

中国特有药用植物距药黄精的叶绿体全基因组解析及黄精属种间系统发育关系分析

程文萍¹, 赵霞¹, 李依民^{1,2}, 张岗^{1,2}, 颜永刚^{1,2}, 高静^{1,2*}, 张明英^{1,2*}

(1. 陕西中医药大学药学院, 陕西省秦岭中草药应用开发工程技术研究中心, 陕西 西安 712046;
2. 陕西中医药大学, 陕西省中医药管理局“秦药”研发重点实验室, 陕西 西安 712046)

摘要: 距药黄精 *Polygonatum franchetii* Hua 为黄精属 *Polygonatum* Mill. 中国特有药用植物。本研究利用 DNBSEQ-T7 高通量测序平台测序获得距药黄精 2 个分别来自不同产地个体的叶绿体全基因组序列, 完成组装注释和特征解析, 并将其与同属已发表植物叶绿体全基因组进行了比较和系统发育分析。距药黄精 2 个个体叶绿体全基因组长度分别为 155 942 和 155 962 bp, 各包含一个大单拷贝区 (large single copy, LSC; 84 670、84 722 bp), 一个小单拷贝区 (small single copy, SSC; 18 564、18 566 bp) 和一对反向重复区 (inverted repeats, IRa/IRb; 26 354、26 337 bp); 两者均编码 113 个基因, 包括 79 个蛋白编码基因、30 个 tRNA 基因和 4 个 rRNA 基因。基因组比较分析结果表明, 距药黄精同种 2 个个体间及黄精属不同种间叶绿体全基因组序列长度、GC (guanine and cytosine) 含量、基因组成及排列顺序均高度保守, 所包含的重复序列类型及分布位置亦较为一致, IRs 区未发生明显的扩张或收缩。距药黄精 2 个个体间叶绿体基因组序列的变异小于不同种间, 种内、种间编码基因序列变异均小于非编码基因序列, IRs 区序列变异均小于 LSC、SSC 区; 共筛选到 8 条核苷酸多样性较高的种间高变异序列, 分别位于 LSC 和 SSC 区。系统发育分析结果显示, 黄精属所有物种以 100% 的支持率构成一个单系, 属内轮叶组单独构成一支, 黄精组和互叶组构成姐妹分支, 距药黄精 2 个个体聚为一支, 位于轮叶组内, 与狭叶黄精 *P. stenophyllum* Maxim. 亲缘关系最近。本研究对距药黄精叶绿体全基因组进行了全面解析并基于叶绿体系统发育基因组学分析阐明了其在黄精属的系统发育位置, 研究结果将为距药黄精的资源开发利用及黄精属药用植物种间分子鉴定、系统发育等研究提供基础。

关键词: 黄精属; 距药黄精; 叶绿体全基因组; 系统发育; 序列变异

中图分类号: R931 文献标识码: A 文章编号: 0513-4870(2023)11-3461-12

Characterization of the complete chloroplast genome of *Polygonatum franchetii* Hua, a Chinese endemic medicinal species, and phylogenetic relationships of *Polygonatum*

CHENG Wen-ping¹, ZHAO Xia¹, LI Yi-min^{1,2}, ZHANG Gang^{1,2}, YAN Yong-gang^{1,2},
GAO Jing^{1,2*}, ZHANG Ming-ying^{1,2*}

(1. College of Pharmacy and Shaanxi Qinling Application Development and Engineering Center of Chinese Herbal Medicine, Shaanxi University of Chinese Medicine, Xi'an 712046, China; 2. Key Laboratory for Research of "Qin Medicine" of Shaanxi Administration of Traditional Chinese Medicine, Shaanxi University of Chinese Medicine, Xi'an 712046, China)

Abstract: *Polygonatum franchetii* Hua is a medicinal plant endemic to China from *Polygonatum* Mill. The chloroplast genomes of two *P. franchetii* individuals sampled from two different habitats were sequenced by using

收稿日期: 2023-05-22; 修回日期: 2023-08-21.

基金项目: 国家自然科学基金资助项目 (82003898); 陕西省自然科学基金基础研究计划项目 (2022JM-458); 陕西中医药大学博士科研启动经费 (104080001); 陕西中医药大学“秦药”品质评价及资源开发学科创新团队项目 (2019-QN01).

*通讯作者 Tel: 86-29-38185165, E-mail: gaojing@sntcm.edu.cn; zhangmy@sntcm.edu.cn

DOI: 10.16438/j.0513-4870.2023-0659

the DNBSEQ-T7 high-throughput sequencing platform. After assembly and annotation, the two complete chloroplast genomes were characterized, and then comparative and phylogenetic analyses were performed with other published chloroplast genome sequences from *Polygonatum*. The whole chloroplast genomes of the two *P. franchetii* individuals were 155 942 and 155 962 bp in length, with a large single copy region (LSC, 84 670 and 84 722 bp), a small single copy region (SSC, 18 564 and 18 566 bp) and a pair of reverse repeats (IRa/IRb, 26 354 and 26 337 bp), respectively. Both of them contained 113 genes, including 79 protein-coding genes (PCGs), 30 transfer RNA (tRNA) genes, and 4 ribosomal RNA (rRNA) genes. Comparative analyses showed that the genome length, the guanine and cytosine (GC) content, genes content and order were highly conserved between the two *P. franchetii* individuals and among different *Polygonatum* species. The detected repeat sequences, including dispersed repeats, tandem repeats and simple sequence repeats (SSRs), were also relatively similar in types and positions, though showing a slightly difference in number. No significant expansion or contraction of the inverted repeat regions was found. Sequences variation between the two *P. franchetii* individuals was lower than that among different *Polygonatum* species. Besides, coding sequences (CDS) showed less divergence than noncoding sequences, and sequence divergence of IRs regions was lower than that of the LSC and SSC regions, both intraspecifically and interspecifically. Eight sequences with high nucleotide diversity among different species were screened, all of which were found located in the LSC and SSC regions. Phylogenetic inference showed that all *Polygonatum* species clustered into a monophyletic clade with a 100% bootstrap value, within which, species in section *Verticillata* formed a distinct group, section *Sibirica* and section *Polygonatum* were sister groups. The two *P. franchetii* individuals grouped together and showed the closest phylogenetic affinity to *P. stenophyllum* Maxim., belonging to the section *Verticillata*. The chloroplast genome of *P. franchetii* and its phylogenetic position in *Polygonatum* were comprehensively investigated and clearly elucidated in this study, the results may lay a foundation for the resource development and utilization of *P. franchetii*, as well as further molecular identification and phylogenetic studies of medicinal *Polygonatum* species.

Key words: *Polygonatum* Mill; *Polygonatum franchetii* Hua; chloroplast genome; phylogeny; sequence variation

叶绿体是植物光合作用和能量转化的细胞器,其自身具有一套完整且独立于核基因组之外的遗传系统,其遗传物质叶绿体基因组(chloroplast genome)在裸子植物中一般为父系遗传,而在被子植物中多为母系遗传^[1,2]。大多数被子植物的叶绿体基因组长度介于115~165 kbp,编码110~130个基因,为双链闭环状四分体结构,由一个大单拷贝区(large single copy, LSC; 81~90 kbp)、一个小单拷贝区(small single copy, SSC; 18~20 kbp)和一对反向重复区(inverted repeats, IRa/b; 20~30 kbp)构成^[3]。叶绿体基因组因单亲遗传、基因组成和排列顺序高度保守、碱基变异速率适中且具有多拷贝和基因组较小等特点^[4,5],基于叶绿体全基因组序列的系统发育基因组学分析已被广泛应用于药用植物种间分类鉴定及系统发育重建研究^[6-8]。

天门冬科 Asparagaceae 黄精属 *Polygonatum* Mill. 为多年生草本植物,全世界约有60多种,中国分布约39种^[9],其中31种具有药用价值^[10],是中药材的重要来源。该属植物因含多糖、生物碱、黄酮类和甾体皂苷类等多种活性成分,具有调节免疫、降血糖血脂、抗肿瘤、抗菌消炎等药理作用^[11],被广泛应用于临床用药及保健品开发。黄精属自建立以来,其属内不同种间的

系统发育关系长期受到关注。依据宏观形态特征、花粉形态特征、染色体数目及分子遗传等证据,开展了大量有关黄精属植物种间分类鉴定、系统演化关系等相关研究,有多个不同的属下分类系统被提出^[12-17]。最新的分子系统发育研究结果将黄精属划分为轮叶组 section *Verticillata*、互叶组 section *Polygonatum* 和黄精组 section *Sibirica*,然而,由于高度的种间、种内形态多样性及广泛的种间地理分布重叠,属下部分种间划分及系统发育关系仍有待进一步研究澄清^[18-20]。

距药黄精为黄精属中国特有物种,主要分布在我国陕西、湖北、湖南、四川等地^[9]。其根状茎入药,具有补气养阴,健脾益肾,润肺等功效^[10],在我国湖南、湖北等地常作为中药材黄精入药使用^[21]。研究^[15]曾依据苞片大小与质地、叶序类型、花被筒长度、花被形状、花药长短以及子房形状等特征将国产黄精属划分为短筒系 series *Alte-lobata*、互叶系 series *Alternifolia* Baker、苞叶系 series *Bracteata* Kom.、独花系 series *Hookeriana*、滇黄精系 series *Kingiana*、对叶系 series *Oppositifolia*、点花系 series *Punctata* 和轮叶系 series *Verticillata*,其中距药黄精因叶互生等特征被置于互叶系, Meng等^[16]基于4个叶绿体片段(*rbcL*、*trnK*、*psbA-trnH*、*trnC-petN*)对

黄精属 30 个种的分子系统发育分析结果曾将距药黄精划分至轮叶组, 但其与组内其他物种之间的系统发育关系并未得到解决, Wang 等^[22]及 Zhang 等^[23]基于叶绿体基因组序列的系统发育分析结果存在争议, 分别将距药黄精置于互叶组或轮叶组, 其他关于黄精属的分类及分子系统学研究鲜有涉及距药黄精。此外, 距药黄精的叶绿体全基因组仍有待进一步研究解析。

本文利用高通量测序方法获取距药黄精同种不同产地个体的叶绿体全基因组序列, 对其基因组成、序列特征及种内变异进行分析比较, 并与同属已发表植物叶绿体基因组序列进行种间比较基因组学和系统发育分析, 旨在全面解析距药黄精叶绿体全基因组序列与结构特征, 探讨其序列变异情况, 并澄清距药黄精在黄精属内的系统发育位置, 为距药黄精遗传多样性保护和资源开发利用及黄精属药用植物分子鉴定、系统演化等相关研究提供基础。

材料与amp;方法

实验材料 采集野外自然状态下生长的距药黄精植株新鲜幼嫩叶片, 放入硅胶中快速脱水干燥, -20 °C 保存, 用于基因组 DNA 提取。本研究距药黄精 2 个个体分别采自陕西省安康市石泉县 (33°2'18.27"N, 108°14'52.39"E, *P. franchetii*_SXSQ) 和四川省巴中市通江县 (31°52'18.48"N, 107°29'33.62"E, *P. franchetii*_SCTJ), 由陕西中医药大学药学院王继涛高级实验师鉴定。凭证标本 (采集号: ZMY_P_17、ZMY_009) 保存于陕西中医药大学中药标本馆。

基因组 DNA 提取与测序 利用植物基因组 DNA 提取试剂盒 (TIANGEN DNasecure Plant Kit, DP320) 提取叶片基因组总 DNA, 用 1% 的琼脂糖凝胶电泳和 Qubit 3.0 (Invitrogen, Carlsbad, California, America) 荧光定量仪检测提得的 DNA 质量及浓度。构建长度约 500 bp 的插入片段文库, 利用 DNBSEQ-T7 测序平台进行双端测序, 测序读长为 150 bp。建库工作在中国西南野生生物种质资源库分子生物学实验平台完成, 测序由深圳华大基因股份有限公司武汉分公司完成。

叶绿体基因组序列组装拼接及注释 利用 NGS QC Tool Kit^[24]对测序得到的原始序列进行过滤 (参数设置为: cut-off value for percentage of read length=80, cut-off value for PHRED quality score=30), 除去接头和低质量序列。以 NCBI 数据库中已收录的黄精 *P. sibiricum* Delar. ex Redoute 叶绿体全基因组序列 (MZ029093) 作为参考, 利用 GetOrganelle 程序^[25]对过滤后的序列进行组装, 再通过 Bandage 软件^[26]将组装得到的序列重叠群 (contigs) 拼接成环并去除冗余序

列, 将拼接结果与上述参考序列 (MZ029093) 在 Geneious 软件中比对, 确定两个反向重复区的方向。在 Geneious 软件中将原始序列重新映射到组装完成的叶绿体基因组序列上, 评估和验证各 contigs 拼接处及 LSC/IRa、IRa/SSC、SSC/IRb 和 IR/LSC 四个节点拼接的正确性, 最终完成距药黄精 2 个不同个体完整叶绿体全基因组序列的组装拼接。

以上述黄精叶绿体基因组序列 (MZ029093) 的注释结果为参考, 利用 PGA (plastid genome annotator) 软件^[27]对距药黄精的叶绿体全基因组序列进行注释, 并在 Geneious 软件中手动检查校正蛋白编码基因的起始、终止位置和基因内含子、外显子边界。利用在线软件 tRNAscan-SE (<http://lowelab.ucsc.edu/tRNAscan-SE/>)^[28]确定 tRNA 基因的边界。最后, 利用 Organellar Genome DRAW (OGDRAW, <https://chlorobox.mpimp-goelm.mpg.de/OGDraw.html>)^[29]绘制叶绿体全基因组图谱。组装并注释完成后的叶绿体全基因组序列提交至 NCBI 数据库, 检索号分别为 ON101705 (*P. franchetii*_SXSQ)、OM219012 (*P. franchetii*_SCTJ)。

重复序列分析检测 利用在线工具 REPuter (<https://bibiserv.cebitec.uni-bielefeld.de/reputer/manual.html>) 检测距药黄精叶绿体全基因组中的散在重复序列 (dispersed repeats), 包括正向重复 (forward match)、反向重复 (reverse match)、互补重复 (complement match) 和回文重复 (palindromic match), 参数设置为: 最小重复序列长度 = 30 bp, 重复序列间的相似度 > 90% (hamming distance = 3, maximum computed repeats = 90, minimal repeat size = 30)^[30]。利用 Tandem Repeats Finder (<https://tandem.bu.edu/trf/trf.unix.help.html>) 检测串联重复序列 (tandem repeats), 参数设置选择默认值^[31]。

为探讨距药黄精不同个体间及其与黄精属不同物种间叶绿体全基因组序列中简单重复序列 (simple sequence repeats, SSRs) 类型、数量及分布模式等的差异, 下载 NCBI 数据库 (<https://www.ncbi.nlm.nih.gov/genome>) 已开放获取的黄精属 25 个物种的叶绿体全基因组序列 (表 1), 利用 MISA 软件^[32] (<https://webblast.ipk-gatersleben.de/misa/>) 对距药黄精 2 个个体及上述 25 个物种 (即共 26 种 27 个个体) 叶绿体全基因组序列的 SSRs 进行检测。参数设置为: 单核苷酸 (mononucleotide) 重复次数不小于 10, 二核苷酸 (dinucleotide) 重复次数不小于 5, 三核苷酸 (trinucleotide) 和四核苷酸 (tetranucleotide) 重复次数均不小于 4, 五核苷酸 (pentanucleotide) 和六核苷酸 (hexanucleotide) 重复次数不小于 3, 两个 SSRs 之间的最小距离为 100 bp^[33]。

叶绿体全基因组序列变异比较分析 为进一步探讨距药黄精种内不同个体间及黄精属不同种间叶绿体全基因组序列的核苷酸变异情况, 利用 mVISTA 软件^[34] (<http://genome.lbl.gov/vista/mvista/submit.shtml>) 以 LAGAN 模型对上述 26 种 27 个个体进行叶绿体全基因组序列比较分析。同时, 将这 27 个个体叶绿体全基因组序列 (去掉一个 IR 区) 构建多序列比对矩阵, 利用 MAFFT 软件完成序列比对, 利用 DnaSP 软件进行滑动窗口分析 (sliding window analysis)。滑动窗口 (windows length) 设置为 200 bp, 步长 (step size) 为 100 bp^[35], 计算核苷酸多样性 (nucleotide diversity)。此外, 利用在线工具 Irscope (<https://irscope.shinyapps.io/irapp/>) 分析其叶绿体基因组 IRs 区收缩/扩张情况。

系统发育分析 将本研究新测序获得的距药黄精 2 个个体与 NCBI 数据库中下载的黄精属其他 25 个物种的叶绿体全基因组序列构建多序列矩阵, 以黄精属近缘类群天门冬科黄精族 Polygonateae 异黄精属 *Heteropolygonatum* M. N. Tamura & Ogisu^[18] 的 *H. alter-*

nicirrhosum (Hand.-Mazz.) Floden (NC_058552) 作为外类群, 利用 MAFFT 软件完成序列比对, 并在 Geneious 中手动对比对结果进行校正调整, 利用 CIPRES (<http://www.phylo.org/>) 在线服务器中的 RaxML (RAXML-HPC BlackBox) 程序以最大似然法 (maximum likelihood, ML) 进行系统发育重建分析, 核苷酸替换模型选择 GTR+G, 其他参数选择默认设置, 系统发育树节点处的靴带支持率 (bootstrap values, BS) 通过进行 1 000 次自展重复分析计算得到。

结果

1 距药黄精叶绿体全基因组拼接与注释

距药黄精 2 个个体叶绿体全基因组长度分别为 155 942 和 155 962 bp, 均为由一个大单拷贝区 (84 670、84 722 bp)、一个小单拷贝区 (18 564、18 566 bp) 和一对反向重复区 (26 354、26 337 bp) 组成的环状四分体结构 (图 1)。整体 GC (guanine and cytosine) 含量分别为 37.7% 和 37.6%, LSC、SSC 及 IRs 区的 GC 含量

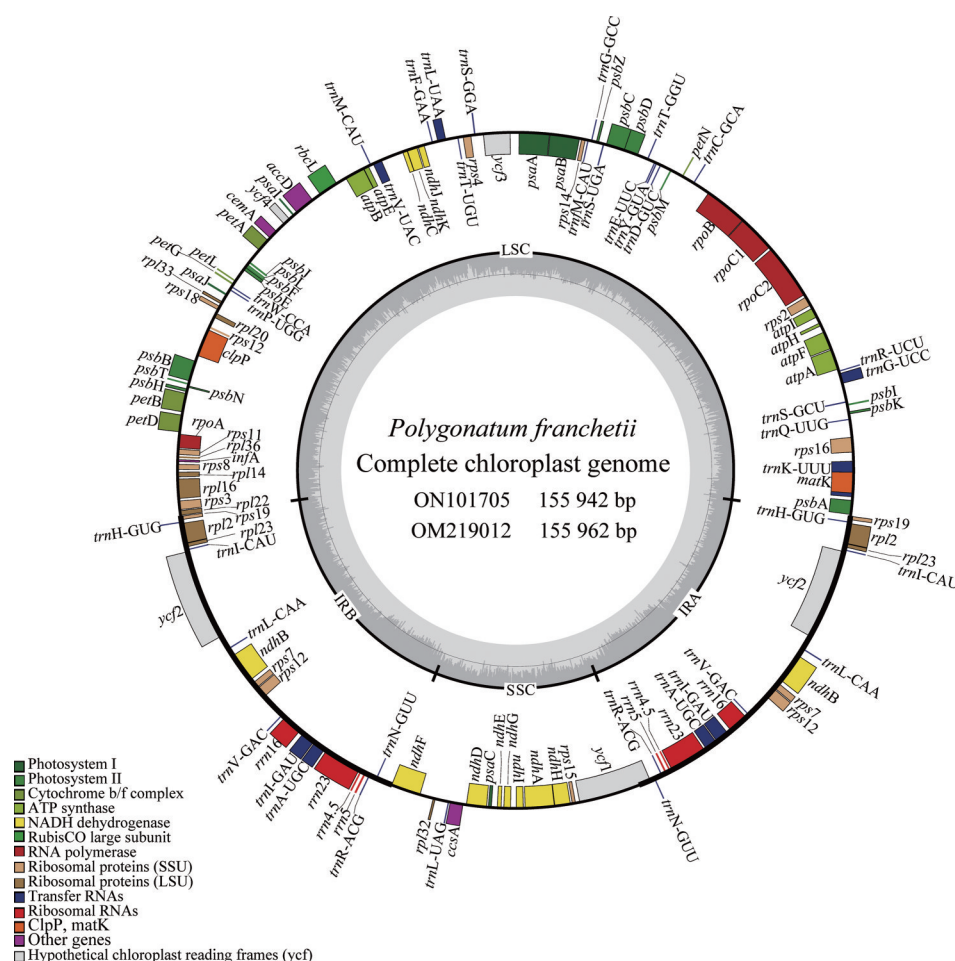


Figure 1 Chloroplast genome map of the two *P. franchetii* individuals. Genes inside of the large circle are transcribed clockwise and those outside are transcribed counterclockwise. All genes are color-coded based on their functions. The grey area in the inner circle represents the guanine and cytosine (GC) content

依次分别为35.7%和35.7%、31.6%和31.5%、43.0%和43.0%(表1)。两者均注释得到113个基因,包括79个蛋白编码基因(其中*infA*为假基因),30个转运RNA基因(tRNA)以及4个核糖体RNA基因(rRNA)(表2)。其中,8个蛋白编码基因(*rps7*、*rps12*、*rps19*、*rpl2*、*rpl23*、*ycf1*、*ycf2*和*ndhB*)、8个tRNA基因(*trnH-GUG*、*trnI-CAU*、*trnI-GAU*、*trnL-CAA*、*trnV-GAC*、*trnA-UGC*、*trnR-ACG*和*trnN-GUU*)以及4个rRNA基因(*rrn4.5*、*rrn5*、*rrn16*和*rrn23*)在两个反向重复区出现重复。由于*ycf1*基因跨SSC/IRa边界,在IRb/SSC边界处存在其对应长度的序列拷贝。此外,*rpl2*、*rpl16*、*rps16*、*rpoC1*、*atpF*、*petB*、*petD*、*ndhA*、*ndhB*、*trnA-UGC*、*trnG-UCC*、*trnI-GAU*、*trnK-UUU*、*trnL-UAA*和*trnV-UAC*基因均含有一个内含子,*ycf3*、*rps12*和*clpP*各含有两个内含子(表2)。

2 重复序列分析

距药黄精2个个体叶绿体全基因组序列(ON101705、OM219012)分别检测到23和22条长度不小于30 bp的散在重复序列。其中,正向重复序列均为8条,回文重复序列分别为15和14条,反向重复和互补重复序列均未检测到。两者分别检测到39和40条串联重复序列,长度范围分别为25~91 bp和25~

145 bp。这些串联重复序列主要分布在LSC区(22、26条)和SSC区(9、7条)。此外,两者SSRs的含量分别为63和66个,其中单核苷酸重复分别为42和44个,二核苷酸重复为16和17个,三核苷酸和五核苷酸重复分别均为3和2个,四核苷酸和六核苷酸重复未检测到。在所有SSRs中,出现频率最高的均为A、T单核苷酸重复(65.08%、66.67%),其次是AT、TA二核苷酸重复(20.63%、21.21%)(表3)。

黄精属26种27个个体叶绿体全基因组序列中SSRs的数量、类型和分布分析比较结果显示,不同个体叶绿体基因组中检测到的SSRs总数在57~67之间(图2A),均以单核苷酸重复最多(含量范围60.0%~69.7%),其次为二核苷酸重复,3、4、5、6核苷酸重复数量为1~5个(含量均小于10%)。单碱基重复单元A、T,二碱基重复单元AT、GA、TA、TC,三碱基重复单元TTA和五碱基重复单元TTTCG是所有个体共有的SSRs类型(图2B)。这些SSRs在不同个体叶绿体基因组LSC区、SSC区和IR区的数量依次为46~56个、6~11个和4~6个(图2C),且主要分布在基因间隔区(intergenic spacer, IGS)中(36~46个,占比59.1%~69.4%),外显子区(extron)和内含子(intron)区含量较少,分别为8~14个和9~13个(图2D)。

Table 1 Chloroplast genome features of the two *P. franchetii* individuals and other 25 *Polygonatum* species. LSC: Large single copy; SSC: Small single copy; IR: Inverted repeats; CDS: Coding sequence; PCGs: Protein coding genes; tRNA: Transfer RNA; rRNA: Ribosomal RNA

Species	Accession	Genome	Length/bp				CDS	Genome	GC/%				Total	Gene number		
			LSC	SSC	IR	CDS			LSC	SSC	IR	CDS		PCGs	tRNA	rRNA
<i>P. franchetii</i>	ON101705	155 942	84 670	18 564	26 354	79 677	37.7	35.7	31.6	43.0	38.0	113	79	30	4	
<i>P. franchetii</i>	OM219012	155 962	84 722	18 566	26 337	79 659	37.6	35.7	31.5	43.0	38.0	113	79	30	4	
<i>P. acuminatifolium</i>	NC_058569	155 354	84 271	18 455	26 314	79 572	37.7	35.7	31.6	43.0	38.0	113	79	30	4	
<i>P. cirrhifolium</i>	NC_053687	155 583	84 412	18 427	26 372	79 626	37.7	35.7	31.5	42.9	38.0	113	79	30	4	
<i>P. curvistylum</i>	NC_058554	155 939	84 563	18 546	26 415	79 809	37.6	35.7	31.5	42.9	38.0	113	79	30	4	
<i>P. cyrtonema</i>	MW248135	155 816	84 618	18 440	26 379	79 629	37.7	35.7	31.6	42.9	38.0	113	79	30	4	
<i>P. filipes</i>	NC_058555	155 334	84 280	18 454	26 300	79 536	37.7	35.7	31.6	43.0	38.0	113	79	30	4	
<i>P. hirtum</i>	NC_058556	155 490	84 385	18 419	26 343	79 572	37.7	35.7	31.6	42.9	38.0	113	79	30	4	
<i>P. hookeri</i>	NC_058557	155 976	84 600	18 546	26 415	79 809	37.6	35.7	31.5	38.0	38.0	113	79	30	4	
<i>P. humile</i>	NC_058558	155 015	83 999	18 388	26 314	79 572	37.7	35.8	31.7	43.0	38.0	113	79	30	4	
<i>P. hunanense</i>	NC_058226	155 609	84 438	18 427	26 372	79 626	37.7	35.7	31.5	42.9	38.0	113	79	30	4	
<i>P. inflatum</i>	NC_058559	154 898	84 270	18 454	26 087	79 575	37.7	35.7	31.6	43.0	38.0	113	79	30	4	
<i>P. involucreatum</i>	NC_058560	155 370	84 280	18 450	26 320	79 572	37.7	35.7	31.6	43.0	38.0	113	79	30	4	
<i>P. jinzhaiense</i>	NC_058561	155 510	84 463	18 292	26 379	79 626	37.7	35.7	31.7	42.9	38.0	113	79	30	4	
<i>P. kingianum</i>	NC_047406	155 399	84 234	18 424	26 372	79 620	37.7	35.7	31.5	42.9	38.0	113	79	30	4	
<i>P. macropodium</i>	NC_058562	154 610	83 554	18 464	26 296	79 536	37.7	35.8	31.6	43.0	38.0	113	79	30	4	
<i>P. multiflorum</i>	NC_058563	155 495	84 391	18 418	26 343	79 572	37.7	35.7	31.6	42.9	38.0	113	79	30	4	
<i>P. nodosum</i>	NC_058564	155 205	84 143	18 422	26 320	79 560	37.7	35.7	31.6	43.0	38.1	113	79	30	4	
<i>P. odoratum</i>	NC_050926	154 468	83 417	18 457	26 297	79 536	37.8	35.8	31.6	43.0	38.1	113	79	30	4	
<i>P. prattii</i>	NC_058565	155 915	84 538	18 547	26 415	79 809	37.7	35.7	31.5	42.9	38.0	113	79	30	4	
<i>P. punctatum</i>	NC_058566	155 657	84 542	18 423	26 346	79 644	37.7	35.7	31.5	43.0	38.0	113	79	30	4	
<i>P. sibiricum</i>	MZ029093	155 585	84 460	18 417	26 354	79 626	37.7	35.7	31.7	43.0	38.1	113	79	30	4	
<i>P. stenophyllum</i>	NC_035995	156 028	84 677	18 561	26 395	79 764	37.7	35.7	31.6	43.0	38.0	113	79	30	4	
<i>P. tessellatum</i>	NC_058567	155 688	84 488	18 564	26 318	79 638	37.6	35.7	31.4	42.9	38.0	113	79	30	4	
<i>P. uncinatum</i>	NC_058568	155 797	84 614	18 531	26 326	79 686	37.7	35.7	31.6	42.9	38.0	113	79	30	4	
<i>P. verticillatum</i>	MZ150865	155 589	84 242	18 523	26 412	79 794	37.7	35.7	31.6	42.9	38.0	113	79	30	4	
<i>P. zanlanscianense</i>	NC_058227	155 911	84 650	18 431	26 415	79 797	37.6	35.6	31.5	42.9	38.0	113	79	30	4	

Table 2 Gene composition of the chloroplast genomes of the two *P. franchetii* individuals. ×2: Genes duplicated in the IR regions; *: Genes with one intron; **: Genes with two introns; #: Pseudogene. NADH: Nicotinamide adenine dinucleotide; ATP: Adenosine triphosphate; CoA: Coenzyme A

Category	Group	Name	
Self-replication	Ribosomal RNA	<i>rrn4.5</i> ×2, <i>rrn5</i> ×2, <i>rrn16</i> ×2, <i>rrn23</i> ×2	
	Transfer RNA	<i>trnC</i> -GCA, <i>trnD</i> -GUC, <i>trnE</i> -UUC, <i>trnF</i> -GAA, <i>trnG</i> -UCC*, <i>trnG</i> -GCC, <i>trnK</i> -UUU*, <i>trnL</i> -UAA*, <i>trnL</i> -UAG, <i>trnM</i> -CAU, <i>trnP</i> -UGG, <i>trnQ</i> -UUG, <i>trnR</i> -UCU, <i>trnS</i> -GCU, <i>trnS</i> -GGA, <i>trnS</i> -UGA, <i>trnT</i> -UGU, <i>trnT</i> -GGU, <i>trnV</i> -UAC*, <i>trnY</i> -GUA, <i>trnW</i> -CCA, <i>trnM</i> -CAU, <i>trnH</i> -GUG×2, <i>trnA</i> -UGC*×2, <i>trnI</i> -CAU×2, <i>trnI</i> -GAU*×2, <i>trnL</i> -CAA×2, <i>trnN</i> -GUU×2, <i>trnR</i> -ACG×2, <i>trnV</i> -GAC×2	
	Ribosomal protein (small subunit)	<i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> ×2, <i>rps8</i> , <i>rps11</i> , <i>rps12</i> **×2, <i>rps14</i> , <i>rps15</i> , <i>rps16</i> *, <i>rps18</i> , <i>rps19</i> ×2	
	Ribosomal protein (large subunit)	<i>rpl2</i> *×2, <i>rpl14</i> , <i>rpl16</i> *, <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> ×2, <i>rpl32</i> , <i>rpl33</i> , <i>rpl36</i>	
	RNA polymerase	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> *, <i>rpoC2</i>	
	Photosynthesis	Subunits of photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i> , <i>ycf3</i> ***, <i>ycf4</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 cytochrome	<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 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>	
Others		Maturase <i>matK</i> Envelope membrane protein <i>cemA</i> Subunit of acetyl-CoA <i>accD</i> Synthesis gene <i>ccsA</i> ATP-dependent protease <i>clpP</i> ** Translational initiation factor <i>infA</i> #	
Unknown function	Proteins of unknown function <i>ycf1</i> ×2, <i>ycf2</i> ×2		

Table 3 Simple sequence repeats (SSRs) types and numbers in the chloroplast genomes of the two *P. franchetii* individuals

Repeat type	Type	Repeat	Number	
			ON101705	OM219012
Mononucleotide	A/T	10	24	26
		11	9	11
		12	2	2
		13	5	3
		15	1	2
		11	1	0
Dinucleotide	C	5	3	3
		5	10	11
	GA/TC	6	2	2
		7	1	1
Trinucleotide	TAA/TTA	4	1	1
		5	1	1
		4	1	1
Pentanucleotide	CGAAA/TTTCG	3	2	2

3 叶绿体全基因组序列种间、种内比较分析

所有 26 种 27 个个体叶绿体全基因组序列 GC 含量高度保守, 所编码的基因数量、类别、排列顺序及基因转录方向均完全相同, 但基因组总长度及 LSC 区、SSC 区和 IR 区长度分别存在一定差异, 且种内差异小于不同物种间。其中, 叶绿体全基因组总长度变化范围为 154 468 [玉竹 *P. odoratum* (Mill.) Druce]~156 028 bp (狭叶黄精 *P. stenophyllum* Maxim.), 相差 1 560 bp, LSC 区、SSC 区和 IR 区长度最大差异依次分别为 1 305、274 和 328 bp, 编码区最大长度差异为 273 bp (表 1)。同

时, 距药黄精同种 2 个个体间及其与黄精属其他 25 种植物叶绿体全基因组序列间 IR/SC 边界均高度保守, LSC/IRb (JLB)、SSC/IRb (JSB)、SSC/IRa (JSA) 和 LSC/IRa (JLA) 4 个边界的侧翼基因完全相同, 尽管各侧翼基因在不同区域的分布长度存在微小差异, IRs 区没有出现明显扩张或收缩情况 (图 3)。此外, 序列在种间、种内均间存在一定程度变异。其中, LSC 区、SSC 区变异程度明显高于 IR 区, 非编码序列的变异高于编码区序列, 且种间变异明显高于距药黄精 2 个不同个体间, 变异较高的区域多集中在 LSC 区, 包括基因间隔区 *rps16_trnQ*-UUG、*atpF_atpH*、*trnT*-UGU、*trnL*-UAA、*ycf4_cemA* 和蛋白编码基因 *psbA* 等 (图 4)。

距药黄精 2 个不同个体间叶绿体全基因组序列的核苷酸多样性指数 (Pi) 变化范围为 0~0.035, 平均值为 0.000 67, 其中, LSC、SSC、IR 区序列的 Pi 变化范围依次为 0~0.035、0~0.010、0~0.005, 平均值为 0.000 78、0.001 30、0.000 04。黄精属不同种间叶绿体全基因组序列的 Pi 值范围为 0~0.038, 平均值为 0.003 54, 其中 LSC、SSC、IR 区 Pi 平均值依次为 0.003 94、0.005 54、0.001 07。黄精属不同种间共筛选到 8 条 Pi 值 > 0.020 的高变异序列, 分别位于 LSC 区、SSC 区, 依次为 *trnS*-GCU_(34 bp)+*trnS*-GCU、*trnS*-GCU、*trnS*-GCU、*rpoB_trnC*-GCA、*trnM*-CAU_(40 bp)+*trnM*-CAU、*atpE*+*atpE*_(16 bp)、*petA_psbJ*+*psbJ*_(32 bp)、*rpl32_trnL*-UAG、

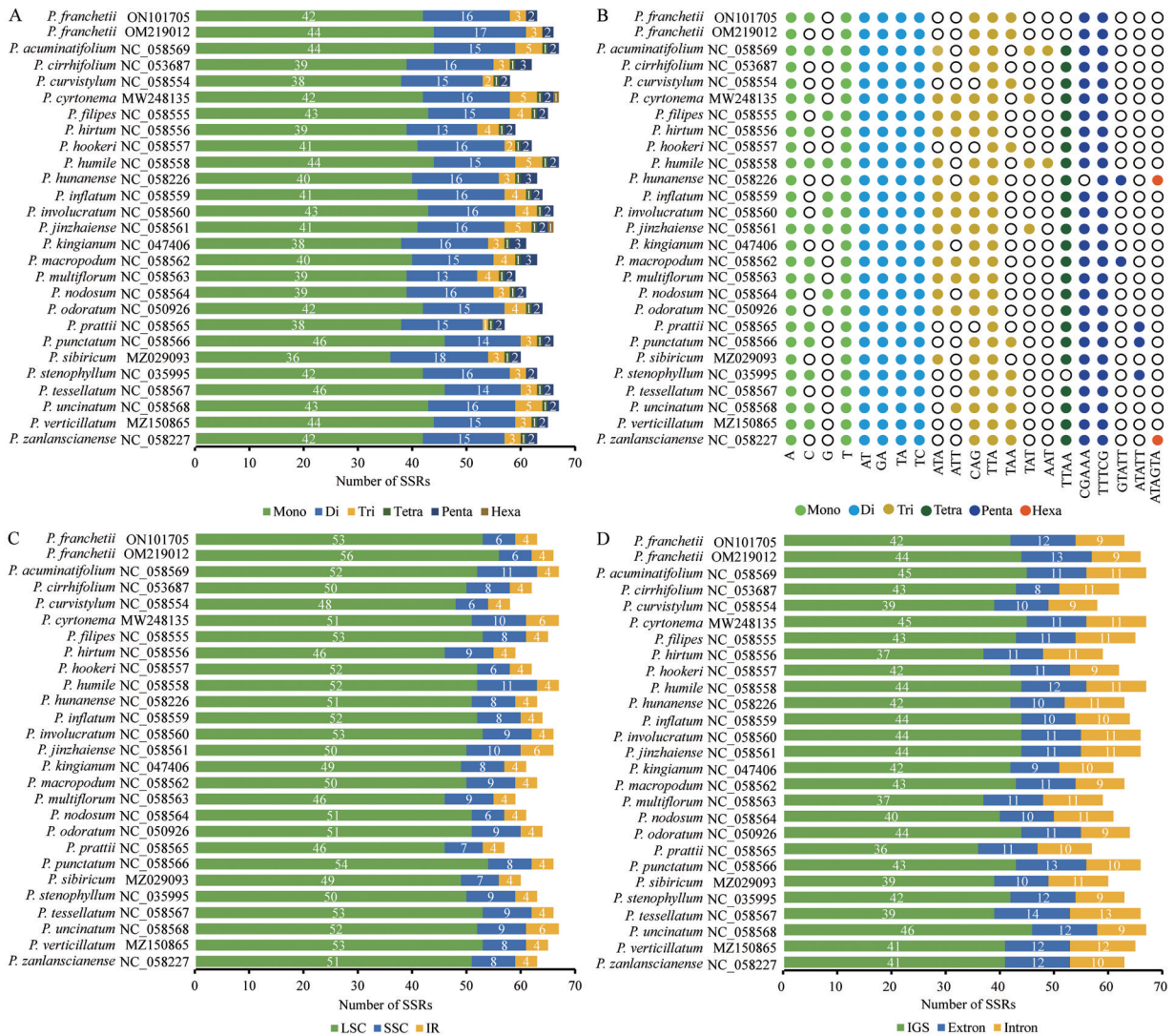


Figure 2 Numbers, types and distribution patterns of SSRs detected in the chloroplast genomes of the two *P. franchetii* individuals and other 25 *Polygonatum* species. A: Number of each of the six SSRs; B: Distribution patterns of SSRs; C: Number of SSRs distributed in LSC, SSC and IR regions; D: Number of SSRs distributed in IGS, exon and intron. Mono: Mononucleotide; Di: Dinucleotide; Tri: Trinucleotide; Tetra: Tetranucleotide; Penta: Pentanucleotide; Hexa: Hexanucleotide; IGS: Intergenic spacer

*ccsA*_(46 bp)+*ccsA_ndhD* 和 *ccsA_ndhD*+*ndhD*_(32 bp) (图5)。

4 系统发育分析

系统发育分析结果显示, 黄精属所有物种构成一个单系, 属下划分为两个次级分支, 其中, 轮叶组以100%的支持率单独构成一支, 黄精组和互叶组以100%的支持率构成姐妹群。距药黄精位于轮叶组, 与狭叶黄精亲缘关系最近, 支持率亦为100% (图6)。

讨论

作为黄精属中国特有物种之一, 距药黄精根状茎在其产地常作为中药材黄精入药使用, 具有一定的研究和开发利用价值。本研究首次对距药黄精叶绿体全基因组进行了全面解析, 并对其种内2个不同产地个

体间及黄精属不同种间叶绿体全基因组序列进行了比较分析。结果显示, 距药黄精种内2个不同个体间及黄精属不同物种间叶绿体全基因组高度保守, 编码基因数量与类别及排列顺序完全一致, 序列的GC含量及长度(包括总长度及LSC、SSC、IRs区和编码基因序列的长度)存在较小差异。其中, 距药黄精种内2个不同个体间叶绿体全基因长度仅相差20 bp, 不同区域长度差异范围为2(SSC区)~52 bp(LSC区), 而包括距药黄精在内的黄精属26个不同物种间叶绿体全基因组序列总长度差异最大值为1 560 bp, LSC区、SSC区和IR区长度差异最大值依次分别为1 305、274和328 bp, 编码区长度相差273 bp, 说明不同种间叶绿体全基因组序列长度的差异主要来自LSC区, 且差异主要出现

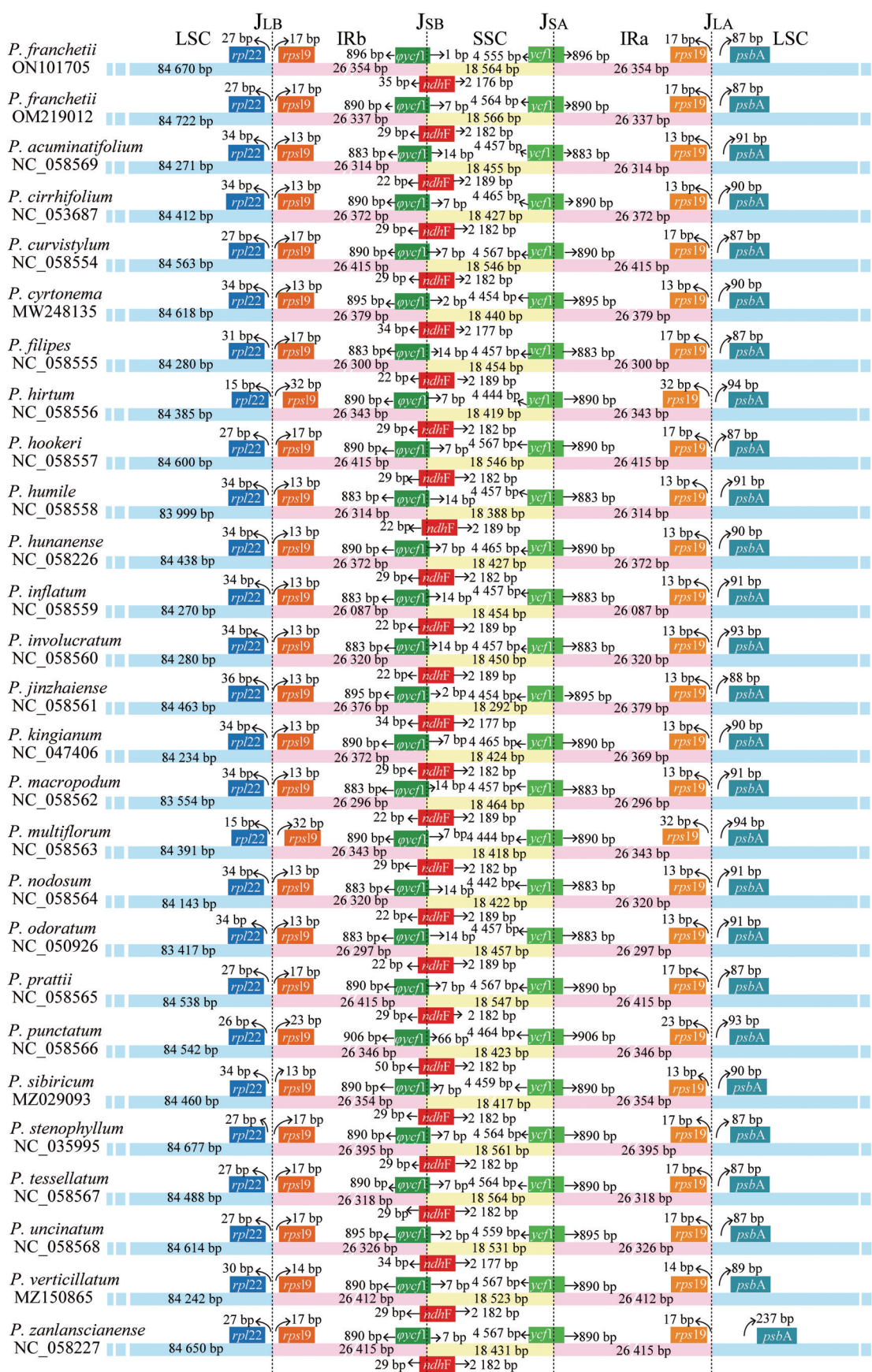


Figure 3 Comparison of the borders of LSC, SSC and IRa/b regions of the whole chloroplast genomes among the two *P. franchetii* individuals and other 25 *Polygonatum* species. JLB: LSC/IRb border; JSB: SSC/IRb border; JSA: SSC/IRa border; JLA: LSC/IRa border

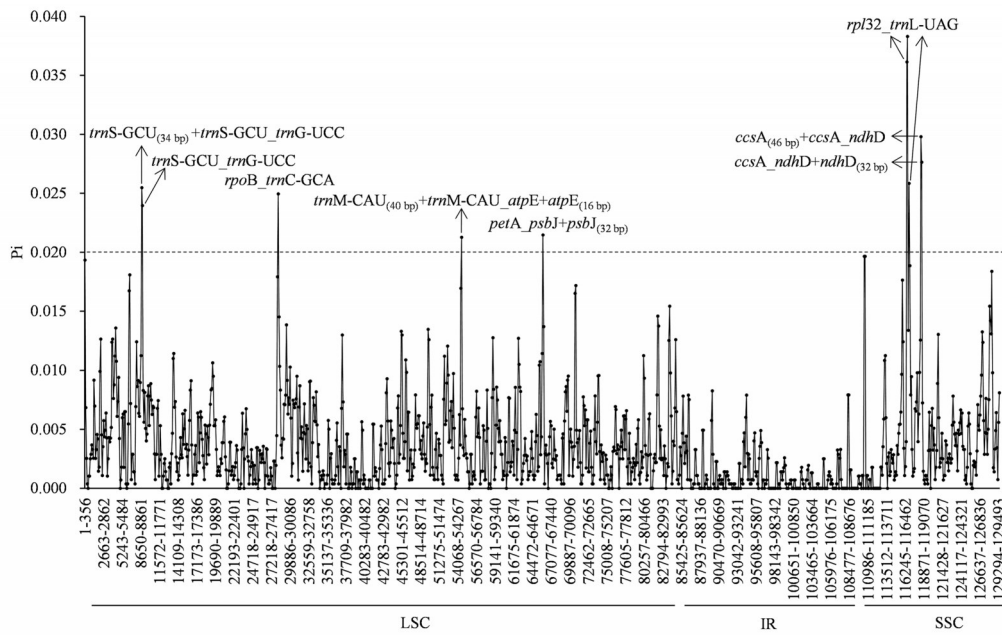


Figure 5 The nucleotide diversity of 26 *Polygonatum* species. Window length: 200 bp; Step size: 100 bp. Pi: Nucleotide diversity

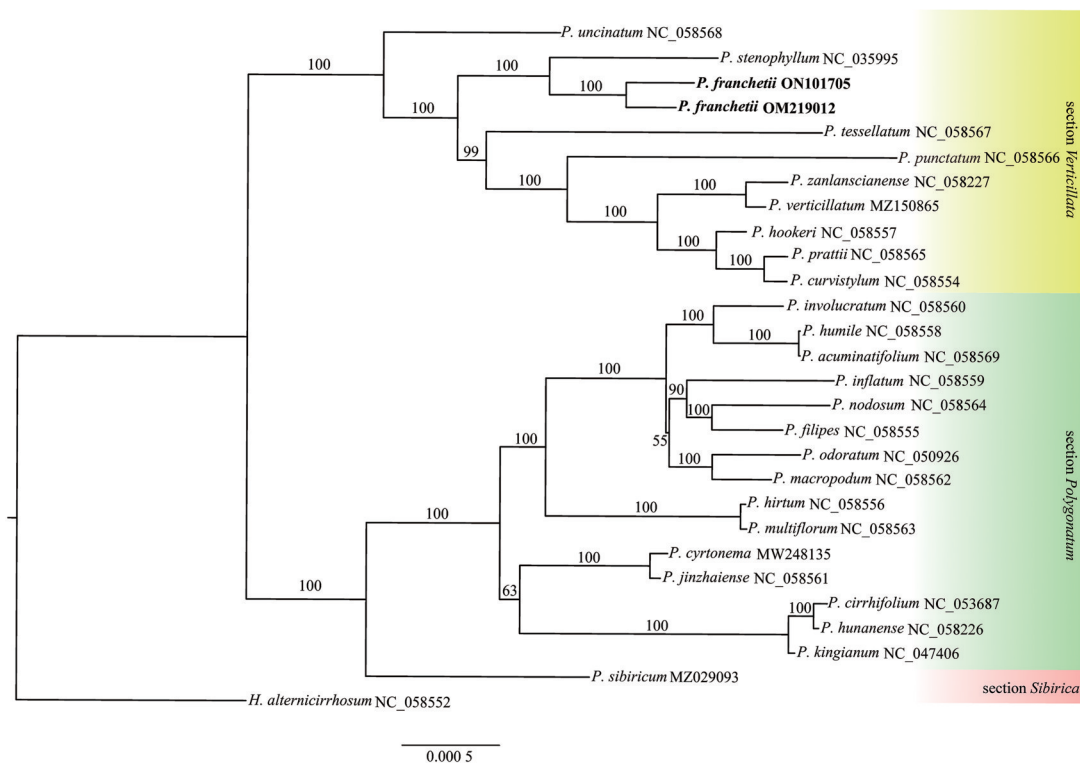


Figure 6 Phylogenetic relationships of 26 *Polygonatum* species (27 individuals) inferred using the whole chloroplast genome sequences based on the maximum likelihood (ML) method. Numbers beside each node are bootstrap values (BS)

在非编码区序列。这也与芍药属 *Paeonia* L.^[36]、柴胡属 *Bupleurum* L.^[33]等其他被子植物类群叶绿体基因组序列长度的种间变异情况相似。

简单重复序列指以长度为1~6 bp的核苷酸为重复单位的首尾串联重复DNA序列,具有多态性高、覆

盖整个基因组、稳定性好等特点,常被开发为分子标记应用于构建基因连锁图分子辅助育种及群体遗传学等相关研究^[37]。本研究结果显示,包括距药黄精及在内的黄精属26种27个个体叶绿体全基因组序列中的SSRs主要类型均为单核苷酸重复(占总数量的

60.00%~69.70%), 其次为二核苷酸 SSRs (22.03%~30.00%), 三、四、五、六核苷酸 SSRs 的含量仅为 0~7.58%。其中, 单核苷酸重复中 A、T 重复占 98.12%, 二核苷酸重复以 AT、TA 为主 (80.53%), 由 A、T 碱基组合组成的三核苷酸重复 (ATA、ATT、TTA、TAA、TAT、AAT) 占三核苷酸 SSRs 总数的 74.47%, 四核苷酸重复全部为 A、T 碱基组合 (TTAA)。说明黄精属植物叶绿体基因组序列中的 SSRs 有明显的 A、T 碱基使用偏好, 与大多数被子植物类群叶绿体基因组序列中 SSRs 的类型相似^[38-41]。

序列变异比较分析结果表明, 距药黄精 2 个不同个体间叶绿体全基因组序列高度保守, 黄精属不同种间存在一定程度的序列变异。其中, LSC 区、SSC 区变异程度显著高于 IR 区, 非编码序列变异高于编码基因序列。核苷酸多样性分析结果进一步证明距药黄精 2 个不同个体间叶绿体全基因组序列的变异小于黄精属不同物种间, 且序列变异主要发生在 LSC 和 SSC 区, 与其他被子植物类群叶绿体基因组序列变异模式相似^[33,36,42]。此外, 筛选到 8 条种间变异较高的区域, 分别为 *trnS-GCU*_(34 bp)+*trnS-GCU_trnG-UCC*、*trnS-GCU_trnG-UCC*、*rpoB_trnC-GCA*、*trnM-CAU*_(40 bp)+*trnM-CAU_atpE+atpE*_(16 bp)、*petA_psbJ+psbJ*_(32 bp)、*rpl32_trnL-UAG*、*ccsA*_(46 bp)+*ccsA_ndhD* 和 *ccsA_ndhD+ndhD*_(32 bp), 分布在 LSC 区和 SSC 区。这些高变异序列可作为黄精属种间分子鉴定的特异性候选分子标记。

本研究基于叶绿体全基因组序列的系统发育分析结果以 100% 的支持率将黄精属 26 个物种划分为 3 个单系分支, 分别对应轮叶组、黄精组和互叶组, 其中黄精组和互叶组构成姐妹关系, 与 Floden 等^[18]、Zhao 等^[19]及 Xia 等^[20]基于形态和分子证据的研究结论一致。目前, 距药黄精在黄精属内的系统发育位置及其与黄精属其他物种之间的系统发育关系仍存在争议。Meng 等^[16]基于 4 个叶绿体片段 (*rbcL*、*trnK*、*psbA-trnH*、*trnC-petN*) 的系统发育分析结果曾将距药黄精划分至轮叶组, 但其与组内其他种间的系统发育关系并未得到解决, 而 Wang 等^[22]及 Zhang 等^[23]基于叶绿体基因组序列的系统发育分析结果分别将距药黄精置于互叶组或轮叶组。本研究以 100% 的支持率明确澄清了距药黄精在黄精属内的系统发育位置, 即位于轮叶组, 并与狭叶黄精互为姐妹关系。此外, 系统树各分支节点的支持率几乎均为 99% 或 100% (仅 2 个节点分别为 55%、63%), 即黄精属下各物种之间的系统发育关系均以较高的支持率得到解决, 说明叶绿体全基因组序列能够有效解决黄精属植物种间系统发育关系问题。

作者贡献: 第一作者程文萍负责论文实验、数据分析及论文撰写; 通讯作者高静、张明英负责论文设计、数据分析和论文指导; 李依民、赵霞参与实验及数据分析; 张岗参与实验设计及论文指导; 颜永刚参与样品采集。所有作者参与论文修改。

利益冲突: 本文所有的作者之间不存在利益冲突。

References

- [1] Hagemann R. Molecular Biology and Biotechnology of Plant Organelles [M]. Dordrecht: Springer, 2004: 93-113.
- [2] Zhang Q, Sodmergen. Why does biparental plastid inheritance revive in angiosperms? [J]. J Plant Res, 2010, 123: 201-206.
- [3] Jansen RK, Raubeson LA, Boore JL, et al. Methods for obtaining and analyzing whole chloroplast genome sequences [J]. Methods Enzymol, 2015, 395: 348-384.
- [4] Jansen RK, Ruhlman TA. Genomics of Chloroplasts and Mitochondria [M]. Dordrecht: Springer, 2012: 103-126.
- [5] Drouin G, Daoud H, Xia J. Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants [J]. Mol Phylogenet Evol, 2008, 49: 827-831.
- [6] Zhu SY, Niu ZT, Xue QY, et al. Accurate authentication of *Dendrobium officinale* and its closely related species by comparative analysis of complete plastomes [J]. Acta Pharm Sin B, 2018, 8: 969-980.
- [7] Liu CK, Lei JQ, Jiang QP, et al. The complete plastomes of seven *Peucedanum* plants: comparative and phylogenetic analyses for the *Peucedanum* genus [J]. BMC Plant Biol, 2022, 22: 101.
- [8] Huang R, Xie X, Chen A, et al. The chloroplast genomes of four *Bupleurum* (Apiaceae) species endemic to Southwestern China, a diversity center of the genus, as well as their evolutionary implications and phylogenetic inferences [J]. BMC Genomics, 2021, 22: 714.
- [9] Chen XQ, Tamura MN. *Polygonatum* Mill. in Flora of China: Vol 24 [M]. Beijing/Saint Louis: Science Press/Missouri Botanical Garden Press, 2000: 223-232.
- [10] Zhang SR. *Polygonatum* Mill. in Medicinal Flora of China: Vol 11 (中国药用植物志: 第十一卷) [M]. Beijing: Peking University Medical Press, 2014: 288-289.
- [11] Zhang J, Wang YZ, Yang WZ, et al. Research progress in chemical constituents in plants of *Polygonatum* and their pharmacological effects [J]. China J Chin Mater Med (中国中药杂志), 2019, 44: 1989-2008.
- [12] Baker JG. Revision of the genera and species of Asparagaceae [J]. Bot J Linn Soc, 1875, 14: 552-561.
- [13] Tamura MN. Biosystematic studies on the genus *Polygonatum* (Liliaceae) III. Morphology of staminal filaments and karyology of eleven Eurasian species [J]. Bot Jahrb Syst, 1993, 115: 1-26.
- [14] Tamura MN, Schwarzbach AE, Kruse S, et al. Biosystematic studies on the genus *Polygonatum* (Convallariaceae) IV. Molecu-

- lar phylogenetic analysis based on restriction site mapping of the chloroplast gene *trnK* [J]. Feddes Repert, 1997, 108: 159-168.
- [15] Editorial Board of Chinese Flora, Chinese Academy of Sciences. Liliaceae in Flora Reipublaris Sinicae: Vol 15 (中国植物志: 第十五卷) [M]. Beijing: Science Press, 1978: 52-80.
- [16] Meng Y, Nie ZL, Deng T, et al. Phylogenetics and evolution of phyllotaxy in the Solomon's seal genus *Polygonatum* (Asparagaceae: Polygonateae) [J]. Bot J Linn Soc, 2014, 176: 435-451.
- [17] Zhang MY, Li YM, Cheng WP, et al. Molecular authentication of medicinal *Polygonatum* species utilizing the universal DNA barcode sequences [J]. Chin Tradit Herb Drugs (中草药), 2023, 54: 235-244.
- [18] Floden A, Schilling EE. Using phylogenomics to reconstruct phylogenetic relationships within tribe Polygonateae (Asparagaceae), with a special focus on *Polygonatum* [J]. Mol Phylogenet Evol, 2018, 129: 202-213.
- [19] Zhao LH, Zhou SD, He XJ. A phylogenetic study of Chinese *Polygonatum* (Polygonateae, Asparagaceae) [J]. Nord J Bot, 2019, 37: 1-10.
- [20] Xia MQ, Liu Y, Liu JJ, et al. Out of the Himalaya-Hengduan Mountains: phylogenomics, biogeography and diversification of *Polygonatum* Mill. (Asparagaceae) in the Northern Hemisphere [J]. Mol Phylogenet Evol, 2022, 169: 107431.
- [21] Fang ZX, Liao CL. Liliaceae in Flora of Medicinal Plants in Enshi, Hubei (湖北恩施药用植物志) [M]. Wuhan: Hubei Science and Technology Press, 2006: 612-671.
- [22] Wang J, Qian J, Jiang Y, et al. Comparative analysis of chloroplast genome and new insights into phylogenetic relationships of *Polygonatum* and tribe Polygonateae [J]. Front Plant Sci, 2022, 13: 882189.
- [23] Zhang DJ, Ren J, Jiang H, et al. Comparative and phylogenetic analysis of the complete chloroplast genomes of six *Polygonatum* species (Asparagaceae) [J]. Sci Rep, 2023, 13: 7237.
- [24] Patel RK, Jain M. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data [J]. PLoS One, 2012, 7: e30619.
- [25] Jin JJ, Yu WB, Yang JB, et al. GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes [J]. Genome Biol, 2020, 21: 241.
- [26] Wick RR, Schultz MB, Zobel J, et al. Bandage: interactive visualization of *de novo* genome assemblies [J]. Bioinformatics, 2015, 31: 3350-3352.
- [27] Qu XJ, Moore MJ, Li DZ, et al. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes [J]. Plant Methods, 2019, 15: 50.
- [28] Chan PP, Lowe TM. tRNAscan-SE: searching for tRNA genes in genomic sequences [J]. Methods Mol Biol, 2019, 1962: 1-14.
- [29] Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes [J]. Nucleic Acids Res, 2019, 47: W59-W64.
- [30] Kurtz S, Choudhuri JV, Ohlebusch E, et al. REPuter: the manifold applications of repeat analysis on a genomic scale [J]. Nucleic Acids Res, 2001, 29: 4633-4642.
- [31] Benson G. Tandem repeats finder: a program to analyze DNA sequences [J]. Nucleic Acids Res, 1999, 27: 573-580.
- [32] Beier S, Thiel T, Münch T, et al. MISA-web: a web server for microsatellite prediction [J]. Bioinformatics, 2017, 33: 2583-2585.
- [33] Zhang MY, Zhang YQ, Li YM, et al. Complete plastid genomes of *Bupleurum chinense* DC. and *B. boissieuianum* H. Wolff, with comparative and phylogenetic analyses of medicinal *Bupleurum* species [J]. Acta Pharm Sin (药学报), 2021, 56: 618-629.
- [34] Frazer KA, Pachter L, Poliakov A, et al. VISTA: computational tools for comparative genomics [J]. Nucleic Acids Res, 2004, 32: W273-W279.
- [35] Zhang J, Lu JH, Wang QQ, et al. Characteristics of the chloroplast genome of *Glycyrrhiza eurycarpa* P. C. Li from Xinjiang with comparison and phylogenetic analysis of the chloroplast genomes of the medicinal plants of *Glycyrrhiza* [J]. Acta Pharm Sin (药学报), 2022, 57: 1516-1525.
- [36] Zhang MY, Wang XF, Gao J, et al. Complete chloroplast genome of *Paeonia mairei* H. Lévl.: characterization and phylogeny [J]. Acta Pharm Sin (药学报), 2020, 55: 168-176.
- [37] 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 U S A, 1995, 92: 7759-7763.
- [38] Kuang DY, Wu H, Wang YL, et al. Complete chloroplast genome sequence of *Magnolia kwangsiensis* (Magnoliaceae): implication for DNA barcoding and population genetics [J]. Genome, 2011, 54: 663-673.
- [39] Zhang SD, Yan K, Ling LZ. Characterization and phylogenetic analyses of ten complete plastomes of *Spiraea* species [J]. BMC Genomic, 2023, 24: 137.
- [40] Zong D, Qiao ZS, Zhou JT, et al. Chloroplast genome sequence of triploid *Toxicodendron vernicifluum* and comparative analyses with other lacquer chloroplast genomes [J]. BMC Genomic, 2023, 24: 56.
- [41] Zhang Z, Zhang DS, Zou L, et al. Comparison of chloroplast genomes and phylogenomics in the *Ficus sarmentosa* complex (Moraceae) [J]. PLoS One, 2022, 17: e0279849.
- [42] Zhou C, Wang PT, Zeng Q, et al. Comparative chloroplast genome analysis of seven extant *Citrullus* species insight into genetic variation, phylogenetic relationships, and selective pressure [J]. Sci Rep, 2023, 13: 6779.