

一年蓬叶绿体基因组特征及系统进化分析

周芸芸¹, 郭敏¹, 彭英权¹, 金晨钟¹, 张雪娇², 胡一鸿¹, 陈功¹, 龚意辉^{1*}

1. 湖南省农田杂草防控技术与应用协同创新中心/湖南人文科技学院农作物有害生物绿色防控重点实验室, 湖南娄底 417000; 2. 邵阳学院, 湖南邵阳 422000

摘要: 一年蓬 (*Erigeron annuus*) 作为一种入侵植物, 对我国农林业发展、本地物种多样性及生态系统具有严重的影响。本研究利用 Illumina NovaSeq 6000 测定一年蓬的 DNA 序列, 采用 SPAdes v3.10.1 软件组装一年蓬叶绿体基因组, 以小蓬草 (*Erigeron canadensis*, MT806101.1) 叶绿体基因组为参考, 对一年蓬叶绿体基因组结构、特征、SSR 位点、IR 边界以及进化树进行分析, 结果表明: 一年蓬叶绿体基因组全长为 153 283 bp, AT、GC 含量分别为 62.95%、37.05%。包括大单拷贝区 (LSC) 84 833 bp, 小单拷贝区 (SSC) 18 412 bp, 反向重复区 (IRa 和 IRb) 25 019 bp, 一年蓬叶绿体基因组共编码到 132 个基因, 包括蛋白编码基因 88 个, 转运 RNA 基因 36 个, 核糖体 RNA 基因 8 个。按照功能分类将它们分成四大类: 光合作用相关基因共 45 个、自我复制相关基因共 73 个、其他基因共 6 个、未知功能基因共 8 个。在一年蓬叶绿体基因组中共查找 200 个符合条件的 SSR 位点, 分布在 LSC、SSC、IR 区域的 SSR 数量分别为 126、40、34 个, 分别占总 SSR 位点数比重为 63%、20%、17%, 以单核苷酸重复占绝对优势, 主要是 A/T 碱基, 其次是三核苷酸重复单元类型, 主要是 TTA、TAA 和 TAT。进化树分析表明, 一年蓬与小蓬草亲缘关系最近。本研究揭示了一年蓬叶绿体基因组信息, 并确定了一年蓬系统发育所属分支, 为进一步开展一年蓬种群遗传学、系统发育地理学、适应性机制及防控与利用等研究提供分子理论支撑。

关键词: 一年蓬; 叶绿体; 基因组; 结构特征; 系统发育

中图分类号: S668.3 文献标识码: A

Characterization and Phylogenetic Analysis of the Complete Chloroplast Genome of *Erigeron annuus*

ZHOU Yunyun¹, GUO Min¹, PENG Yingquan¹, JIN Chenzhong¹, ZHANG Xuejiao², HU Yihong¹, CHEN Gong¹, GONG Yihui^{1*}

1. Hunan Provincial Collaborative Innovation Center for Field Weeds Control / Key Laboratory of Green Control of Crop Pests in Hunan Higher Education, Hunan University of Humanities, Science and Technology, Loudi, Hunan 417000, China; 2. Shaoyang University, Shaoyang, Hunan 422000, China

Abstract: *Erigeron annuus*, a member of the Asteraceae family, is widely distributed in tropic, subtropical and temperate regions of China. It is a plant native to North America, which was introduced to China at the end of the nineteenth century as an ornamental plant. *E. annuus* has great medicinal value. Its whole plant has been used in folk medicine for the treatment of indigestion, enteritis, epidemic hepatitis, lymphadenitis, hematuria, acute inflammation, malaria and obesity. *E. annuus* could be a phytoremediator for metal- and metalloloid-contaminated soils. It can degrade diesel effectively, has strong tolerance and accumulation for uranium, cadmium, zinc, lead, and antimony, and can be distributed in soils with high pollution levels. Hence, In this study, it was to explore the characteristics of chloroplast (cp) genome and its phylogenetic relationship, laying a foundation for the further development of species identification, control and full

收稿日期 2022-06-21; 修回日期 2022-07-11

基金项目 湖南省科技创新计划项目 (No. 2019NK4170); 湖南省教育厅科学研究重点项目 (No. 20A281); 湖南省教育厅重点项目 (No. 21A0480)。

作者简介 周芸芸 (1986—), 女, 博士, 讲师, 研究方向: 野生动植物资源管理和分子生态学。*通信作者 (Corresponding author): 龚意辉 (GONG Yihui), E-mail: gyhzhg@163.com。

utilization of *E. annuus*. The complete cp genome of *E. annuus* was assembled and sequenced by Illumina Nova Seq 6000 Sequencing platform and SPAdes v3.10. software. Taking the cp genome of *E. canadensis* (MT806101.1) as a reference, the genome feature and phylogenetic analysis study were investigated. The cp genome of *E. annuus* was a circular and quadripartite structure with a total length of 153 283 bp. The GC and AT content of the cp genome was 36.77% and 63.23%, respectively. The cp genome of *E. annuus* had four typical regions, containing a pair of inverted repeats (IRs, 25 019 bp) that were separated by the large single-copy (LSC, 84 833 bp) and small single-copy (SSC, 18 412 bp). In total, 132 genes were predicted, consisting of 88 protein-coding genes, 36 transfer RNA genes, and 8 ribosomal RNA genes. The genes were divided into 4 categories according to the functions, including 45 photosynthesis genes, 73 self-replication genes, 6 other genes and 8 genes of unknown function. SSR site analysis revealed that the cp genome sequence contained 200 SSR loci in total, which were spaced disproportionately through the genome with the largest number of SSRs distributed in the LSC region, followed by the SSC and IR regions. LSC, SSC and IR regions had 126, 40, and 34 SSRs, respectively. The number of SSRs in LSC, SSC and IR regions was accounted with 63%, 20% and 17%, respectively. The SSRs were mononucleotides, especially for A/T. The second most abundant motif type was the trinucleotide type, especially TAA, TAT and TTA. Phylogenetic analysis showed that *E. annuus* was closely related to the plant of *E. canadensis*. This study revealed the cp genomic information of *E. annuus* and would provide essential data for further studies on genetics, phylogeny, utilization and prevention of this species. This study revealed the chloroplast genome information and identified the phylogenetic branches of *E. annuus*, which would provide molecular theoretical support for further research on population genetics, phylogenetic geography, adaptive mechanism and management of *E. annuus*.

Keywords: *Erigeron annuus*; chloroplast; genome; structure characteristics; phylogenetic

DOI: 10.3969/j.issn.1000-2561.2023.04.004

一年蓬 [*Erigeron annuus* (L.) Pers.] 隶属于菊科 (Asteraceae), 紫菀族 (Astereae), 飞蓬属 (*Erigeron*), 一年或两年生植物, 原产于北美洲, 早年作为观赏植物进入我国, 现广泛分布于我国^[1]。一年蓬的花形似小雏菊, 民间用全草或根可入药, 具有清热解毒、消食止泻的功效^[1]。近年来有研究报道一年蓬植株提取液具有抗炎、抑菌、止血、镇痛的效果^[2-5], 在一年蓬水提物的研究中还发现其对高脂肪饮食的小鼠具有抗肥胖的作用^[6]。一年蓬被认为是一种修复金属和类金属污染土壤的潜能植物, 对 Cd 和 Zn 等金属及柴油具有较好的富集和耐受效果^[7-8]。然而它作为一种入侵植物, 对我国农林业发展、本地物种多样性及生态系统具有严重的影响, 被我国生态环境部列入第三批外来入侵物种名单^[1, 9]。目前国内外对一年蓬活性化合物^[10-12]、入侵及适应机制^[13-15]、防除^[16-17]等方面研究较多, 而有关一年蓬叶绿体全基因组特征、序列分析及其相关系统发育分析等尚未见报道。

叶绿体 (chloroplast, CP) 是为绿色植物提供能量的光合细胞器, 它们在光合作用及相关代谢活动中起重要作用^[18]。被子植物中典型的叶绿体基因组为双链共价闭合环状分子, 具有典型的四分体结构, 其中一个大单拷贝区 (large single copy

region, LSC), 小单拷贝区 (small single copy region, SSC), 反向重复区 A (inverted repeat region A, IRA), 反向重复区 B (inverted repeat region B, IRB)^[19]。基因组的长度一般在 120~170 kb 之间, 通常编码 110~130 个基因, 约 40 个基因是参与光合作用、转录和翻译^[20]。叶绿体基因组具自我复制的特点, 在大多数植物中为母系遗传, 进化过程中不经历重组, 能直接反映植物在长期的进化过程中积累的遗传变异, 已被广泛用于阐明物种遗传多样性、进化历程、系统发育地位及近缘种间和种内的遗传关系等相关研究^[21-22]。随着新一代测序技术的发展, 植物基因组得到了广泛的研究, 多种菊科植物的叶绿体基因组研究都已被报道, 涉及到的属有紫菀属 (*Aster*)、向日葵属 (*Helianthus*)、飞蓬属 (*Erigeron*)、蒿属 (*Artemisia*)、菜蓟属 (*Cynara*)、莴苣属 (*Lactuca*)、禾叶金菀属 (*Pityopsis*) 等^[23-27]。本研究以一年蓬为试材, 利用高通量测序, 对组装后的一年蓬叶绿体基因组结构、序列特征、与近缘物种序列的差异, 在紫菀族中的地位及与该属其他植物的亲缘关系等方面进行分析与研究, 旨在为后续一年蓬的遗传、系统发育、进化历程及资源开发利用与防除等研究奠定分子理论基础。

1 材料与方法

1.1 材料

供试材料一年蓬采集于湖南省娄底市娄星区湖南人文科技学院校园内 (27°53'31"N, 111°54'2"E), 植物样本经湖南人文科技学院金晨钟、胡一鸿教授进行植物分类学鉴定。采摘一年蓬健康植株幼嫩绿色叶片, 用液氮速冻后放置于-80 °C 的超低温冰箱中, 备用。

1.2 方法

1.2.1 DNA 提取与测序分析 植物全基因组 DNA 用天根生化科技(北京)有限公司的植物基因组 DNA 提取试剂盒 (DP305) 提取, 并通过 NanoDrop2000 分光光度计和琼脂糖凝胶电泳检测总 DNA 的质量和浓度。将高质量的 DNA 样本送南京集思慧远生物科技有限公司进行测序, 测序采用 Illumina NovaSeq 6000 平台, 使用 fastp 0.20.0 (<https://github.com/OpenGene/fastp>) 软件对原始数据进行过滤得到高质量 reads, 得到 18 635 007 条 reads (Q₂₀ 为 97.49%, Q₃₀ 为 93.02%) 供后续组装和注释用。

1.2.2 叶绿体基因组组装与注释 从 NCBI 上下载已发表的小蓬草 (*Erigeron canadensis*, MT 806101.1) 的叶绿体全基因组序列为参照, 采用 SPAdes^[28]v3.10.1 (<http://cab.spbu.ru/software/spades/>), SSPACE v2.0 (<https://www.baseclear.com/services/bioinformatics/basetools/sspace-standard/>), Gapfillerv2.1.1^[29] (<https://sourceforge.net/projects/gapfiller/>) 在线组装软件对叶绿体基因组进行组装、校正, 得到完整的一年蓬叶绿体环状基因组序列。使用在线软件 prodigal v2.6.3 (<https://www.github.com/hyatt/Prodigal>) 注释叶绿体的 CDS, 使用 hmmer v3.1b2 (<http://www.hmmer.org/>) 在线软件预测 rRNA, 使用在线软件 aragorn v1.2.38 (http://130.235.244.92/ARA_GORN/) 预测 tRNA, 再使用 blastv2.6 (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) 比对组装的序列, 获得最终的注释。将注释完成的一年蓬叶绿体基因组序列, 用在线软件 OGDRAW (<https://chlorobox.mpimp-golm.mpg.de/OGDraw.html>) 制作叶绿体基因组图谱。

1.2.3 叶绿体基因组特征及对比分析 使用 MISA v1.0 (MICROSatellite identification tool, <http://pgrc.ipk-gatersleben.de/misa/misa.html>) 在线软件分析一年蓬叶绿体全基因组 SSR 位点, 利用 Find repeats 工具获得小蓬草、多舌飞蓬

(*E. multiradiatus*)、短葶飞蓬 (*E. breviscapus*)、香丝草 (*E. bonariensis*) 等 4 种基因组的 LSC 序列、SSC 序列和 IR 序列, 采用 MAFFT 软件^[30] 进行多序列比对, 最后使用 RAxML v8.2.10 (<https://cme.h-its.org/exelixis/software.html>) 软件, 选用 GTRGAMMA 模型, rapid Bootstrap 分析, bootstrap=1000, 构建最大似然系统发育树。

2 结果与分析

2.1 一年蓬叶绿体基因组结构

一年蓬叶绿体基因组是一个典型的四分体结构, 其基因组全长为 153 283 bp, AT、GC 含量分别为 62.95%、37.05%。其中包括大单拷贝区 (LSC) 84 833 bp, GC 含量为 34.87%; 小单拷贝区 (SSC) 18 412 bp, GC 含量为 30.95%; 反向重复区 (IRa 和 IRb) 长 25 019 bp, GC 含量为 42.98% (表 1, 图 1)。一年蓬叶绿体基因组共注释出 132 种基因, 编码蛋白质基因、tRNA 基因和 rRNA 基因分别为 88、36、8 (表 1)。

表 1 一年蓬叶绿体基因组特征

Tab. 1 Characteristics of chloroplast genome of *E. annuus*

类别 Category	项目 Item	描述 Description
叶绿体基因组结构大小/bp	大单拷贝区	84 833
	小单拷贝区	18 412
	反向重复区	25 019
	基因组大小	153 283
基因组成	基因总数	132
	编码基因	88
	tRNAs	36
	rRNAs	8
GC 含量/%	大单拷贝区	34.87
	小单拷贝区	30.95
	反向重复区	42.98
	平均值	36.27
单碱基含量/%	A	31.46
	T	31.49
	C	18.48
	G	18.57

2.2 叶绿体基因的组成和特点

按照功能分类将已注释到的这些基因分为光合作用相关基因 (photosynthesis genes)、自我复制基因 (self-replication genes)、其他基因 (other genes)、未知功能基因 (genes of unknown function) 四大类 (表 2)。光合作用相关基因主要包括 5 个光

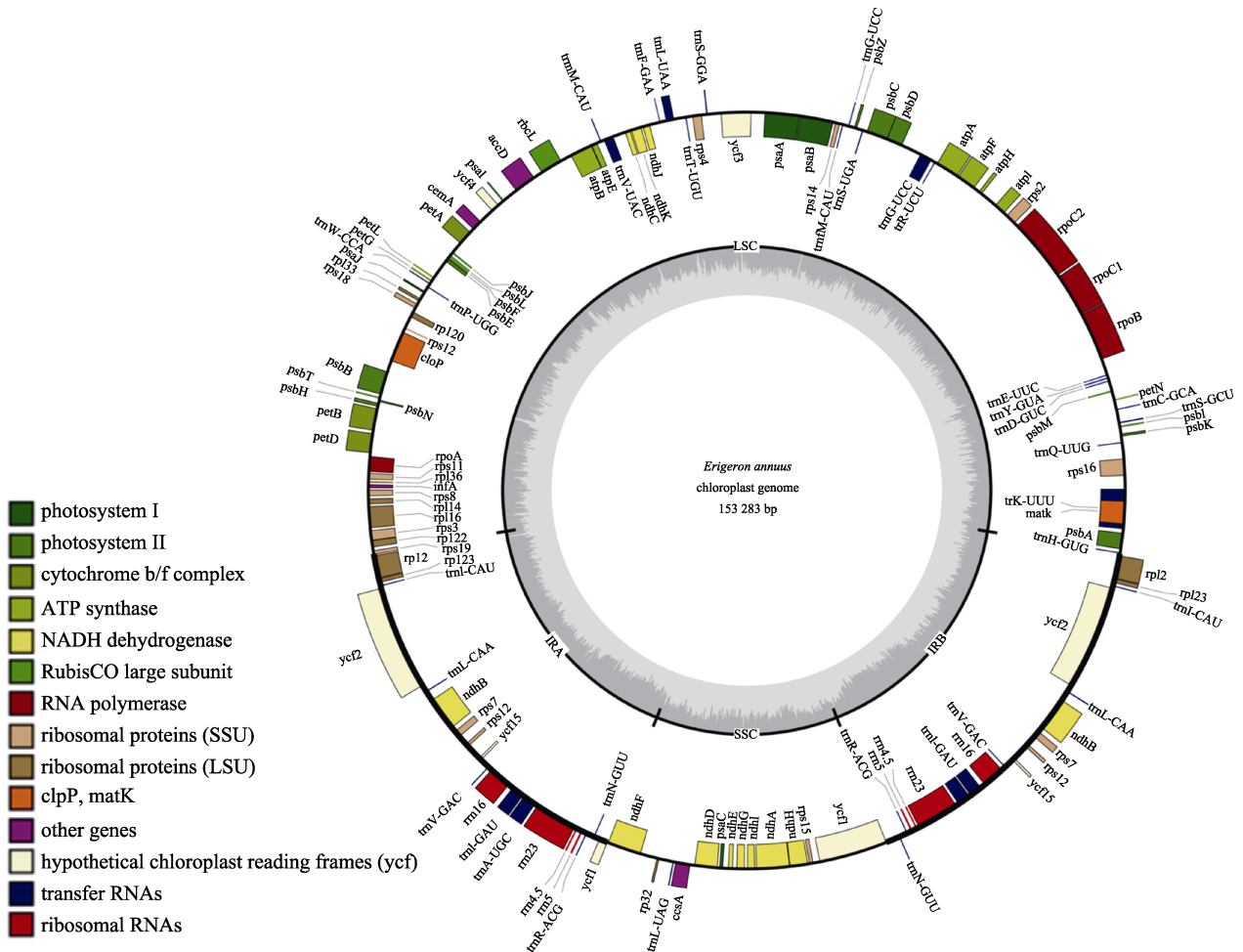


图 1 一年蓬叶绿体基因组

Fig. 1 chloroplast genome of *E. annuus*

合系统 I 基因 (subunits of photosystem I)、15 个光合系统 II 基因 (subunits of photosystem II)、12 个 NADH 脱氢酶基因 (subunits of NADH dehydrogenase)、6 个细胞色素复合物编码基因 (subunits of cytochrome b/f complex)、6 个 ATP 合酶基因 (subunits of ATP synthase)、1 个二磷酸核酮糖羧化酶大亚基基因 (large subunit of rubisco); 自我复制基因主要包括 11 个核糖体蛋白大亚基基因 (proteins of large ribosomal subunit)、14 个核糖体小亚基基因 (proteins of small ribosomal subunit)、4 个 RNA 聚合酶亚基基因 (subunits of RNA polymerase)、8 个 rRNA 基因和 36 个 tRNA 基因; 此外还有 6 个其他功能基因及 8 个未知功能基因。其中 *ndhB*、*rpl2*、*rpl23*、*rps12*、*rps7*、*rml16*、*rml23*、*rml4.5*、*rml5*、*trnA-UGC*、*trnI-CAU*、*trnI-GAU*、*trnL-CAA*、*trnN-GUU*、*trnR-ACG*、*trnV-GAC*、*ycf1*、*ycf15* 和 *ycf2* 均含 2 个拷贝基因。18 个基因含有 1 个内含子, 分别为 *ndhA*、

ndhB(2)、*petB*、*petD*、*atpF*、*rps12*^{*}(2)、*rps16*、*rpoC1*、*trnA-UGC*(2)、*trnG-UCC*、*trnI-GAU*(2)、*trnK-UUU*、*trnL-UAA*、*trnV-UAC*。2 个基因含有 2 个内含子, 分别为 *clpP* 和 *ycf3*。值得注意的是, 本研究所注释到 8 个未知功能基因 (*ycf1*(2)、*ycf15*(2)、*ycf2*(2)、*ycf3*、*ycf4*) 仍需利用先进的分子技术手段对其功能进行鉴定。

2.3 IR 边界的特征

使用在线网站 IRscope (<https://irscope.shinyapps.io/irapp/>) 绘制对比图, 比较一年蓬与多舌飞蓬 (*E. multiradiatus*)、香丝草 (*E. bonariensis*)、短蕊飞蓬 (*E. breviscapus*)、小蓬草 (*E. canadensis*) 4 个物种的 LSC、SSC 和 IR 区的边界位置 (图 2), 这 5 个菊科物种叶绿体基因组 4 个区域中, IR 区较为保守, 其片段大小为 24 326~25 019 bp。*rps19* 基因位于 LSC/IRb 区域中, 其中一年蓬、香丝草、短蕊飞蓬和小蓬草中 LSC 和 IRb 区向扩张程度相同, 均为 217、62 bp, 多舌飞蓬在 LSC 和 IRb 区

表 2 一年蓬叶绿体基因组编码的基因
Tab. 2 Genes present in chloroplast genome of *E. annuus*

基因分类 Gene category	基因分组 Group of gene	基因名称 Gene name
光合作用相关基因	subunits of photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i>
	subunits of 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>
	subunits of NADH dehydrogenase	<i>ndhA</i> [*] , <i>ndhB</i> ⁽²⁾ , <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>
	subunits of cytochrome b/f complex	<i>petA</i> , <i>petB</i> [*] , <i>petD</i> [*] , <i>petG</i> , <i>petL</i> , <i>petN</i>
	subunits of 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>
	subunits photochlorophyllide reductase	—
自我复制基因	proteins of large ribosomal subunit	<i>rpl14</i> , <i>rpl16</i> [*] , <i>rpl2</i> ⁽²⁾ , <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> (2), <i>rpl32</i> , <i>rpl33</i> , <i>rpl36</i>
	proteins of small ribosomal subunit	<i>rps11</i> , <i>rps12</i> ⁽²⁾ , <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> (2), <i>rps8</i>
	subunits of RNA polymerase	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> [*] , <i>rpoC2</i>
	Ribosomal RNAs	<i>rrn16</i> (2), <i>rrn23</i> (2), <i>rrn4.5</i> (2), <i>rrn5</i> (2)
	Transfer RNAs	<i>trnA-UGC</i> ⁽²⁾ , <i>trnC-GCA</i> , <i>trnD-GUC</i> , <i>trnE-UUC</i> , <i>trnF-GAA</i> , <i>trnG-UCC</i> , <i>trnG-UCC</i> [*] , <i>trnH-GUG</i> , <i>trnI-CAU</i> (2), <i>trnI-GAU</i> ⁽²⁾ , <i>trnK-UUU</i> [*] , <i>trnL-CAA</i> (2), <i>trnL-UAA</i> [*] , <i>trnL-UAG</i> , <i>trnM-CAU</i> , <i>trnN-GUU</i> (2), <i>trnP-UGG</i> , <i>trnQ-UUG</i> , <i>trnR-ACG</i> (2), <i>trnR-UCU</i> , <i>trnS-GCU</i> , <i>trnS-GGA</i> , <i>trnS-UGA</i> , <i>trnT-UGU</i> , <i>trnV-GAC</i> (2), <i>trnV-UAC</i> [*] , <i>trnW-CCA</i> , <i>trnY-GUA</i> , <i>trnY-M-CAU</i>
其他基因	maturase	<i>matK</i>
	protease	<i>clpP</i> ^{**}
	envelope membrane protein	<i>cemA</i>
	acetyl-CoA carboxylase	<i>accD</i>
	c-type cytochrome synthesis gene	<i>ccsA</i>
	translation initiation factor	<i>infA</i>
	other	—
未知功能基因	conserved hypothetical chloroplast ORF	<i>ycf1</i> (2), <i>ycf15</i> (2), <i>ycf2</i> (2), <i>ycf3</i> ^{**} , <i>ycf4</i>

注: *表示含有 1 个内含子; **表示含有 2 个内含子; (2) 表示拷贝数为 2 的基因。

Note: * indicates one intron in the gene; ** indicates two introns in the gene; (2) means twice copy gene.

向扩张分别为 212、67 bp。这 5 种植物中的 *rpl2* 基因均位于 IR 区域内。一年蓬中的 *tmN* 基因位于 IRa 区域内, 而多舌飞蓬、香丝草、短荳飞蓬和小蓬草中位于 IRb 区域内; 一年蓬存在 2 个 *ycf1* 基因, 分别在 SSC/IRb 边界扩张, 距离 JSB 边界 567、12 bp, SSC/IRa 边界区域中扩张, 距离 JSA 边界分别为 4482 bp 和 15 bp, 而其他 4 种植物只存在 1 个 *ycf1* 基因; SSC/IRb 边界显示, 一年蓬、多舌飞蓬、短荳飞蓬和小蓬草中的 *ndhF* 基因全部在 SSC 区域内扩张, 而香丝草并没有检测到 *ndhF* 基因。这 5 种植物中的 *tmH* 基因全部在 LSC 区域内扩张。

2.4 叶绿体基因组 SSR 位点分析

使用 MISA v1.0 (<http://pgrc.ipk-gater-sleben.de/misa/misa.html>) 在线软件对一年蓬叶绿体基因组进

行重复单元 (SSR) 位点分析 (图 3)。共在一年蓬叶绿体基因组中查找到 200 个符合条件的 SSR 位点, 分布在 LSC、SSC、IR 区域的 SSR 数量分别为 126、40、34 个, 分别占总 SSR 位点数比重为 63%、20%、17%。单核苷酸重复单元共 139 个, 主要以 A 和 T 为主, 双核苷酸重复单元 20 个, 主要以 AT 和 TA 为主, 三核苷酸重复单元 106 个, 主要以 TTA、TAA 和 TAT 为主。四核苷酸重复单元 13 个, 主要以 TATT 和 TTAT 为主, 五核苷酸重复单元 3 个, 分别为 TACTA、TATAT、TTAGA。说明一年蓬叶绿体基因组中 A/T 碱基占优势, 具有碱基偏好性。这与该物种中 AT 含量占 62.95% 的结果相一致。

2.5 系统发育分析

将一年蓬与菊科管状花亚科紫菀族 21 种植

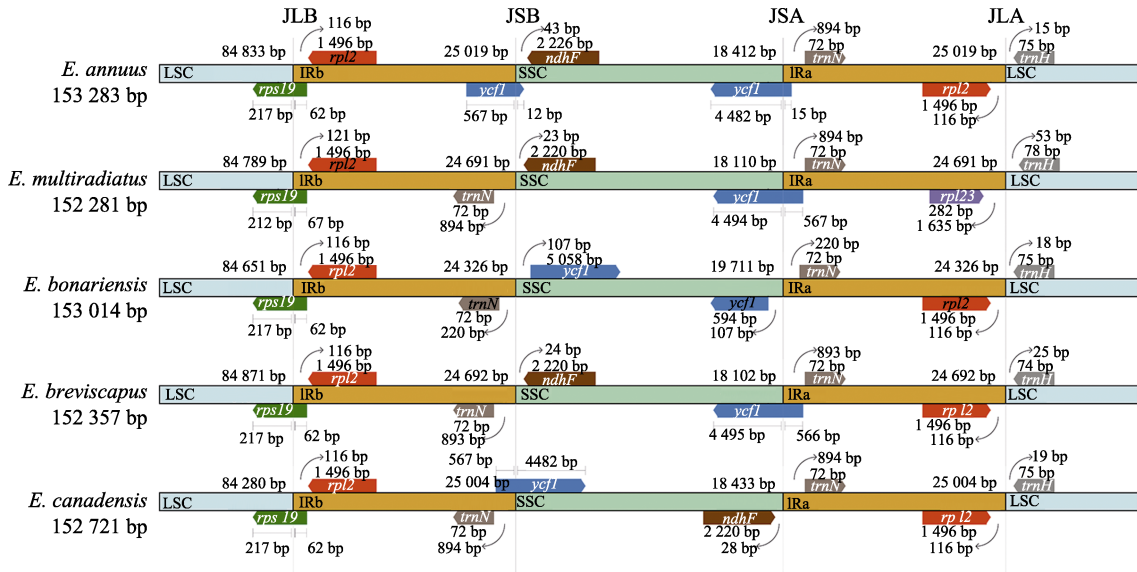


图 2 一年蓬及 4 个菊科植物的叶绿体全基因组边界图

Fig. 2 Border regions of complete chloroplast genomes of *E. annuus* and four Asteraceae plants

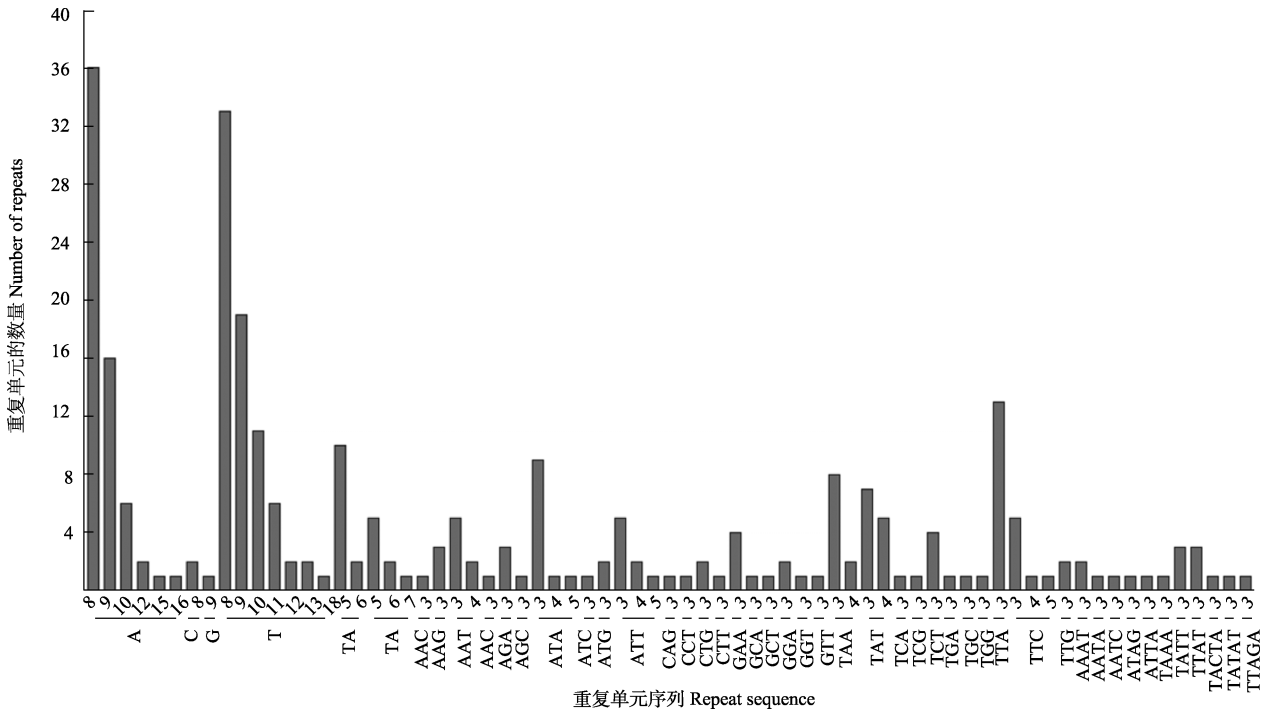


图 3 一年蓬重复序列 SSR 位点统计

Fig. 3 Number of SSR motifs in *E. annuus*

物叶绿体基因组序列，以舌状花亚科苦苣菜属植物苦苣菜 (*Ixeris polycephala*) 为外群构建发育树 (图 4)。进化树结果显示，一年蓬与一枝黄花 (*Solidago decurrens*)、钻叶紫菀 (*Symphytotrichum subulatum*)、多舌飞蓬、短蕊飞蓬、香丝草、小蓬草、瓶头草 (*Lagenophora cuchumatana*) 聚在一个分支上，紫菀属植物聚类在另外一支。从飞蓬属的 5 个物种的亲缘关系来看，一年蓬与同

属的小蓬草 (MT806101.1) 聚在一起，形成 2 个姊妹类群。

3 讨论

叶绿体细胞器参与高等植物光合作用和碳固定等生理活动，是承担能量转换的重要细胞器^[31]。叶绿体基因组是植物细胞内 3 种基因组中最为保守的一类，具有高度的保守性和较慢的进化速率，

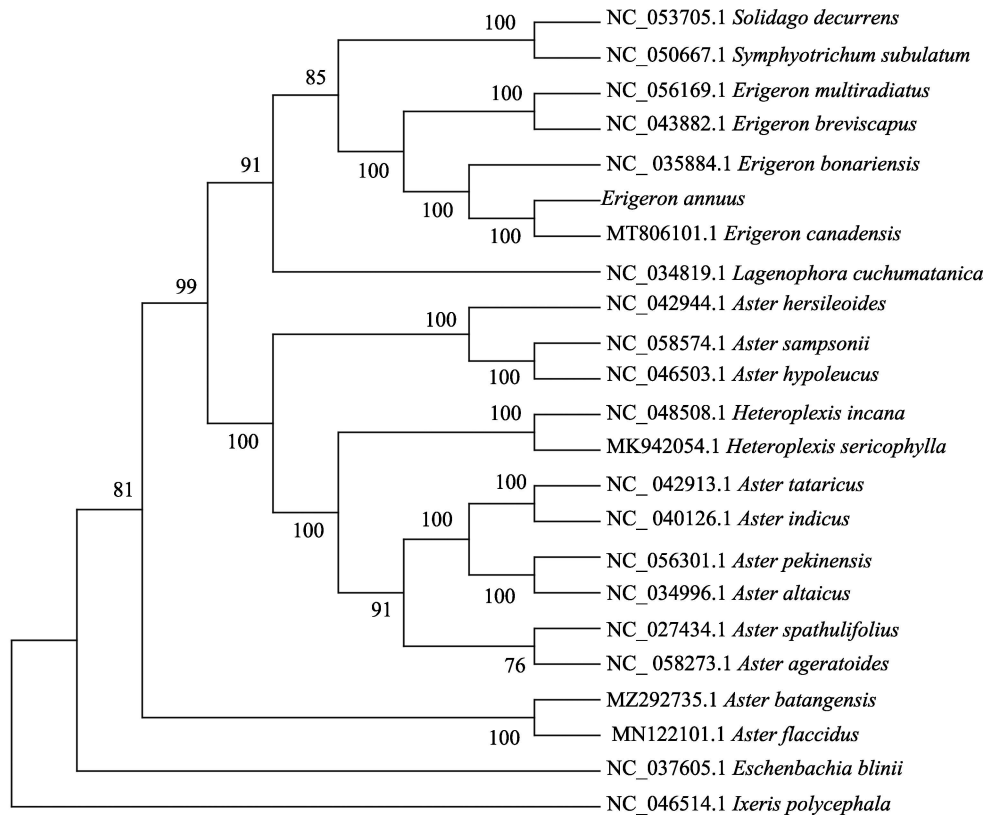


图 4 基于叶绿体全基因组构建的系统发育树

Fig. 4 Phylogenetic tree based on complete chloroplast genome

在不同物种间的系统进化研究中有更好的统一性^[32]。植物叶绿体基因组长度一般为 107~218 kb, 不同物种之间长度略有不同, 其长度的变化主要是由 IR 区的收缩和扩张导致^[33]。本研究所获得的一年蓬叶绿体基因组大小与结构与已报道的被子植物研究结果相符。一年蓬叶绿体基因组是一个典型的四分体结构, 全长为 153 283 bp, 与同属植物小蓬草 (152 721 bp)、多舌飞蓬 (152 281 bp)、短葶飞蓬 (152 357 bp)、香丝草 (153 014 bp) 差异小, 与已报道的菊科植物叶绿体基因组长度 (149.51~153.20 kb) 相比较, 一年蓬及其所属的飞蓬属植物叶绿体基因组长度较长^[23-34]。一年蓬叶绿体基因组中 AT、GC 含量分别为 62.95%、37.05%, 其中反向重复区 (IRs) 中 GC 含量达到最高 (42.98%), 这与其他菊科植物叶绿体基因组特征一致^[25]。一年蓬基因组大单拷贝区 (LSC) 84 833 bp, 小单拷贝区 (SSC) 18 412 bp, 反向重复区 (IRs) 25 019 bp, 与多舌飞蓬 (AT: 62.90%、GC: 37.10%; LSC: 84 881 bp、SSC: 18 102 bp、IRs: 24 692 bp)、短葶飞蓬 (AT: 62.80%、GC: 37.20%; LSC: 84 789 bp、SSC:

18 110 bp、IRs: 24 691 bp) 遗传差异很小, 主要差异在 SSC 和 IRs 区域, 相差 300 bp 左右^[35]。通过 IRscope 比较分析, 发现飞蓬属的几个物种叶绿体基因组结构和大小高度保守, 但是 SC 和 IR 边界仍会有细微区别。这与苏玥等^[36]对乳苣叶绿体基因组特征研究结果一致, 乳苣与其他 5 种菊科近缘种植物叶绿体基因组进行比较分析时发现基因组序列整体高度相似, 但是在 *rps19*、*rpl2*、*ycf1*、*trnN*、*ndhF*、*trnH* 和 *rpl23* 等基因上有变化。

一年蓬叶绿体基因组共注释出 132 种基因, 其中编码蛋白质基因、tRNA 基因和 rRNA 基因数量分别为 88、36、8 个。梁凤萍等^[23]对 27 种菊科植物叶绿体基因组进行注释, 发现菊科植物叶绿体基因组共有基因一般为 125~138 个, 编码蛋白基因为 83~90 个, tRNA 基因 32~43 个, rRNA 基因 7~9 个。同属多舌飞蓬和短葶飞蓬叶绿体基因组共注释 129 个基因, 编码蛋白质基因 84 个、tRNA 基因 37 个和 rRNA 基因 8 个^[35]。一年蓬叶绿体基因组基因数目特征与其他菊科植物一致, 值得提出的是本研究注释到 8 个未知功能基因 [*ycf1*(2)、*ycf15*(2)、*ycf2*(2)、*ycf3*、*ycf4*], 需进一

步利用先进的分子技术手段对其功能进行鉴定。

叶绿体 SSR (cp SSR) 通常是短的单核苷酸串联重复序列, 当位于线粒体基因组的非编码区域时, 通常显示种内变异重复次数^[37]。cp SSR 可以在同一物种内表现出高度变异, 被广泛应用于群体遗传学和系统发育分析等领域^[38-39]。在一年蓬叶绿体基因组 LSC、SSC、IR 区域的 SSR 数量分别为 126、40、34 个, 单核苷酸重复单元最多, 其次是三核苷酸重复单元, A/T 碱基占优势, 在菊科^[40-41]和其他植物科^[42-43]中研究结果一致。这些 SSR 可为飞蓬属物种分子标记开发及物种鉴定提供理论依据。

随着叶绿体基因组测序数量的增加, 很多学者通过该类基因组重建了菊科植物的系统发育关系。本研究在 NCBI 上下载了 22 种紫菀族植物, 以舌状花亚科苦苣菜属植物苦苣菜为外群, 构建系统发育树。一年蓬在内的飞蓬属与联毛紫菀属、一枝黄花属和瓶头草属的植物聚类在一个分支上, 紫菀属植物另聚类为一支。这与基于 ITS 等分子标记建立的分子系统发育树显示的小蓬草与飞蓬属、一枝黄花属等构成一个分支的结果一致^[44]。一年蓬与同属小蓬草聚在一起, 认为基于叶绿体基因组数据, 同属植物的亲缘关系最近。本研究对一年蓬叶绿体基因组基因结构、SSRs 数量及分布和基因功能等进行了分析, 结合已公布的紫菀族物种叶绿体基因组序列构建了系统发育树, 阐明了一年蓬和紫菀族不同物种之间的系统发育关系, 丰富了菊科植物的进化关系, 为植物分类、遗传进化、系统发育及一年蓬的开发和防除等研究奠定了分子基础。

参考文献

- [1] 范建军, 乙杨敏, 朱珣之. 入侵杂草一年蓬研究进展[J]. 杂草学报, 2020, 38(2): 1-8.
FAN J J, YI Y M, ZHU X Z. Advances on the invasive weed, *Erigeron annuus*[J]. Weed Science, 2020, 38(2): 1-8. (in Chinese)
- [2] 徐琅, 贾元超, 龚祝南. 一年蓬挥发油的气相色谱-质谱分析及体外抑菌活性研究[J]. 时珍国医国药, 2009, 20(5): 1171-1172.
XU L, JIA Y C, GONG Z N. GC-MS analysis and *in vitro* antibacterial activity of volatile oil from *Erigeron annuus*[J]. Lishizhen Medicine and Materiamedic Research, 2009, 20(5): 1171-1172. (in Chinese)
- [3] 唐红艳, 黄光伟, 刘文佳, 韦宝韩. 含一年蓬中药提取液牙膏的抗炎止血镇痛作用研究[J]. 口腔护理用品工业, 2018, 28(4): 26-28.
TANG H Y, HUANG G W, LIU W J, WEI B H. Study on the anti-inflammatory, hemostatic and analgesic effects of toothpaste containing the extract of traditional Chinese medicine *Erigeron annuus*[J]. Oral Care Industry, 2018, 28(4): 26-28. (in Chinese)
- [4] JO M J, LEE J R, CHO I J. Roots of *Erigeron annuus* attenuate acute inflammation as mediated with the inhibition of NF- κ B-associated nitric oxide and prostaglandin E2 production[J]. Evid-Based Complement and Alternative Medicine, 2013: 297427.
- [5] YOO N H, JANG D S, YOO J L, LEE Y M, KIM Y S, CHO J H, KIM J S. *Erigeron flavanone*, a flavanone derivative from the flowers of *Erigeron annuus* with protein glycation and a aldose reductase inhibitory activity[J]. Journal of Natural Products, 2008, 71(4): 713-715.
- [6] ZHENG Y L, CHOI Y H, LEE J H, LEE S Y, KANG I J. Anti-obesity effect of *Erigeron annuus* (L.) Pers. extract containing phenolic acids[J]. Foods, 2021, 10(6): 1266.
- [7] XI Y, SONG Y Z, LIU H G, JOHNSON D M, HUANG, Y P. Selenium enhanced degradation of diesel by *Erigeron annuus*[J]. Journary of Soils Sediments, 2018, 18(5): 1906-1914.
- [8] ZUO R Z, LIU H G, XI Y, GU Y, REN D, YUAN X, HUANG Y P. Nano-SiO₂ combined with a surfactant enhanced phenanthrene phytoremediation by *Erigeron annuus* (L.) Pers[J]. Environmental Science and Pollution Research, 2020, 27(19): 20538-20544.
- [9] 闫小玲, 寿海洋, 马金双. 浙江省外来入侵植物研究[J]. 植物分类与资源学报, 2014, 36(1): 77-88.
YAN X L, SHOU H Y, MA J S. The alien invasive plants in Zhejiang Province, China[J]. Plant Diversity and Resources, 2014, 36(1): 77-88. (in Chinese)
- [10] LEE J Y, PARK J Y, KIM D H, KIM H D, JI Y J, SEO K H. *Erigeron annuus* protects PC12 neuronal cells from oxidative stress induced by ROS-Mediated apoptosis[J]. Evidence-based Complementary and Alternative Medicine, 2020, 2020(6): 3945194.
- [11] ZHANG L Y, XU Q, LI L, LIN L, YU J L, ZHU J Y, ZHANG H, XIA G Q, ZANG H. Antioxidant and enzyme-inhibitory activity of extracts from *Erigeron annuus* flower[J]. Industrial Crops and Products, 2020, 148: 112283.
- [12] LEE T K, PARK J H, KIM B, PARK Y E, LEE J C, AHN J H, PARK C W, NOH Y H, LEE J W, KIM S S, KIM J D, WON M H. YES-10, a combination of extracts from *Clematis mandshurica* RUPR. and *Erigeron annuus* (L.) PERS, prevents ischemic brain injury in a gerbil model of transient forebrain ischemia[J]. Plants, 2020, 9(2): 154.

- [13] EZAPVA V, EZA M, LIBLOYA Z, MICHALOVA T, HENEBERG P. Stable colonization of native plants and early invaders by arbuscular mycorrhizal fungi after exposure to recent invaders from the Asteraceae family[J]. *Invasive Plant Science and Management*, 2021, 14(3): 1-29.
- [14] ZHOU Y Y, GUO M, LIU X, FU G F, ZOU D, JIN C Z. Research progress on the invasion mechanism of *Erigeron annuus*[J]. *Agricultural Biotechnology*, 2021, 10(6): 25-28, 34.
- [15] PSHEGUSOV R, NAZRANOV K, CHADAEVA V, NAZRANOV B. Invasion of *Erigeron annuus* (L.) Pers. in the hay and pasture meadows of the Greater Caucasus: modeling of the current and predicted distribution[J]. *E3S Web of Conferences*, 2020, 222(1): 4001.
- [16] 林娟, 杨柳, 王艾迪, 徐娜, 冯俊姣. 沪瀾湿地外来物种一年蓬的入侵扩散特性及防治对策[J]. *价值工程*, 2012, 31(1): 315-316.
LIN J, YANG L, WANG A D, XU N, FENG J J. Invasion and spread characteristics of alien species *Erigeron annuus* in Chanba Ecological District and control measures[J]. *Value Engineering*, 2012, 31(1): 315-316. (in Chinese)
- [17] 祝宏, 李博, 蒋轶纯. 氯氟吡氧乙酸十二酯的合成及生物活性测定[J]. *农药*, 2010, 49(11): 794-797.
ZHU H, LI B, JIANG Y C. Synthesis and bioactivity of fluroxypyr-dodecyl[J]. *Agrochemicals*, 2010, 49(11): 794-797. (in Chinese)
- [18] BOBIK K, BURCH-SMITHUR T M. Chloroplast signaling within, between and beyond cells[J]. *Frontiers in Plant Science*, 2015, 6: 781.
- [19] WICKE S, SCHNEEWEISS G M, DEPAMPHILIS C W, KAI F M, QUANDT D. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function[J]. *Plant Molecular Biology*, 2011, 76: 273-297.
- [20] WOLFE K H, LI W, SHARP P M. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs[J]. *Proc Eeding of the National Academy of Sciences*, 1987, 84(24): 9054-9058.
- [21] HOWE C J, BARBROOK A C, KOUMANDOU V L, NISBET R E R, SYMINGTON H A, WIGHTMAN T F. Evolution of the chloroplast genome[J]. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 2003, 358(1429): 99-107.
- [22] ASAF S, WAQAS M, KHAN A L, KHAN M A, KANG S M, IMRAN Q M, SHAHZAD R, BILAL S, YUN B W, LEE I J. The complete chloroplast genome of wild rice (*Oryza minuta*) and its comparison to related species[J]. *Frontiers in Plant Science*, 2017, 8: 304.
- [23] 梁凤萍, 文祥宁, 高赫一, 张颖. 菊科植物叶绿体基因组特征分析[J]. *基因组学与应用生物学*, 2018, 37(12): 5437-5447.
LIANG F P, WEN X N, GAO H Y, ZHANG Y. Analysis of chloroplast genomes features of Asteraceae species[J]. *Genomics and Applied Biology*, 2018, 37(12): 5437-5447. (in Chinese)
- [24] CHOI K S, PARK S. The complete chloroplast genome sequence of *Aster spathulifolius* (Asteraceae) genomic features and relationship with Asteraceae[J]. *Gene*, 2015, 572: 214-221.
- [25] SHEN X F, GUO S, YIN Y, ZHANG J J, YIN X M, LIANG C L, WANG Z W, HUANG B F, LIU Y H, XIAO SM. Complete chloroplast genome sequence and phylogenetic analysis of *Aster tataricus*[J]. *Molecules*, 2018, 23: 2426.
- [26] HATMAKER E A, WADL P A, RINEHART T A, CARROLL J, LANE T S, TRIGIANO R N. Complete chloroplast genome comparisons for *Pityopsis* (Asteraceae)[J]. *PLoS One*, 2020, 15(12): e0241391.
- [27] JIANG M, LI Y N, CHEN H M, WANG B. Comparative and phylogenetic analysis of the complete chloroplast genome sequences of *Lactuca raddeana* and *Lactuca sativa*[J]. *Mitochondrial DNA Part B*, 2021, 6(4): 1498-1506.
- [28] BANKEVICH A, NURK S, ANTIPOV D, GUREVICH A A, DVORKIN M, KULIKOV A S, LESIN V M, NIKOLENKO S I, PHAM S, PRJIBELSKI A D, PYSHKIN A V, SIROTKIN A V, VYAHHI N, TESLER G, ALEKSEYEV M A, PEVZNER P A. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing[J]. *Journal of Computational Biology*, 2012, 19(5): 455-477.
- [29] BOETZER M, PIROVANO W. Toward almost closed genomes with Gap Filler[J]. *Genome biology*, 2012, 13(6): 1-9.
- [30] KATOH K, STANDLEY D M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability[J]. *Molecular Biology and Evolution*, 2013, 30(4): 772-780.
- [31] GREVICH J J, DANIELL H. Chloroplast genetic engineering: recent advances and future perspectives[J]. *Critical Reviews in Plant Sciences*, 2005, 24: 83-107.
- [32] NEUHAUS H, EMES M. Nonphotosynthetic metabolism in plastids[J]. *Annual Review of Plant Biology*, 2000, 51: 111-140.
- [33] 樊守金, 郭秀秀. 植物叶绿体基因组研究及应用进展[J]. *山东师范大学学报(自然科学版)*, 2022, 37(1): 22-31.
FAN S J, GUO X X. Advances in research and application of plant chloroplast genome[J]. *Journal of Shandong Normal University (Natural Science)*, 2022, 37(1): 22-31. (in Chinese)
- [34] YU J, XIA M, XU H, HAN S, ZHANG F. The complete chloroplast genome sequence of *Neopallasia pectinata* (Asteraceae)[J]. *Mitochondrial DNA Part B*, 2021, 6(2): 430-431.
- [35] LI Z J, LIU Y Y, YANG C W, QIAN Z G, LI G D. The complete chloroplast genome sequence of *Erigeron breviscapus* and *Erigeron multiradiatus* (Asteraceae)[J]. *Mitochondrial DNA Part B*, 2019, 4(2): 3826-3827.
- [36] 苏玥, 刘娟娟, 完斌, 张鹏举, 陈正根, 宿俊吉, 王彩香. 乳苣叶绿体基因组特征及其系统发育分析[J]. *中国农业*

- 科技导报, 2021, 23(6): 33-42.
- SU Y, LIU J J, WAN B, ZHANG P J, CHEN Z G, SU J J, WANG C X. Chloroplast genome structure characteristic and phylogenetic analysis of *Mulgedium tataricum*[J]. Journal of Agricultural Science and Technology, 2021, 23(6): 33-42. (in Chinese)
- [37] PROVAN J, POWELL W, HOLLINGSWORTH P M. Chloroplast microsatellites: new tools for studies in plant ecology and evolution[J]. Trends in Ecology and Evolution, 2001, 16: 142-147.
- [38] EBRAHIMI A, ZAREI A, FARDADONBEH M Z, LAWSON S. Evaluation of genetic variability among “Early Mature” *Juglans regia* using microsatellite markers and morphological traits[J]. Peer J, 2017, 5: e3834.
- [39] GRASSI F, LABRA M, SCIENZA A, IMAZIO S. Chloroplast SSR markers to assess DNA diversity in wild and cultivated grapevines[J]. Vitis, 2002, 41(3): 157-158.
- [40] NIE X, LV S, ZHANG Y, DU X, WANG L, BIRADAR S S. Complete chloroplast genome sequence of a major invasive species, crofton weed (*Ageratina adenophora*)[J]. PLoS One, 2012, 7: e36869
- [41] TYAGI S, JUNG J A, KIM J S, WON S Y. Comparative analysis of the complete chloroplast genome of mainland *Aster spathulifolius* and other aster species[J]. Plants, 2020, 9: 568.
- [42] MARTIN G, BAURENS F C, CARDI C, AURY J M, D'HONT A. The complete chloroplast genome of banana (*Musa acuminata*, Zingiberales): insight into plastid monocotyledon evolution[J]. PLoS One, 2013, 8: e67350.
- [43] 金玲, 张颖敏, 张雪梅, 李国栋. 藿香叶绿绒蒿叶绿体全基因组特征分析[J/OL]. 分子植物育种: 1-16. [2022-06-10]. <http://kns.cnki.net/kcms/detail/46.1068.S.20220607.1742.006.html>.
- JIN L, ZHAGN Y M, ZHANG X M, LI G D. Analysis on chloroplast genomic characteristics of *Meconopsis betonici-folia*[J/OL]. Molecular Plant Breeding: 1-16. [2022-06-10]. <http://kns.cnki.net/kcms/detail/46.1068.S.20220607.1742.006.html>. (in Chinese)
- [44] 钟彩霞, 黎维平, 杨秀林, 唐明, 廖威, 陈三茂. 基于 2 个核糖体 DNA 序列的国产白酒草属、小舌菊属和歧伞菊属(菊科紫菀族)分子系统学研究(英文)[J]. 植物科学学报, 2014, 32(3): 216-227.
- ZHONG C X, LI W P, YANG X L, TANG M, LIAO W, CHEN S M. Molecular phylogeny of Chinese *Conyza*, *Microglossa* and *Thespis* (Asteraceae: Astereae) based on two nuclear ribosomal DNA regions[J]. Plant Science Journal, 2014, 32(3): 216-227. (in Chinese)