

# 相近石韦叶绿体基因组学分析

杨楚虹<sup>1</sup>, 崔英贤<sup>2</sup>, 聂丽萍<sup>2</sup>, 姚辉<sup>2</sup>, 王瑀<sup>2\*</sup>, 刘霞<sup>1\*</sup>

(1. 武汉理工大学 化学化工与生命科学学院, 武汉 430070;

2. 中国医学科学院 & 北京协和医学院 药用植物研究所

国家中医药管理局中药资源保护重点实验室, 北京 100193)

**[摘要]** 目的: 测定相近石韦 *Pyrrosia assimilis* 叶绿体基因组, 分析其序列特征并探讨相近石韦本草基因组学研究。方法: 应用高通量测序技术对相近石韦进行了叶绿体全基因组测序, 利用生物信息学方法分析其结构特征和系统发育关系。结果: 相近石韦叶绿体基因组呈环形双链结构, 全长 154 964 bp, 鸟嘌呤和胞嘧啶 (GC) 总量 41.2%; 共注释到 131 个基因, 包括 88 个蛋白编码基因, 35 个转运 RNA (tRNA) 基因和 8 个核糖体 RNA (rRNA) 基因; 共检测到 43 个散在重复序列和 56 个简单重复序列 (SSR); 编码亮氨酸的密码子使用频率最高, 编码色氨酸的密码子数最少; 叶绿体基因组全局比对分析筛选出 5 个高变异区 (*psbA*, *rrn16*, *petA-psbJ*, *ndhC-trnM* 和 *psbM-petN*); 系统发育树显示相近石韦与波氏石韦 *P. bonii* 亲缘关系较近。结论: 相近石韦叶绿体基因组中非编码区变异高于编码区, 大单拷贝区 (LSC) 和小单拷贝区 (SSC) 变异大于反向重复区 (IR), 筛选出的 5 个高变区可作为石韦属物种鉴定的候选 DNA 条形码。相近石韦的叶绿体基因组学研究为其他石韦属药用植物在分子鉴定、遗传基因转化、抗性蛋白表达及次生代谢途径解析等方面的研究提供了参考。

**[关键词]** 本草基因组学; 相近石韦; 叶绿体基因组; 石韦属; 结构特征; DNA 条形码; 系统发育树

**[中图分类号]** R22; Q732; R28; Q34; R931 **[文献标识码]** A **[文章编号]** 1005-9903(2020)05-0123-09

**[doi]** 10.13422/j.cnki.syfx.20200747

**[网络出版地址]** <http://kns.cnki.net/kcms/detail/11.3495.R.20191216.1718.002.html>

**[网络出版时间]** 2019-12-17 09:04

## Chloroplast Genomic Analysis of *Pyrrosia assimilis*

YANG Chu-hong<sup>1</sup>, CUI Ying-xian<sup>2</sup>, NIE Li-ping<sup>2</sup>, YAO Hui<sup>2</sup>, WANG Yu<sup>2\*</sup>, LIU Xia<sup>1\*</sup>

(1. School of Chemistry, Chemical Engineering and Life Sciences, Wuhan University of Technology, Wuhan 430070, China; 2. Key Laboratory of Chinese Medicine Resources Conservation, National Administration of Traditional Chinese Medicine, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100193, China)

**[Abstract]** **Objective:** The complete chloroplast genome of *Pyrrosia assimilis* was sequenced, its sequence characteristics was analyzed and herbgonomics of *P. assimilis* was discussed. **Method:** Its complete chloroplast genome sequence was determined through high-throughput sequencing technology, and its structural characteristics and phylogenetic relationships were analyzed by bioinformatics. **Result:** The chloroplast genome of *P. assimilis* was a circular double-chain structure with a total length of 154 964 bp, and the total content of guanine and cytosine (GC) was 41.2%. A total of 131 genes were annotated, including 88 protein-coding genes, 35 transfer RNA (tRNA) genes and 8 ribosomal RNA (rRNA) genes. A total of 43 dispersed repetitive sequences and 56 simple sequence repeats (SSRs) were detected. The frequency of codon encoding leucine was the highest, while

**[收稿日期]** 20191104(029)

**[基金项目]** 国家“重大新药创制”科技重大专项(2018ZX09711001-008-007); 武汉理工大学专业学位研究生团队指导项目(201701)

**[第一作者]** 杨楚虹, 在读硕士, 从事中药资源研究, E-mail: ych0613fcw@163.com

**[通信作者]** \* 刘霞, 博士, 副教授, 从事中药资源和中药分子鉴定研究, E-mail: lrx1125@126.com;

\* 王瑀, 副主任技师, 从事中药资源研究, E-mail: ywang@implad.ac.cn

the number of codon encoding tryptophan was the lowest. Five highly divergent regions (*psbA*, *rrn16*, *petA-psbJ*, *ndhC-trnM*, and *psbM-petN*) were screened, phylogenetic analysis showed that *P. assimilis* was closely related to *P. bonii*. **Conclusion:** Comparative analysis of the complete chloroplast genome of *P. assimilis* reveals that non-coding regions exhibited a higher divergence than the coding regions, the large single copy region (LSC) and small single copy region (SSC) are more divergent than the reverse repeat region (IR), the selected five highly variable regions can be used as specific DNA barcodes for identification of *Pyrrosia* species. Study on the chloroplast genome of *P. assimilis* can provide a reference for the molecular identification, genetic transformation, expression of resistance protein and secondary metabolism pathway analysis of other *Pyrrosia* medicinal plants.

[ **Key words** ] herbgenomics; *Pyrrosia assimilis*; chloroplast genome; *Pyrrosia*; structural characteristics; DNA barcoding; phylogenetic tree

本草基因组学是利用组学技术研究中药基原物种的遗传信息及其调控网络,阐明中药防治人类疾病分子机制的学科,是从基因组水平研究中药及其对人体作用的前沿科学<sup>[1-2]</sup>。其研究包括结构基因组学、功能基因组学、表观基因组学和宏基因组学等领域<sup>[3]</sup>,而叶绿体基因组研究是本草基因组学的重要组成部分。叶绿体基因组作为植物特有的且相对独立于核基因组之外的基因组<sup>[4]</sup>,具有相对分子质量小、结构稳定、多拷贝、序列高度保守等特点<sup>[5]</sup>,与已有的 DNA 条形码片段相比,叶绿体基因组具有更多的遗传和变异信息,具有更强的分辨力<sup>[6]</sup>。随着高通量测序技术的不断发展,叶绿体基因组广泛应用于分子标记、物种鉴定、基因工程及系统发育等方面研究<sup>[7-10]</sup>。李滢等<sup>[11]</sup>对贝母属 4 个物种叶绿体基因组进行全局分析,筛选出 7 个基因间区可作为潜在的贝母属植物鉴定的特异性 DNA 条形码。李西文等<sup>[12]</sup>采用 454FLX 测序技术对药用厚朴的叶绿体基因组测序分析,成功区分了厚朴及同属近缘物种并推动了厚朴叶绿体基因工程的发展。LI 等<sup>[13]</sup>通过叶绿体分区工程化策略,首次在植物底盘中实现了紫杉醇关键中间体 5 $\alpha$ -羟基紫杉二烯的异源合成,这是叶绿体基因工程在烟草工程体系的一大突破。叶绿体基因组学的发展不仅丰富了叶绿体基因组数据库,而且对物种鉴定、遗传基因的转化、优良抗性植物的选育等方面的研究具有重要意义。

相近石韦是水龙骨科石韦属蕨类植物,也是药用石韦的近缘物种,具有一定的观赏价值和药用价值。石韦属物种丰富,分布范围较广,叶片形态相似,药用历史悠久,长期以来近缘物种混用现象较为严重<sup>[14]</sup>。据文献报道,相近石韦又名相异石韦,临床上常与药材石韦混淆<sup>[15]</sup>。传统形态学方法和 DNA 条形码序列在石韦物种分类与鉴定中存在困难,张雅琴等<sup>[16]</sup>基于 *psbA-trnH* 序列能准确鉴定石

韦药材及其混伪品,但无法区分石韦和有柄石韦。CAI 等<sup>[17]</sup>测定了波氏石韦叶绿体基因组并进行了结构解析和 ML 系统发育分析。WEI 等<sup>[18]</sup>利用 5 个叶绿体片段 (*rbcL*, *matK*, *psbA-trnH* 和 *rps4 + rps4-trnS*) 和一个单(低)拷贝核基因片段 (*LEAFY*) 对石韦属的 51 个物种进行了系统进化研究。已有学者对石韦属药用植物的化学成分和药理作用做了深入研究,但在分子鉴定方面的研究甚少,而且石韦属叶绿体基因组仅波氏石韦一条。因此,本研究采用高通量测序技术对相近石韦进行了叶绿体全基因组测序和分析,成功注释了 131 个叶绿体基因,为后续功能基因组学的研究提供基础信息;筛选出的 5 个高变区可作为特异性条形码,为石韦属种间鉴定提供参考;基于叶绿体全基因组分析系统进化关系,为追溯物种起源、种群结构及进化关系奠定基础。

## 1 材料与方法

**1.1 材料** DK-S24 型电热恒温水浴锅(上海森信实验仪器有限公司), 1-14 型高速离心机(德国 Sigma 公司), Nanodrop 2000 型超微量分光光度计(美国基因有限公司), SIM-F140 型制冰机(日本三洋公司), DYY-8C 型琼脂糖凝胶电泳仪(北京六一仪器厂), Universal Hood II 型凝胶成像仪(美国 Bio-Rad 公司), DL-CJ-IND 型超净工作台(哈尔滨市东联电子技术开发有限公司), PL203 型电子天平(瑞士梅特勒-托利多仪器公司), MM721NH1-PW 型微波炉(美的集团有限公司)。

植物基因组 DNA 提取试剂盒(DNeasy Plant Mini Kit, 德国 QIAGEN 公司, 批号 16004553), 2  $\times$  Taq 聚合酶链式反应(PCR) MasterMix 和 RNA 酶(北京艾德莱生物科技有限公司, 批号分别为 301447AX, 282232AX),  $\beta$ -巯基乙醇(北京博奥拓达科技有限公司, 批号 MC0674), 氨基丁三醇(Tris)-硼酸电泳缓冲液(5  $\times$  TBE, 自制), 硼酸(国药集团

化学试剂有限公司,批号 10004818),Tris(北京索莱宝科技有限公司,批号 1124Z072),乙二胺四乙酸(EDTA,BBI 生命科学有限公司,批号 20170110),溴化乙锭(EB)染料和 BM15000 DNA Marker(北京博迈德基因技术有限公司,批号分别为 18Ga2102,28455AC),6 × Loading Buffer(DNA 凝胶加样缓冲液,日本 Takara 公司,批号 AGZ0504A),引物由苏州金唯智生物科技有限公司合成,琼脂糖(西班牙 Biowest 公司,批号 111860),水为无菌去离子水。相近石韦 *Pyrrosia assimilis* 的新鲜叶片采自福建武夷山,样品经福建中医药大学黄泽豪教授和中国医学科学院药用植物研究所林余霖研究员鉴定,凭证标本保存于中国医学科学院药用植物研究所。

## 1.2 方法

### 1.2.1 总 DNA 的提取、叶绿体基因组测序及组装

取新鲜幼嫩的相近石韦叶片,液氮研磨,采用植物基因组 DNA 提取试剂盒提取总 DNA。将检测合格的总 DNA 构建约 500 bp 的测序文库,通过 Illumina HiSeq 2000 双末端测序策略进行测序,得到约 6 G 的原始测序数据。对低质量数据进行过滤,利用生物信息学技术对测序数据进行拼接和组装,拼接方法和应用软件同文献[19]一致。

设计引物采用 PCR 方法对反向重复区边界进行验证,序列为大单拷贝区-反向重复区 b(LSC-IRb)正向引物:5'-CTCGACTGCTGTTGATCCCT-3',反向引物:5'-CCAGCAAAGAACCATTCCGC-3';IRb-小单拷贝区(SSC)正向引物:5'-TGCGCATAGCAACCCCATTT-3',反向引物:5'-TCCTGTCTTTGTCCGAAGTTCA-3';SSC-IRa 正向引物:5'-GCTGCAAAAAGAGCGTCGAA-3',反向引物:5'-TCCCCCGACATGTTGTGAAG-3';IRa-LSC 正向引物:5'-CAGCAAAGAACCATTCCGCC-3',反向引物:5'-TGCGGCTTTCGTGTGATAGT-3'。

### 1.2.2 叶绿体全基因组注释及分析

相近石韦叶绿体全基因组的注释,简单重复序列(SSR),长重复序列及密码子使用情况的分析方法与文献[19]一致。

### 1.2.3 叶绿体全基因组 mVISTA 全局比对和系统发育树分析

以相近石韦为参考基因组,应用在线基因组比对工具 mVISTA<sup>[20]</sup>对水龙骨科波氏石韦 *P. bonii* (MH352390),网眼瓦韦 *Lepisorus clathratus* (KY419704)及日本水龙骨科 *Polypodiodes niponica* (MH319944)的叶绿体基因组序列进行全局比对。从美国国家生物技术信息中心(NCBI)数据库下载

了 16 条叶绿体基因组序列,包括水龙骨科波氏石韦 *P. bonii* (MH352390),网眼瓦韦 *L. clathratus* (KY419704),日本水龙骨科 *P. niponica* (MH319944),断线蕨 *Leptochilus hemionitideus* (MH319943),扇蕨 *Neocheiropteris palmatopedata* (NC\_044794),*Pecluma dulcis* (MK705755),肾蕨科长叶肾蕨 *Nephrolepis biserrata* (NC\_044683),鳞毛蕨科全缘贯众 *Cyrtomium falcatum* (NC\_028705),披针贯众 *Cyrtomium devexiscapulae* (NC\_028542),迷人鳞毛蕨 *Dryopteris decipiens* (KY427348),蹄盖蕨科宿蹄盖蕨 *Athyrium anisopterum* (KY419703),中华蹄盖蕨 *Athyrium sinense* (KY427333),叉蕨科 *Tectaria panamensis* (MK705757),岩蕨科大囊岩蕨 *Woodsia macrochlaena* (NC\_035864),耳羽岩蕨 *Woodsia polystichoides* (KY427359),桫欏科中华桫欏 *Alsophila costularis* (MH684489),采用最大似然法(maximum-likelihood,ML)构建系统发育树,以中华桫欏为外类群。

## 2 结果与分析

### 2.1 叶绿体基因组的基本特征

相近石韦叶绿体基因组(GenBank 登录号 MN617019)呈环状双链结构,由 2 个 IR,1 个 LSC 和 1 个 SSC 组成,见图 1。相近石韦叶绿体基因组全长 154 964 bp,其中 LSC 长 83 421 bp,SSC 长 21 589 bp,两者被长度为 24 977 bp 的 IRa 和 IRb 分隔。相近石韦叶绿体基因组鸟嘌呤和胞嘧啶(GC)总量 41.2%;IR,LSC 和 SSC 的 GC 总量分别为 44.9%,40.2%,37.1%。共注释了 131 个基因,包括 88 个蛋白编码基因,35 个 tRNA 基因和 8 个 rRNA 基因。其中,10 个蛋白编码基因(*ndhA*,*ndhB*,*rps16*,*rps12*,*atpF*,*rpoC1*,*petB*,*petD*,*rpl16* 和 *rpl2*)和 6 个 tRNA 基因(*trnA-UGC*,*trnG-UCC*,*trnI-GAU*,*trnT-UGU*,*trnV-UAC*,*trnL-UAA*)均有 1 个内含子,而 *clpP* 和 *ycf3* 基因含有 2 个内含子,见表 1。

### 2.2 叶绿体基因组重复序列分析

分析相近石韦叶绿体基因组中散在重复序列(重复单元长度 ≥ 30 bp),预测了 43 个长重复序列,包括 19 个正向重复(F),24 个回文重复(P),分别占总重复序列的 44.2%,55.8%,未发现反向重复(R)和互补重复(C)。其中,长度在 30~39 bp 的正向重复有 13 个,回文重复 18 个;长度在 40~49 bp 的正向重复有 15 个,回文重复 5 个;长度在 50~59 bp 的正向和回文重复均只含有 1 个。分布于 LSC 的重复序列最多,共有 31 个,占总比例的 72.1%;而 IRa,SSC 和 IRb

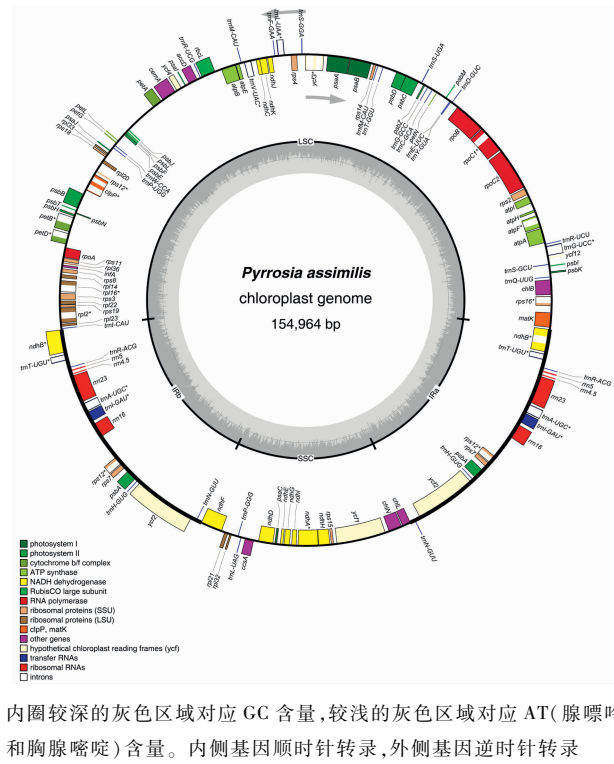


图 1 相近石韦叶绿体基因组的物理谱

Fig. 1 Physical spectrum of complete chloroplast genome of *Pyrrosia assimilis*

表 1 相近石韦叶绿体基因组内含子的长度和位置

Table 1 Lengths and locations of introns in chloroplast genome of *Pyrrosia assimilis*

基因	位置	外显子 I /bp	内含子 I /bp	外显子 II /bp	内含子 II /bp	外显子 III /bp
<i>atpF</i>	LSC	145	722	416		
<i>clpP</i>	LSC	71	699	292	544	240
<i>ndhA</i>	SSC	558	952	558		
<i>ndhB</i>	IR	40	832	600		700
<i>petB</i>	LSC	6	807	642		
<i>petD</i>	LSC	8	640	472		
<i>rp116</i>	LSC	9	762	405		
<i>rp12</i>	IR	397	723	437		
<i>rpoC1</i>	LSC	432	737	1 650		
<i>rps12</i>	IR	114	588	232		26
<i>rps16</i>	LSC	48	707	243		
<i>trnA-UGC</i>	IR	37	820	36		
<i>trnG-UCC</i>	LSC	23	902	48		
<i>trnI-GAU</i>	IR	36	960	36		
<i>trnT-UGU</i>	IR	33	543	40		
<i>trnV-UAC</i>	LSC	34	637	40		
<i>trnL-UAA</i>	SSC	34	584	51		
<i>ycf3</i>	LSC	125	749	229	642	162

分布较少, 分别有 7 个 (16.3%), 1 个 (2.3%) 和 4 个 (9.3%), 由此可见, 长重复序列在 4 个区域分布极不均衡。散在重复序列在基因组重排中起着重要作用, 为物种的系统发育研究提供了信息<sup>[21]</sup>。

SSR 在整个基因组中被称为微卫星, 是由 1 ~ 6 bp 的重复单元组成的核苷酸序列<sup>[22]</sup>。SSR 具有丰富的多态性, 是一种高效的分子标记, 广泛应用于物种鉴定以及群体和个体水平的遗传差异分析<sup>[23-25]</sup>。在相近石韦叶绿体基因组中, 共检索到 56 个符合条件的 SSR 位点, 包括 30 个单碱基重复单元, 9 个二碱基重复单元, 5 个三碱基重复单元, 10 个四碱基重复单元以及 2 个五碱基重复单元, 未发现六碱基重复, 见表 2。在相近石韦的微卫星类型中最丰富的是单碱基重复, 占 SSR 总数的 53.6%, 五碱基重复最少, 占 SSR 总数的 3.6%。在所有 SSR 中, 单碱基重复单元腺嘌呤 (A)/胸腺嘧啶 (T) (16 次) 和胞嘧啶 (C)/鸟嘌呤 (G) (14 次) 数量相当, 共有 8 处连续 A, 8 处连续 T, 8 处连续 C 及 6 处连续 G 的单碱基重复; 二碱基重复单元以 AT/AT (8 次) 为主, 占 SSR 总数的 14.3%。三碱基重复单元以 AAG/CTT (4 次) 为主, 占 SSR 总数的 7.1%。

表 2 相近石韦叶绿体基因组 SSR 类型、数量及比例

Table 2 Types, amounts and ratios of SSR in chloroplast genome of *Pyrrosia assimilis*

SSR 类型	碱基重复单元	数量/次	比例%
单碱基重复	A/T	16	28.6
	C/G	14	25.0
二碱基重复	AG/CT	1	1.8
	AT/AT	8	14.3
三碱基重复	AAG/CTT	4	7.1
	ACT/AGT	1	1.8
四碱基重复	AAAG/CTTT	1	1.8
	AAAT/ATTT	2	3.6
	AAGC/CTTG	1	1.8
	AATC/ATTG	1	1.8
五碱基重复	AATT/AATT	3	5.4
	AGAT/ATCT	2	3.6
	AACAG/CTGTT	1	1.8
	AGAGG/CCTCT	1	1.8

### 2.3 叶绿体基因组密码子使用分析 相对同义密码子使用度 (RSCU) 是特定密码子的使用频率和

预期频率之比<sup>[26]</sup>。通过 MEGA 10.0 软件计算相近石韦叶绿体基因组密码子使用率,共有 51 654 个密码子。结果发现编码亮氨酸(Leu)的密码子最多,为 5 223 个,而编码色氨酸(Trp)的密码子最少,为

706 个。除了蛋氨酸(Met)和 Trp,其他氨基酸的 RSCU 均 >1,密码子具有偏好性。其中密码子 AGA 的 RSCU 最高,为 1.61,密码子 GCG/GAC 的 RSCU 最低,为 0.67,见表 3。

表 3 相近石韦叶绿体基因组相对同义密码子使用度分析

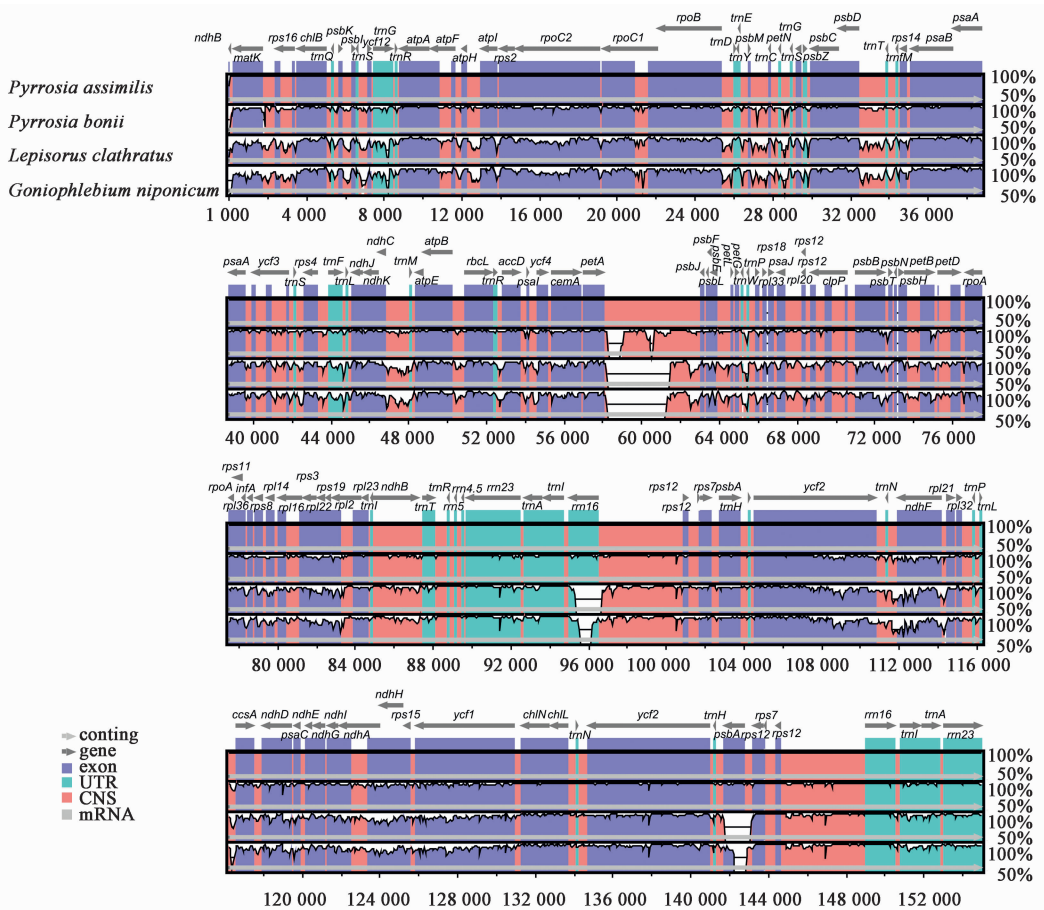
Table 3 Analysis of relative synonymous codon usage in chloroplast genome of *Pyrrosia assimilis*

氨基酸	密码子	密码子数/个	RSCU	氨基酸	密码子	密码子数/个	RSCU
苯丙氨酸(Phe)	UUU	1 733	1.18	酪氨酸(Tyr)	UAU	1 077	1.23
Phe	UUC	1 194	0.82	Tyr	UAC	681	0.77
Leu	UUA	1 021	1.17	终止(Stop)	UAA	1 053	1.20
Leu	UUG	1 033	1.19	Stop	UAG	720	0.82
Leu	CUU	1 032	1.19	Stop	UGA	863	0.98
Leu	CUC	736	0.85	Trp	UGG	706	1.00
Leu	CUA	811	0.93	谷氨酰胺(Gln)	CAA	1 059	1.32
Leu	CUG	590	0.68	Gln	CAG	543	0.68
异亮氨酸(Ile)	AUU	1 482	1.21	天冬酰胺(Asn)	AAU	1 523	1.29
Ile	AUC	957	0.78	Asn	AAC	843	0.71
Ile	AUA	1 231	1.01	赖氨酸(Lys)	AAA	1 759	1.28
Met	AUG	730	1.00	Lys	AAG	988	0.72
缬氨酸(Val)	GUU	802	1.30	天冬氨酸(Asp)	GAU	1 015	1.33
Val	GUC	452	0.73	Asp	GAC	507	0.67
Val	GUA	788	1.28	谷氨酸(Glu)	GAA	1 221	1.30
Val	GUG	423	0.69	Glu	GAG	658	0.70
丝氨酸(Ser)	UCU	1 111	1.27	脯氨酸(Pro)	CCU	732	1.06
Ser	UCC	945	1.08	Pro	CCC	757	1.09
Ser	UCA	966	1.11	Pro	CCA	779	1.13
Ser	UCG	686	0.79	Pro	CCG	500	0.72
Ser	AGU	880	1.01	精氨酸(Arg)	CGU	503	0.80
Ser	AGC	644	0.74	Arg	CGC	430	0.68
半胱氨酸(Cys)	UGU	689	1.08	Arg	CGA	683	1.08
Cys	UGC	583	0.92	Arg	CGG	499	0.79
苏氨酸(Thr)	ACU	862	1.29	Arg	AGA	1 014	1.61
Thr	ACC	696	1.05	Arg	AGG	651	1.03
Thr	ACA	639	0.96	组氨酸(His)	CAU	740	1.26
Thr	ACG	467	0.70	His	CAC	438	0.74
丙氨酸(Ala)	GCU	699	1.36	甘氨酸(Gly)	GGU	656	1.06
Ala	GCC	443	0.86	Gly	GGC	422	0.68
Ala	GCA	572	1.11	Gly	GGA	746	1.21
Ala	GCG	347	0.67	Gly	GGG	644	1.04

2.4 相近石韦的叶绿体基因组全局比对 以相近石韦为参考基因组,通过 mVISTA 软件对相近石韦、波氏石韦、网眼瓦韦和日本水龙骨科的 4 条叶绿体基

因组序列进行全局比对分析,见图 2。

由图 2 可知,非编码区(红色部分)变异较大,编码区(蓝色部分)相对保守,非编码区变异程度高



每个基因上的灰色箭头表示该基因的方向;不同颜色的波形图表示叶绿体基因组不同区域;Y轴表示各物种叶绿体基因组与参考序列 50% ~ 100% 相似性;X轴表示叶绿体基因组序列长度,单位为 bp(碱基对)

图 2 相近石韦及其他物种叶绿体基因组全局比对

Fig. 2 Global alignment of chloroplast genomes among *Pyrrosia assimilis* and other species

于编码区,LSC 和 SSC 变异程度大于 IR,同属物种的变异程度小于不同属物种。编码区中除了 *psbA*, *rrn16*, *matK*, *ndhF*, 和 *rpl22* 等基因变异程度较大,其他基因相对保守。基因区的差异小于基因间隔区,如 *petA-psbJ*, *ndhC-trnM* 和 *psbM-petN*, 差异大的区域可作为候选 DNA 条形码,为石韦属物种鉴定提供信息参考。

**2.5 系统进化分析** 从 NCBI 下载的 16 条叶绿体基因组序列及相近石韦叶绿体基因组构建 ML 树,以中华桫欏作为外类群,见图 3。系统发育树显示,绝大部分的节点均具有较高的支持率,水龙骨科 7 个物种聚为一个支,相近石韦与波氏石韦聚为一支且彼此分开,支持率 100%,表明相近石韦和波氏石韦亲缘关系更近,网眼瓦韦和日本水龙骨科分支成姐妹群关系,表明其系统发育关系较近,这与 CAI 等<sup>[17]</sup> 和 MIN 等<sup>[27]</sup> 的研究结果一致。水龙骨科物种与叉蕨科 Tectariaceae 聚为一支,支持率为 100%,表明水龙骨科与叉蕨科有着更近的系统发育关系。

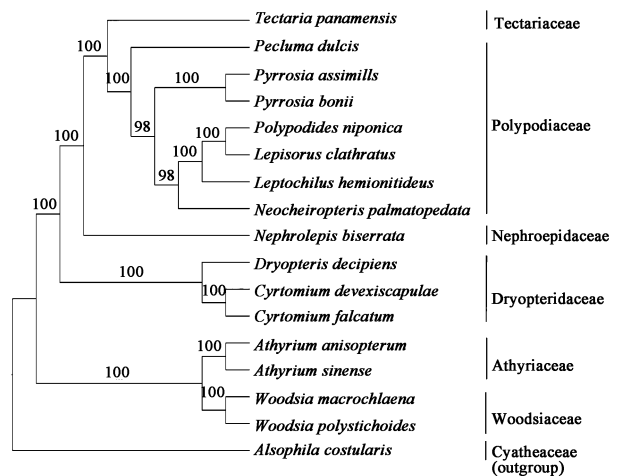


图 3 基于叶绿体全基因组构建水龙骨科和其他植物的 ML 树  
Fig. 3 Phylogenetic tree constructed by maximum-likelihood based on complete chloroplast genomes of Polypodiaceae and other species

### 3 讨论

本草基因组学作为一门新兴学科,从基因组水

平研究中药对人体的作用<sup>[1]</sup>。涉及中草药结构基因组、中草药基因组学、中药分子鉴定、药用模式生物和基因辅助分子育种等多个领域<sup>[1]</sup>,同时人参<sup>[28]</sup>、三七<sup>[29]</sup>、丹参<sup>[30-31]</sup>等药用植物基因组的测序工作中取得了阶段性成果。我国中药基因资源的紧缺限制了大多数药用植物分子遗传学研究,严重影响了中药材新品种选育和种质资源保护<sup>[1]</sup>。随着高通量技术的迅速发展,叶绿体基因组测序方法更加便捷,目前水龙骨科已发表的叶绿体基因组序列较少,石韦属的叶绿体基因组只有波氏石韦。相近石韦叶绿体基因组学研究可为石韦属物种鉴定、基因转化及遗传改良等方面探索提供指导。

超级条形码作为鉴定物种的一种手段,受到越来越多人关注。由于传统的 DNA 条形码序列包含的遗传信息有限,LI 等<sup>[32]</sup>提出将叶绿体全基因组序列或部分特异性片段作为物种鉴定的超级条形码,叶绿体全基因组序列作为超级条形码在藁吾属<sup>[33]</sup>、薯蕷属<sup>[34]</sup>、铁皮石斛<sup>[35]</sup>等药用植物鉴定中已有应用。相近石韦叶绿体基因组全局比对分析,筛选出 5 个高变区(*psbA*, *rrn16*, *petA-psbJ*, *ndhC-trnM* 和 *psbM-petN*)可作为石韦属物种鉴定的候选条形码。以叶绿体全基因组作为超级条形码对石韦属进行鉴定,或通过叶绿体基因组序列比对筛选高变异区作为潜在的 DNA 条形码对石韦属物种进行鉴定,尚需增加更多石韦属物种样本进行验证。

叶绿体基因转化技术已成为叶绿体基因工程研究的热点之一。叶绿体基因组富含大量的遗传信息且具有多拷贝性<sup>[36]</sup>,叶绿体遗传转化与核遗传转化相比具有很多优点,叶绿体外源基因能高度表达,可同时转化多个基因<sup>[37]</sup>。将外源基因定点整合到叶绿体基因组的方法在烟草<sup>[38]</sup>、马铃薯<sup>[39]</sup>、拟南芥<sup>[40]</sup>等多种植物中已成功应用。叶绿体转化技术虽然取得了阶段性成果<sup>[41-44]</sup>,但大部分药用植物还缺乏叶绿体基因组,相近石韦在叶绿体基因转化方面的研究仍是空白,也是相近石韦叶绿体基因组学研究的一大难点,需进一步深入研究,扩大应用范围,建立更加全面的叶绿体基因组转化技术体系<sup>[45]</sup>。

石韦属植物在干燥的季节,叶片会干燥卷缩,降雨后叶片也能吸水复苏,具有明显的抗旱性。李胜男<sup>[46]</sup>对庐山石韦、有柄石韦及柔软石韦进行了干旱胁迫和恢复实验,发现庐山石韦最耐旱,在干旱胁迫(含水量 0~10%)至 30 d 时叶片出现卷曲;水分恢复试验表明,柔软石韦无法恢复且植株全部死亡;而

庐山石韦和有柄石韦能恢复正常。卷柏又名复活草,在极度干旱的情况下也能遇水复苏,XU 等<sup>[47]</sup>通过比较卷柏同属植物的叶绿体基因组发现,卷柏的耐旱性与叶绿体基因组中 *NDH* 基因的缺失有关。但在相近石韦叶绿体基因组中并未发现 *NDH* 基因的缺失,所以石韦属物种的抗旱性可能与卷柏的耐旱机制不同,相近石韦及石韦属抗旱性是否与某个叶绿体基因有关,还需通过测定更多的石韦属叶绿体基因组来确认。

相关功能的叶绿体基因能组成复合操纵子结构,这一结构特点有利于叶绿体基因组基因的表达调控<sup>[48]</sup>。STAUB 等<sup>[49]</sup>将外源基因导入烟草叶绿体基因组中,烟草转基因表达出具有生物活性的生长激素蛋白。侯丙凯等<sup>[50]</sup>和张中林等<sup>[51]</sup>分别将 *Bt* 蛋白基因导入油菜和烟草叶绿体基因组,获得了具有抗虫性的植株。叶绿体作为一种生物反应器,使外源基因在叶绿体基因组高效表达并合成具有生物活性的药用蛋白<sup>[52]</sup>,可以克服花粉传递过程中基因漂移和环境安全性等问题<sup>[53-54]</sup>。叶绿体基因组学的研究使人们对叶绿体基因表达、叶绿体基因组功能的认识更加深入<sup>[55]</sup>,相近石韦叶绿体全基因组的测定为石韦属物种外源基因的表达研究提供了平台。叶绿体基因组学的研究将推进相近石韦及其他石韦属药用植物在分子鉴定、遗传基因转化、抗性蛋白表达及次生代谢途径等叶绿体基因工程方面的研究进程,为发掘石韦属植物抗病、抗逆基因和培育优良性状的药材奠定理论基础。

#### [参考文献]

- [1] 陈士林,宋经元.本草基因组学[J].中国中药杂志,2016,41(21):3881-3889.
- [2] CHEN S L, SONG J Y, SUN C, et al. Herbal genomics: examining the biology of traditional medicines [J]. *Sciences*, 2015, 347(6219): S27-S29.
- [3] 陈士林.本草基因组学[M].北京:科学出版社,2018.
- [4] WU M, LAN S, CAI B, et al. The complete chloroplast genome of *Guadua angustifolia* and comparative analyses of neotropical-paleotropical bamboos [J]. *PLoS One*, 2015, 10(12): e0143792.
- [5] JANSEN R K, RAUBESON L A, BOORE J L, et al. Methods for obtaining and analyzing whole chloroplast genome sequences [J]. *Methods Enzymol*, 2005, 395: 348-384.
- [6] 林小涵,刘志华,李卿,等.药用植物叶绿体基因组研究[J].世界科学技术—中医药现代化,2010,12(3):

- 442-446.
- [ 7 ] WU F H, CHAN M T, LIAO D C, et al. Complete chloroplast genome of *Oncidium* Gower Ramsey and evaluation of molecular markers for identification and breeding in *Oncidiinae* [ J ]. BMC Plant Biol, 2010, 10:68.
- [ 8 ] KUANG D Y, WU H, WANG Y L, et al. Complete chloroplast genome sequence of *Magnolia kwangsiensis* (Magnoliaceae): implication for DNA barcoding and population genetics [ J ]. Genome, 2011, 54 ( 8 ): 663-673.
- [ 9 ] NOCK C J, WATERS D L E, EDWARDS M A, et al. Chloroplast genome sequences from total DNA for plant identification [ J ]. Plant Biotechnol J, 2011, 9 ( 3 ): 328-333.
- [ 10 ] TAKANO A, OKADA H. Phylogenetic relationships among subgenera, species, and varieties of Japanese *Salvia* L. (Lamiaceae) [ J ]. J Plant Res, 2011, 124 ( 2 ): 245-252.
- [ 11 ] 李滢,姚辉,宋经元,等.基于叶绿体全基因组的贝母属特异性 DNA 条形码的筛选 [ J ].世界科学技术—中医药现代化,2016,18(1):24-28.
- [ 12 ] 李西文,胡志刚,林小寒,等.基于 454FLX 高通量技术的厚朴叶绿体全基因组测序及应用研究 [ J ].药学学报,2012,47(1):124-130.
- [ 13 ] LI J H, MUTANDA I, WANG K B, et al. Chloroplastic metabolic engineering coupled with isoprenoid pool enhancement for committed taxanes biosynthesis in *Nicotiana benthamiana* [ J ]. Nat Commun, 2019, 10 ( 1 ):4850.
- [ 14 ] 韩丽颖,赵琳,牛野,等.石韦原植物、药材品种考证 [ J ].中国民族民间医药,2019,28(15):16-20.
- [ 15 ] 文瑞良,伍兰.相异石韦性状及显微鉴别 [ J ].中药材,2000,23(8):446-447.
- [ 16 ] 张雅琴,石钺,宋明,等.蕨类药材石韦及其混伪品的 *psbA-trnH* 序列鉴定 [ J ].中国中药杂志,2014,39(12):2222-2226.
- [ 17 ] CAI S, CAI X, LI S, et al. The complete chloroplast genome of *Pyrrosia bonii* (Polypodiaceae), an important ornamental and medical fern [ J ]. Mitochondrial DNA B, 2018,3(2):801-802.
- [ 18 ] WEI X, QI Y, ZHANG X, et al. Phylogeny, historical biogeography and characters evolution of the drought resistant fern *Pyrrosia* Mirbel (Polypodiaceae) inferred from plastid and nuclear markers [ J ]. Sci Rep, 2017, 7 ( 1 ):12757.
- [ 19 ] ZHOU J, CHEN X, CUI Y, et al. Molecular structure and phylogenetic analyses of complete chloroplast genomes of two *Aristolochia* medicinal species [ J ]. Int J Mol Sci, 2017,18(9):E1839.
- [ 20 ] FRAZER K A, PACHTER L, POLIAKOV A, et al. Vista; computational tools for comparative genomics [ J ]. Nucleic Acids Res, 2004, doi:10.1093/nar/gkh458.
- [ 21 ] PARK I, YANG S, GOYA C, et al. The complete chloroplast genome sequences of *Aconitum pseudolaeve* and *Aconitum longecassidatum*, and development of molecular markers for distinguishing species in the *Aconitum* subgenus *Lycocotnum* [ J ]. Molecules, 2017, 22 ( 11 ):E2012.
- [ 22 ] JIAO Y, JIA H M, LI X W, et al. Development of simple sequence repeat (SSR) markers from a genome survey of Chinese bayberry (*Myrica rubra*) [ J ]. BMC Genomics, 2012,13:201.
- [ 23 ] YANG A H, ZHANG J J, YAO X H, et al. Chloroplast microsatellite markers in *Liriodendron tulipifera* (Magnoliaceae) and cross-species amplification in *L. chinense* [ J ]. Am J Bot, 2011, 98(5):e123-e126.
- [ 24 ] TANG D Q, LU J J, FANG W, et al. Development, characterization and utilization of GenBank microsatellite markers in *Phyllostachys pubescens* and related species [ J ]. Mol Breeding, 2010, 25 ( 2 ):299-311.
- [ 25 ] POWELL W, MORGANTE M, MCDEVITT R, et al. Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines [ J ]. Proc Natl Acad Sci USA, 1995, 92 ( 17 ):7759-7763.
- [ 26 ] ZHOU J, CUI X, CHEN X, et al. Complete chloroplast genomes of *Papaver rhoeas* and *Papaver orientale*: molecular structures, comparative analysis, and phylogenetic analysis [ J ]. Molecules, 2018, 23 ( 2 ):E437.
- [ 27 ] MIN Y, CAI S, XIAO H, et al. The complete chloroplast genome of *Pyrrosia calvata* (Polypodiaceae), a traditional Chinese medicinal fern only restricted to Guangxi, China [ J ]. Mitochondrial DNA B, 2019, 4 ( 1 ):1757-1758.
- [ 28 ] XU J, CHU Y, LIAO B, et al. *Panax ginseng* genome examination for ginsenoside biosynthesis [ J ]. GigaScience, 2017, 6 ( 11 ):1-15.
- [ 29 ] ZHANG D, LI W, XIA E H, et al. The medicinal herb *Panax notoginseng* genome provides insights into ginsenoside biosynthesis and genome evolution [ J ]. Mol Plant, 2016, 10 ( 6 ):903-907.
- [ 30 ] XU H, SONG J, LUO H, et al. Analysis of the genome sequence of the medicinal plant *Salvia miltiorrhiza* [ J ]. Mol Plant, 2016, 9 ( 6 ):949-952.
- [ 31 ] XU Z, PETERS R J, WEIRATHER J, et al. Full-length

- transcriptome sequences and splice variants obtained by a combination of sequencing platforms applied to different root tissues of *Salvia miltiorrhiza* and tanshinone biosynthesis [J]. *Plant J*, 2015, 82 (6): 951-961.
- [32] LI X, YANG Y, HENRY R J, et al. Plant DNA barcoding: from gene to genome [J]. *Biol Rev*, 2015, 90 (1): 157-166.
- [33] CHEN X, ZHOU J, CUI Y, et al. Identification of *Ligularia* herbs using the complete chloroplast genome as a super-barcode [J]. *Front Pharmacol*, 2018, 9: 695.
- [34] 马双姣, 周建国, 李滢, 等. 薯蓣和叉蕊薯蓣叶绿体基因组及特异性 DNA 条形码鉴定序列筛选研究 [J]. *中国科学: 生命科学*, 2018, 48(5): 571-582.
- [35] YANG P, ZHOU H, QIAN J, et al. The complete chloroplast genome sequence of *Dendrobium officinale* [J]. *Mitochondrial DNA A DNA Mapp Seq Anal*, 2016, 27(2): 1262-1264.
- [36] QIAN J, SONG J, GAO H, et al. The complete chloroplast genome sequence of the medicinal plant *Salvia miltiorrhiza* [J]. *PLoS One*, 2013, 8(2): e57607.
- [37] 钱雪艳, 杨向东, 郭东全, 等. 植物叶绿体遗传转化及研究进展 [J]. *分子植物育种*, 2008, 6(5): 959-966.
- [38] HAJDUKIEWICZ P T, GILBERTSON L, STAUB J M, et al. Multiple pathways for Cre/lox-mediated recombination in plastids [J]. *Plant J*, 2001, 27(2): 161-170.
- [39] SIDOROV V A, KASTEN D, PANG S Z, et al. Stable chloroplast transformation in potato: use of green fluorescent protein as a plastid marker [J]. *Plant J*, 1999, 19(2): 209-216.
- [40] SIKDAR S R, SERINO G, CHAUDHURI S, et al. Plastid transformation in *Arabidopsis thaliana* [J]. *Plant Cell Rep*, 1998, 18(1/2): 20-24.
- [41] KUMAR S, DHINGRA A, DANIELL H. Plastid-expressed betaine aldehyde dehydrogenase gene in callus, roots, and leaves confer enhanced salt tolerance [J]. *Plant Physiol*, 2004, 136 (1): 2843-2854.
- [42] HOU B K, ZHOU Y H, WAN L H, et al. Chloroplast transformation in oilseed rape [J]. *Transgenic Res*, 2003, 12(1): 111-114.
- [43] 何道一, 夏仁杰, 周秀杰, 等. 大豆 *Lea5* 基因过表达提高烟草抗旱和耐盐性的研究 [J]. *激光生物学报*, 2018, 27(1): 53-59.
- [44] 高岩. 拟南芥 *fad8* 基因在烟草叶绿体基因组表达研究 [D]. 哈尔滨: 哈尔滨师范大学, 2010.
- [45] 王永飞, 马三梅, 王莹. 高等植物叶绿体基因组转化的应用 [J]. *遗传*, 2004, 26(6): 977-983.
- [46] 李胜男. 三种石韦属植物的水分胁迫研究 [D]. 长沙: 中南林业大学, 2016.
- [47] XU Z C, XIN T Y, Bartels D, et al. Genome analysis of the ancient tracheophyte *Selaginella tamariscina* reveals evolutionary features relevant to the acquisition of desiccation tolerance [J]. *Mol Plant*, 2018, 11 (7): 983-994.
- [48] 苏宁, 孟昆, 沈桂芳. 植物叶绿体基因组基因表达调控的研究 [J]. *生物技术通报*, 1999(1): 14-17.
- [49] STAUB J M, GARCIA B, GRAVES J, et al. High-yield production of a human therapeutic protein in tobacco chloroplasts [J]. *Nat Biotechnol*, 2000, 18(3): 333-338.
- [50] 侯丙凯, 陈正华. 苏云金芽孢杆菌杀虫蛋白基因克隆及油菜叶绿体遗传转化研究 [J]. *遗传*, 2001, 23(1): 39-40.
- [51] 张中林, 任延国, 沈燕新, 等. 苏云金芽孢杆菌 (Bt) 晶体毒蛋白基因在烟草叶绿体中的表达 [J]. *遗传学报*, 2000, 27(3): 270-277.
- [52] 杜美丽, 刘秀明, 江莺, 等. 叶绿体表达系统的研究进展 [J]. *中国农学通报*, 2011, 27(9): 336-341.
- [53] 王金辉, 李轶女, 倪丕冲, 等. 叶绿体转化体系研究进展 [J]. *生物技术通报*, 2012(1): 1-6.
- [54] 许柯, 王玉伟, 商艳芳. 叶绿体基因组的转化研究 [J]. *安徽农学通报*, 2007, 13(17): 29-30.
- [55] 肖钢, 官春云. 植物基因工程的新方向—叶绿体基因工程 [J]. *作物研究*, 2004(5): 389-392.

[责任编辑 刘德文]