大简约（maximum parsimony）树，自举检测值为 1000。

## 3 结果与分析

### 3.1 叶绿体基因组的边界验证

利用 NOVOPlasty 对射干叶片 DNA 的 Clean reads 进行拼接，并通过 PCR 手段对 4 个边界进行克隆和测序验证。电泳结果表明，4 对引物均能扩增出单 1 条带，PCR 产物分别位于 800、1300、1800、1300 bp 处，大小与预期一致（图 1）。序列测定结果表明，4 条序列的长度分别为 812、1328、1859、1325 bp，碱基排列与组装完成的叶绿体基因组边界序列完全一致。




M-Marker 1～4-LSC/IRb、IRB/SSC、SSC/IRa 和 IRa/LSC 边界序列

M-Marker 1—4-boundary sequences of LSC/IRb、IRb/SSC、SSC/IRa and IRa/LSC

**图 1 PCR 产物的凝胶电泳**

**Fig. 1 Gel electrophoresis of PCR products**

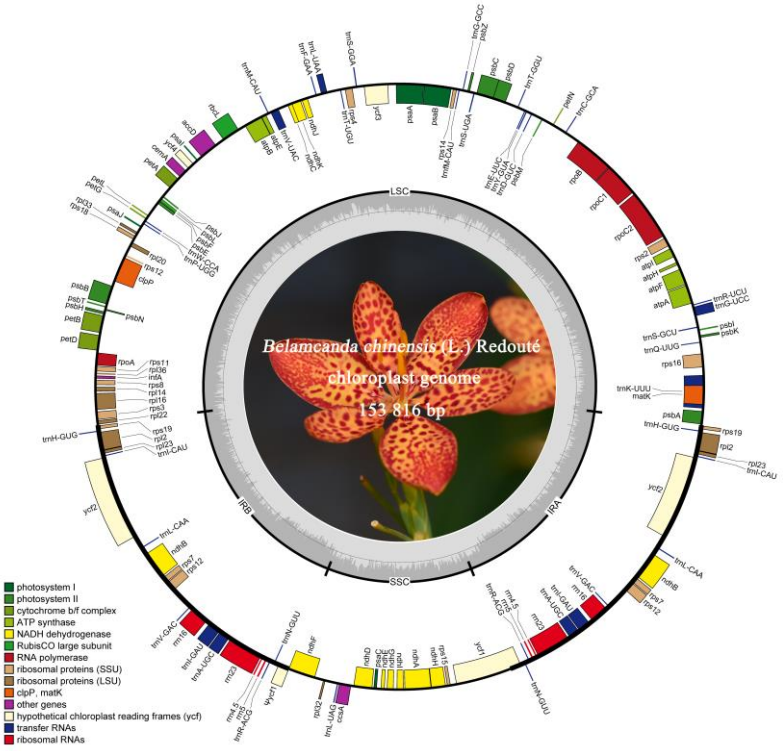
### 3.2 叶绿体基因组的特点

射干叶绿体基因组全长为 153 816 bp，具有一个典型的四分体结构，即分别由 LSC、SSC、IRa 和 IRb 4 个部分组成（图 2）。LSC 与 SSC 的大小为 83 143 bp 和 18 245 bp，IRa 和 IRb 长度均为 26 214 bp（表 1）。利用 Excel 计算 GC 值，结果表明，反向重复区的 GC 值最大，达 43.0%，LSC 次之，为 36.0%，而 SSC 的 GC 值最小，仅 31.4%；射干整个叶绿体基因组的 GC 值为 37.8%。

### 3.3 叶绿体基因的组成和特点

射干叶绿体基因组上共有 133 个基因，包括 38 个转运 RNA（tRNA）、8 个核糖体 RNA（rRNA）、14 个核糖体蛋白小亚基基因、11 个核糖体蛋白大亚基基因、4 个 RNA 聚合酶基因、12 个 NADH 脱氢酶亚基基因、20 个光系统 I/光系统 II 亚基基因、6 个细胞色素 b/f 复合物亚基基因和 6 个 ATP 合成酶亚基基因等。另有 7 个未知功能基因，它们是 *ycf1*、*ycf2*、*ycf3* 和 *ycf4*，除 *ycf3* 外，其它各有 2 份拷贝。

rRNA 基因中，*4.5rrn*、*5rrn*、*16rrn* 和 *23rrn* 各有 2 份拷贝，分别位于 2 个反向重复区域，4 种 rRNA 的长度分别为 103、121、1491、2810 bp。tRNA 中，*trnA-UGC*、*trnI-GAU*、*trnH-GUG*、*trnI-CAU*、*trnL-CAA*、*trnR-ACG*、*trnN-GUU* 和 *trnV-GAC* 各有 2 份拷贝，其余 tRNA 均只有一个。具有 2 份拷贝的基因还有 *rps7*、*rpl2*、*rpl23*、*ndhB*、*ycf4* 和 *ycf2* 基因等，*ycf1* 也有 2 份拷贝，但其中 1 个为假基因（表 2）。




内圈深色部分为 GC 含量

The darker gray area in inner circle is GC contents

**图 2　射干叶绿体基因组图谱**

**Fig. 2　Map of *B. chinensis* chloroplast genome**

**表 1　射干叶绿体基因组的碱基组成**

**Table 1　Base composition of chloroplast genome in *B. chinensis***

| 区域 | A/% | T/% | C/% | G/% | 长度/bp | GC/% |
|---|---|---|---|---|---|---|
| LSC | 31.3 | 32.7 | 18.5 | 17.5 | 83 143 | 36.0 |
| SSC | 34.2 | 34.4 | 16.7 | 14.8 | 18 245 | 31.4 |
| IRa/IRb | 28.5 | 28.4 | 22.3 | 20.8 | 26 214 | 43.0 |
| 合计 | 30.7 | 31.5 | 19.3 | 18.5 | 153 816 | 37.8 |

射干叶绿体基因组中，大部分基因没有内含子，少部分具 1～2 个内含子。*ycf3*、*rps12* 和 *clpP* 基因有 2 个内含子，*trnA-UGC*、*trnG-UCC*、*trnI-GAU*、*trnV-UAC*、*rpl2*、*rpl16*、*rps16*、*trnK-UUU*、*trnL-UAA*、*rpoC1*、*ndhA*、*ndhB*、*atpF* 和 *petB* 具 1 个内含子（表 2）。

### 3.4　基因组特征比较分析

从 NCBI 下载了溪荪、龙须菜、蔫第百子莲、惠普尔丝兰、马褂木和鹿葱 6 种植物的叶绿体基因组，它们的全长分别为 152 408、156 875、157 055、157 832、159 429 和 158 459 bp。马褂木叶绿体基因组的 GC 值最大，为 39.2%，溪荪次之，为 38.0%，蔫第百子莲的 GC 值最小，仅 37.5%（表 3）。马褂木的 LSC 和 SSC 最长，分别为 87 766、18 997 bp，但它的 IR 较短，为 26 333 bp。蔫第百子莲的 IR 最长，长度为 26 869 bp，鹿葱次之，为 26 764 bp，而溪荪的 IR 最短，仅 26 026 bp。

射干与其他 6 种植物的 LSC/IRb、IRb/SSC、SSC/IRa 和 IRa/LSC 边界及基因分布如图 3 所示，基因在边界的排列情况基本相同。除马褂木外，LSC/IRb 边界两侧分布有 *rpl22* 和 *rps19* 基因，而马褂木 LSC/IRb 两侧为 *rps19* 和 *rpl2* 基因。*ycf1* 和 *ndhF* 基因位于 IRb/SSC 的边界，7 种植物在该位置的 *ycf1* 基因全长为 899～1112 bp，显著短于正常的 *ycf1* 基因，它们均为假基因。另一个 *ycf1* 位于 SSC/IRa 交界处，长度为 5276～5489 bp，均为正常基因。马褂木 IRa/LSC 边界两侧分布有 *rpl2* 和 *trnH* 基因，而其他 6 种植物在该处为 *rps19* 和 *psbA* 基因。

表 2 射干叶绿体基因组的基因

**Table 2 Genes harbored in chloroplast genome of *B. chinensis***

| 基因类别 | 基因名称 | 数量 |
|---|---|---|
| 转运 RNA | *trnA-UGC*（×2）*、*trnC-GCA*、*trnD-GUC*、*trnE-UUC*、*trnF-GAA*、*trnfM-CAU*、*trnG-GCC*、*trnG-UCC**、*trnH-GUG*（×2）、*trnI-CAU*（×2）、*trnI-GAU*（×2）*、*trnK-UUU**、*trnL-CAA*（×2）、*trnL-UAA**、*trnL-UAG*、*trnM-CAU*、*trnN-GUU*（×2）、*trnP-UGG*、*trnQ-UUG*、*trnR-ACG*（×2）、*trnR-UCU*、*trnS-GCU*、*trnS-GGA*、*trnS-UGA*、*trnT-GGU*、*trnT-UGU*、*trnV-GAC*（×2）、*trnV-UAC**、*trnW-CCA*、*trnY-GUA* | 38 |
| 核糖体 RNA | *rrn4.5*（×2）、*rrn5*（×2）、*rrn16*（×2）、*rrn23*（×2） | 8 |
| 核糖体蛋白小亚基 | *rps2*、*rps3*、*rps4*、*rps7*（×2）、*rps8*、*rps11*、*rps12***、*rps14*、*rps15*、*rps16**、*rps18*、*rps19*（×2） | 14 |
| 核糖体蛋白大亚基 | *rpl2*（×2）*、*rpl14*、*rpl16**、*rpl20*、*rpl22*、*rpl23*（×2）、*rpl32*、*rpl33*、*rpl36* | 11 |
| RNA 聚合酶 | *rpoA*、*rpoB*、*rpoC1**、*rpoC2* | 4 |
| NADH 脱氢酶亚基 | *ndhA**、*ndhB*（×2）*、*ndhC*、*ndhD*、*ndhE*、*ndhF*、*ndhG*、*ndhH*、*ndhI*、*ndhJ*、*ndhK* | 12 |
| 光系统 I 亚基 | *psaA*、*psaB*、*psaC*、*psaI*、*psaJ* | 5 |
| 光系统 II 亚基 | *psbA*、*psbB*、*psbC*、*psbD*、*psbE*、*psbF*、*psbH*、*psbI*、*psbJ*、*psbK*、*psbL*、*psbM*、*psbN*、*psbT*、*psbZ* | 15 |
| 细胞色素 b/f 复合物亚基 | *petA*、*petB**、*petD*、*petG*、*petL*、*petN* | 6 |
| ATP 合成酶亚基 | *atpA*、*atpB*、*atpE*、*atpF**、*atpH*、*atpI* | 6 |
| Rubisco 大亚基 | *rbcL* | 1 |
| 成熟酶 | *matK* | 1 |
| 蛋白酶 | *clpP*** | 1 |
| 被膜蛋白 | *cemA* | 1 |
| 乙酰 CoA 羧化酶 | *accD* | 1 |
| 细胞色素 C 合成酶 | *ccsA* | 1 |
| 翻译起始因子 | *infA* | 1 |
| 未知功能蛋白 | *ycf1*、Ψ*ycf1*、*ycf2*（×2）、*ycf3***、*ycf4*（×2） | 7 |
| 合计 | | 133 |

×2-2 份拷贝 Ψ-假基因 *-一个内含子 **-2 个内含子

×2-2 copies Ψ-pseudogene *-one intron **-2 introns

表 3 7 种植物叶绿体基因组的特征

**Table 3 Chloroplast genome characteristics of seven plants**

| 植物名称 | 叶绿体基因组长度/bp | GC 值/% | LSC 长度/bp | SSC 长度/bp | IR 长度/bp |
|---|---|---|---|---|---|
| 射干 | 153 816 | 37.8 | 83 143 | 18 245 | 26 214 |
| 溪荪 | 152 408 | 38.0 | 82 377 | 18 016 | 26 026 |
| 龙须菜 | 156 875 | 37.6 | 85 118 | 18 645 | 26 556 |
| 蔹第百子莲 | 157 055 | 37.5 | 86 648 | 18 113 | 26 869 |
| 马褂木 | 159 429 | 39.2 | 87 766 | 18 997 | 26 333 |
| 惠普尔丝兰 | 157 832 | 37.8 | 86 170 | 18 228 | 26 717 |
| 鹿葱 | 158 459 | 37.8 | 86 431 | 18 500 | 26 764 |

### 3.5 系统发育分析

利用在线工具 PhyML-SMS 获得最佳分子进化模型，结果表明，7 条叶绿体基因组的最佳模型为 GTR＋G＋I，该模型的赤池信息标准（Akaike information criterion，AIC）与贝叶斯信息标准（Bayesian information criterion，BIC）分别为 820 620.543 72 和 820 831.677 20，与其他模型相比，它们的数值最低。以马褂木叶绿体基因组为外类群，构建系统发育树（图 4）。7 种植物叶绿体基因组在系统发育树上可分为 4 组，石蒜科的鹿葱和蔹第百子莲聚于组 I，支持率为 100%；百合科的龙须菜和惠普尔丝兰聚于组 II，支持率达 100%；鸢尾科的溪荪和射干聚于组 III，支持率也为 100%；而外类群马褂木单独处于分支 IV。同时，利用 Mega X 构建了最大简约树，结果与最大似然树基本一致，除龙须菜和惠普尔丝兰的支持率为 99%外，其余均达 100%（图 4）。

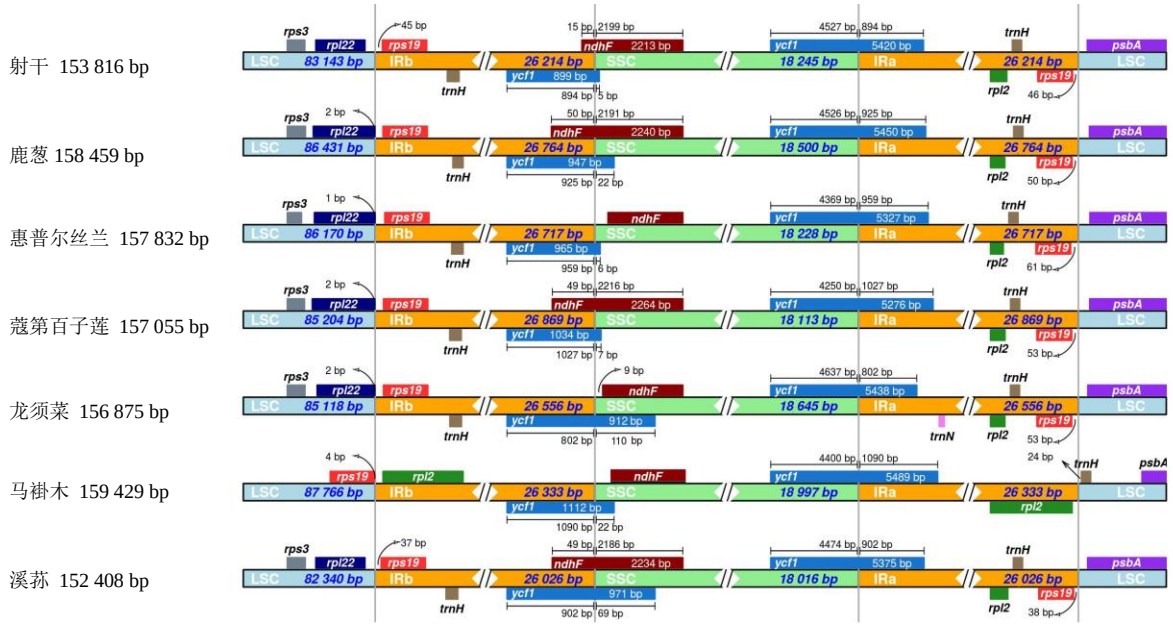



图 3　射干叶绿体基因组的反向重复区域/单拷贝区域边界

**Fig. 3　Inverted repeat region/single-copy region borders of *B. chinensis* chloroplast genome**

图 4　基于叶绿体全基因组构建的系统发育树

**Fig. 4　Phylogenetic tree constructed based on complete chloroplast genomes**

## 4　讨论

被子植物叶绿体基因组的结构通常十分保守，为双链环状，具 2 个反向重复序列 IRa 和 IRb，它们将 LSC 和 SSC 隔开，最终形成四分体结构[27]。叶绿体基因组的长度也十分保守，大部分陆生植物的叶绿体基因组大小在 135～160 kb [28]。禾本科（Gramineae）植物剪股颖 *Agrostis stolonifera* L.的叶绿体基因组大小为 136 584 bp[29]；百合科洋葱 *Allium cepa* L.的叶绿体基因大小为 153 538 bp[30]。有些植物的叶绿体基因组较小，如药用植物木贼麻黄 *Ephedra equisetina* Bge.仅 109 518 bp[31]；而天竺葵 *Pelargonium hortorum* Bailey 的叶绿体基因组大小达 217～942 bp，在目前已完成测序的陆生植物中最大[32]。本研究以射干为材料，在高通量测序的基础上，对其叶绿体基因组进行了组装，经 PCR 克隆和测序验证，得到射干叶绿体基因组的全长序列，它的长度为 153 816 bp。

叶绿体基因组的基因组成十分保守，通常包含 130 个左右的基因，这些基因的功能涉及光合作用、转录和翻译等[33]。叶绿体基因组通常有 114 种基因，包括 4 种 rRNA 基因、30 种 tRNA 基因和 80 种蛋白质编码基因[34]。在一些寄生或半寄生植物中，基因丢失现象十分普遍。山毛榉寄生 *Epifagus virginiana* (L.) W. P. C. Barton 的叶绿体基因组大小仅 70 kb，大部分基因丢失，仅剩下 42 个，与光合作用和叶绿体呼吸相关的基因全部缺失[35]。在半寄生植物广寄生 *Taxillus chinensis* (DC) Danser 和桑寄生 *T. sutchuenensis* (Lecomte) Danser 叶绿体基因组中，仅注释到 106 个基因，包括 66 个蛋白编码基因、28 个 tRNA、8 个 rRNA 和 4 个假基因[36]。在射干叶绿体基因组中，共有 4 种 rRNA 基因，它们各有 2 份拷贝，分布在不同的 IR 区域，另有 30 种 tRNA 基因，其中的 *trnA-UGC*、*trnH-GUG* 和 *trnI-CAU* 等各有 2 份拷贝。

植物在进化过程中，叶绿体基因组 LSC/IRb、IRb/SSC、SSC/IRa 和 IRa/LSC 边界常发生扩张或收缩事件，使叶绿体基因组大小出现一定的差异，甚至发生假基因化。比较鲁桑 *Morus multicaulis* Perr.、蒙桑 *M. mongolica* Schneid.、印度桑 *M. indica* L.和川桑 *M. notabilis* Schneid.的叶绿体基因组发现，它们的 SSC/IRb 交界处均存在 1 个 *ycf1*

假基因；在鼠尾草 *Salvia japonica* Thunb. SSC/IRb 处的 *ycf1* 也是一个假基因。本研究中，也发现类似的现象，在射干叶绿体基因组中共有 2 份 *ycf1* 拷贝，位于 SSC/IRb 边界的 *ycf1* 为假基因，而另一份正常。叶绿体基因组可用于植物的进化分析和亲缘关系鉴定，目前已应用于不同分类级别的系统进化分析。张慧等[37]利用 15 个野芝麻亚科（Lamioideae）65 个共有叶绿体蛋白序列构建最大似然法树（maximum likelihood，ML），发现益母草 *Leonurus artemisia* (Laur.) S. Y. Hu 和水苏属 *Stachys* Linn.的亲缘关系较近，大部分节点的支持率达 100%。射干与其他 6 种植物叶绿体基因组的系统发育分析结果表明，射干与溪荪的关系最为接近，支持率达 100%。

射干种质资源十分丰富，但不同种源药材的质量参差不齐。叶绿体基因组的组装和序列分析，为开发高分辨率的遗传标记提供了依据，也为后续开展该植物的群体遗传学和遗传多样性研究奠定了基础。

**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 [第 16(1)卷] [M]. 北京: 科学出版社, 1985: 26.

[2] 张婧涵, 张晓瑞, 李国信, 等. 线性回归色谱峰定位法在射干药材多组分同时测定中的应用 [J]. 药物分析杂志, 2014, 34(7): 1149-1155.

[3] 廉伟伟, 熊维政, 贾玉梅, 等. 射干及其方剂临床应用探讨 [J]. 中医学报, 2018, 33(11): 2184-2190.

[4] 王迪, 辛旭阳, 尤献民, 等. 射干效用演变探析 [J]. 辽宁中医药大学学报, 2015, 17(9): 67-69.

[5] 展锐, 焦正花, 王红丽, 等. 射干的药理作用研究概况 [J]. 甘肃中医, 2011, 24(1): 78-80.

[6] 王晓荣, 朱序弼, 刘金. 耐寒耐旱射干 西部开发佳卉 [J]. 中国花卉盆景, 2002(9): 33.

[7] 王继东, 卢杉. 延庆区射干栽培技术 [J]. 农业科技通讯, 2019(1): 192-193.

[8] 燕晨宇, 张翔, 秦民坚. 射干真叶、花蕾诱导愈伤组织的研究 [J]. 中国野生植物资源, 2018, 37(5): 16-19.

[9] 莫雪梅, 焦明姚, 文永刚, 等. 射干叶枯病防治药剂筛选试验初探 [J]. 耕作与栽培, 2017(4): 36-37.

[10] Li J Y, Ni G, Li L, *et al*. New iridal-type triterpenoid derivatives with cytotoxic activities from *Belamcanda chinensis* [J]. *Bioorg Chem*, 2019, 83: 20-28.

[11] Tian M, Zhang X, Zhu Y, *et al*. Global transcriptome analyses reveal differentially expressed genes of six organs and putative genes involved in (iso) flavonoid biosynthesis in *Belamcanda chinensis* [J]. *Front Plant Sci*, 2018, 9: 1160.

[12] Suzuki N, Koussevitzky S, Mittler R, *et al*. ROS and redox signalling in the response of plants to abiotic stress [J]. *Plant Cell Environ*, 2012, 35(2): 259-270.

[13] Liu X M, Zhou Y L, Xiao J W, *et al*. Effects of chilling on the structure, function and development of chloroplasts [J]. *Front Plant Sci*, 2018, 9: 1715.

[14] Miller G, Shulaev V, Mittler R. Reactive oxygen signaling and abiotic stress [J]. *Physiol Plant*, 2008, 133(3): 481-489.

[15] Du Y P, Bi Y, Yang F P, *et al*. Complete chloroplast genome sequences of Lilium: Insights into evolutionary dynamics and phylogenetic analyses [J]. *Sci Rep*, 2017, 7(1): 5751.

[16] Xiong A S, Peng R H, Zhuang J, *et al*. Gene duplication, transfer, and evolution in the chloroplast genome [J]. *Biotechnol Adv*, 2009, 27(4): 340-347.

[17] Daniell H, Lin C S, Yu M, *et al*. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering [J]. *Genome Biol*, 2016, 17(1): 134.

[18] Twyford A D, Ness R W. Strategies for complete plastid genome sequencing [J]. *Mol Ecol Resour*, 2017, 17(5): 858-868.

[19] Patel R K, Jain M. NGS QC Toolkit: A toolkit for quality control of next generation sequencing data [J]. *PLoS One*, 2012, 7(2): e30619.

[20] Dierckxsens N, Mardulyn P, Smits G. NOVOPlasty: de novo assembly of organelle genomes from whole genome data [J]. *Nucleic Acids Res*, 2017, 45(4): e18.

[21] Wyman S K, Jansen R K, Boore J L. Automatic annotation of organellar genomes with DOGMA [J]. *Bioinformatics*, 2004, 20(17): 3252-3255.

[22] 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.

[23] Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences [J]. *Nucleic Acids Res*, 2004, 32(1): 11-16.

[24] Lohse M, Drechsel O, Kahlau S, *et al*. OrganellarGenomeDRAW: A suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets [J]. *Nucleic Acids Res*, 2013, 41: W575-W581.

[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] Bendich A J. Circular chloroplast chromosomes: The grand illusion [J]. *Plant Cell*, 2004, 16(7): 1661-1666.

[28] Raubeson L A, Peery R, Chumley T W, *et al*. Comparative chloroplast genomics: Analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus* [J]. *BMC Genomics*, 2007, 8: 174.

[29] Saski C, Lee S B, Fjellheim S, *et al*. Complete chloroplast genome sequences of Hordeum vulgare, *Sorghum bicolor* and *Agrostis stolonifera*, and comparative analyses with other grass genomes [J]. *Theor Appl Genet*, 2007, 115(4): 571-590.

[30] von Kohn C, Kiełkowska A, Havey M J. Sequencing and annotation of the chloroplast DNAs and identification of polymorphisms distinguishing normal male-fertile and male-sterile cytoplasms of onion [J]. *Genome*, 2013, 56(12): 737-742.

[31] Wu C S, Lai Y T, Lin C P, *et al*. Evolution of reduced and compact chloroplast genomes (cpDNAs) in gnetophytes: Selection toward a lower-cost strategy [J]. *Mol Phylogenet Evol*, 2009, 52(1): 115-124.

[32] Chumley T W, Palmer J D, Mower J P, *et al*. The complete chloroplast genome sequence of *Pelargonium x hortorum*: Organization and evolution of the largest and most highly rearranged chloroplast genome of land plants [J]. *Mol Biol Evol*, 2006, 23(11): 2175-2190.

[33] Daniell H, Lee S B, Grevich J, *et al*. Complete chloroplast genome sequences of *Solanum bulbocastanum*, *Solanum lycopersicum* and comparative analyses with other *Solanaceae genomes* [J]. *Theor Appl Genet*, 2006, 112(8): 1503-1518.

[34] 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.

[35] 李巧丽, 延娜, 宋琼, 等. 鲁桑叶绿体基因组序列及特征分析 [J]. 植物学报, 2018, 53(1): 94-103.

[36] 何懿菡, 韩立敏, 刘玉萍, 等. 鼠尾草叶绿体基因组序列分析 [J]. 植物研究, 2017, 37(4): 572-578.

[37] 张慧, 何帅兵, 孔繁德, 等. 益母草叶绿体基因组序列与系统进化位置分析 [J]. 中医药信息, 2018, 35(4): 21-27.

[责任编辑 时圣明]