2023, 36(1): 100-108
http://www.lykxyj.com

DOI:10.12403/j.1001-1498.20220277

九叶青花椒叶绿体基因组结构及系统进化分析

刘　霞[1]，孙　冲[1,2]，黄勤琴[1,3]，谢润泸[1]，刘浩文[1]，陈泽雄[1*]

(1. 重庆文理学院园林与生命科学学院，重庆　402160；2. 长江大学园艺园林学院/香辛作物研究院，湖北 武汉　434025；
3. 重庆三峡学院生物与食品工程学院，重庆　404000)

摘要:［**目的**］揭示九叶青花椒叶绿体基因组的结构特征及其与其他物种的系统进化关系，为花椒属种质资源的鉴定、新品种的培育及品种间的遗传分析提供参考。［**方法**］使用改良 CTAB 法提取九叶青花椒叶片总 DNA。利用 BGISeq-500 平台进行高通量测序，SPAdes v3.13.0 软件组装叶绿体基因组，通过 GeSeq 注释九叶青花椒的全叶绿体基因组信息，对其叶绿体基因组序列的结构特征、重复序列、密码子偏好性及系统进化关系进行分析。［**结果**］九叶青花椒叶绿体基因组为典型的四分体组成结构，全长序列为 158 579 bp，编码 133 个基因；测到 19 个串联重复序列，49 个长片段重复，70 个简单重复序列；九叶青花椒叶绿体基因组中的蛋白编码基因共有 26 398 个密码子（不包括终止密码子），在密码子第三位碱基上有较强的 A/T 碱基偏好性。［**结论**］本研究首次组装了九叶青花椒叶绿体基因组全序列，明确了花椒属为单系类群，九叶青花椒与野花椒（*Z. simulans*）关系密切，这将为花椒的遗传信息、花椒种质资源评价、分子育种、cpSSR 分子标记的开发及遗传多样性等研究提供了重要参考。

关键词: 高通量测序；花椒属；叶绿体基因组；九叶青花椒；系统进化
中图分类号: S718.46　　**文献标志码:** A　　**文章编号:** 1001-1498(2023)01-0100-09

九叶青花椒（*Zanthoxylum armatum* ‘Jiuyeqing’）为芸香科（Rutaceae）花椒属（*Zanthoxylum* L.）竹叶花椒的主要栽培品种，具有矮化、丰产、质优、抗逆等优良特性[1]，其果实清香，麻味纯正，备受消费者喜爱，市场需求旺盛[2]。我国花椒属植物的基础研究十分薄弱，花椒属物种间的系统发育关系尚待进一步解析。

长期以来，国内外有关花椒属植物的研究大都集中在其化学成分的提取和分离[3-4]，以及提取物种所具有的生物活性[5]和相应的药用价值方面[6]，仅有少数学者基于少量的核基因（ETS 和 ITS）[7-8]、叶绿体基因（*rps16* 和 *trnL-trnF*）[8]、分子标记（SSR、RAPD 和 SRAP）[9-10]开展了花椒属内几个物种间的亲缘关系研究。但基于单个片段或分子标记其有效信息位点不足，花椒属内物种间的系统位置支持率不高，花椒属内种间系统进化关系未能很好的解析。因此，解析九叶青花椒的叶绿体基因组结构特征与系统进化关系，为花椒种质资源的分子鉴定、品种选育和农艺经济性状的遗传分析具有重要意义，亦为花椒属系统进化研究提供依据。

叶绿体是植物进行光合作用的重要细胞器，在植物生长发育过程中具有重要作用[11]。叶绿体基因组一般具有保守的四分体结构，即包括大单拷贝区（Large Single Copy, LSC），小单拷贝区（Small

收稿日期：2022-05-30　　修回日期：2022-08-29
基金项目：国家自然科学基金项目（31901324）；重庆市自然科学基金面上项目（cstc2019jcyj-msxmX0693）；重庆市教委青年项目（KJQN202201337；KJQN202101320）
作者简介: 刘霞，博士，副教授。主要研究方向：经济林种质资源开发与利用。电话：19942239287。Email：liuxiavip8@163.com
* 通讯作者: 陈泽雄，博士，教授。主要研究方向：园艺植物创新与利用 电话：15923962393。Email：chenzexiong1979@163.com

Single Copy, SSC），2 个反向重复区（Inverted Repeat Region, IRs）。大多数被子植物叶绿体基因组为母系遗传，且叶绿体基因组的核苷酸替换率适中[12]，因此，叶绿体基因组是进行系统进化和分化时间分析的理想标记[13]。近年来，随着叶绿体基因组测序成本降低，使其在植物系统进化分析中广泛应用[14]。本研究利用高通量测序技术进行主栽品种九叶青花椒叶绿体全基因组测序，通过组装、拼接和注释得到其全长叶绿体基因组序列，并对其叶绿体基因组结构、组成及系统进化关系进行分析，旨在丰富花椒的遗传信息，为未来研究花椒属植物间的系统进化、亲缘关系及品种鉴定提供理论依据。

1　材料与方法

1.1　试验材料

九叶青花椒叶片采于重庆文理学院特色植物研究院花椒示范基地，植物标本（2021023jyq）放置于重庆文理学院园林与生命科学学院标本馆。采用改良 CTAB 法对九叶青花椒提取基因组 DNA，-20 ℃ 冰箱保存，备用。

1.2　试验方法

1.2.1　叶绿体基因组测序和注释　利用提取的 DNA 构建 shotgun 文库，并在 BGISEQ-500 平台 (中国，广州) 进行测序。PE150 双末端测序得到 raw data，过滤掉低质量序列后得到 reads 进行拼接组装。使用 SOAPnuke 软件[15] 进行质量控制后，以花椒（*Zanthouxylum bungeanum* Maxim.）的叶绿体基因组[16] 作为参考，用 SPAdes v3.13.0 软件[17] 组装叶绿体基因组。基因组注释采用 GeSeq（https://chlorobox.mpimp-golm.mpg.de/geseq.html）软件[18]。注释后的叶绿体基因组序列提交至 GenBank 数据库中。

1.2.2　叶绿体基因组结构组成分析　九叶青叶绿体基因组序列的 GC 含量采用 Editseq v7.1.0 软件[19] 计算，OGDRAW 软件[20] 制作九叶青花椒叶绿体基因组图谱。CodonW v.1.4.2 软件[21] 分析相对密码子使用度（Relative Synonymous Codon Usage, RSCU），研究其偏好性。

1.2.3　重复序列预测　通过 Tandem Repeats Finder（https://tandem.bu.edu/trf/trf.html）搜寻基因组序列中的串联重复序列，用 REPuter（https://bibiserv.cebitec.uni-bielefeld.de/reputer/）在线服务分析叶绿体基因组中的重复序列 (正向、反向、回文和互补重复)，采用 MISA (https://webblast.ipk-gatersleben.de/misa/) 在线预测九叶青花椒叶绿体基因组包含的简单重复序列（Simple Sequence Reapeat, SSR），其中，一、二、三、四、五、六碱基重复的检测标准分别设置为 10，6，5，5，5，5。2 个 SSR 之间的最小距离设置为 100 bp，小于 100 bp 时，2 个微卫星组成 1 个复合微卫星。

1.2.4　系统进化分析　从 NCBI 上下载得到 12 个芸香科物种的完整叶绿体基因组序列，将这些叶绿体基因组序列与本研究得到的九叶青花椒的完整叶绿体基因组序列通过 MAFFT 软件[22] 进行比对分析，其中黄檗（*Phellodendron amurense* Rupr.）和川黄檗（*Phellodendron chinense* Schneid.）的叶绿体基因组序列作为外类群。采用 RAxML 软件[23] 构建 ML (Maximum Likelihood）树，建树过程中通过 GTR 模型确定最优树。MrBayes3.1.2 软件[24] 构建 BI（Bayesian）树，采用 GTR 替代模型和伽马分布率，每 1 000 代进行 1 次计算。

2　结果与分析

2.1　九叶青花椒叶绿体全基因组结构及特征

注释组装获得九叶青花椒完整的叶绿体基因组图谱见图 1。该叶绿体基因组序列全长 158 579 bp，包括大单拷贝区域（LSC）长 85 780 bp，小单拷贝区域（SSC）长 17 603 bp，2 个反向重复区（IRa 和 IRb）相等长度为 27 598 bp。共注释了 133 个基因，其中，88 个蛋白编码基因，37 个 tRNA 基因和 8 个 rRNA 基因。叶绿体基因组总的 GC 含量为 38.5%，IR 区域（42.5%）的 GC 含量明显高于 LSC 区域（36.8%）和 SSC 区域（33.6%），这可能主要是 IR 区域具有高 GC 含量的 rRNA 基因造成的。在注释的九叶青花椒叶绿体基因组中 18 个特异基因具有内含子，包括 12 个特异编码基因和 6 个特异 tRNA 基因。除 *pafI* 具有 2 个内含子外，其余均具有 1 个内含子。*rps12* 基因被截断为 2 个片段，其中，1 个外显子位于 LSC 区域，另外 2 个外显子位于 IR 区域。最长内含子位于 *trnK-UUU* 基因，长 2 503 bp，主要是因为内部包含 *matK* 基因；*rps12* 基因内含子最短（538 bp）。

2.2　九叶青花椒叶绿体基因组的重复序列和简单重复序列分析

通过软件测算在九叶青花椒叶绿体基因组中存

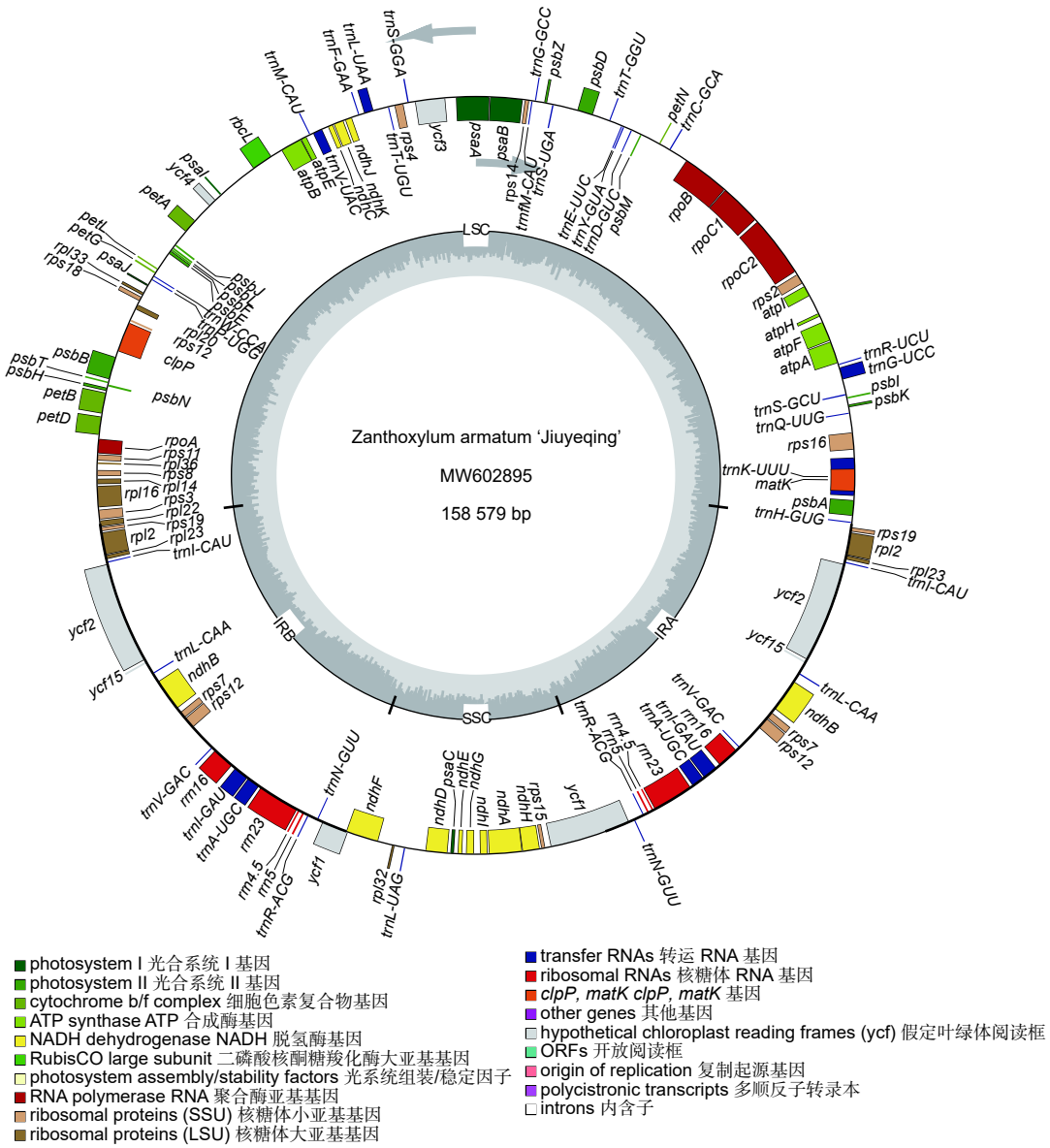


图 1 九叶青花椒叶绿体基因组图谱

Fig. 1 Gene map of the chloroplast genome of *Z. armatum*'Jiuyeqing'

在 19 个串联重复序列，其中，长度最大为 25 bp，最小为 14 bp，长度为 18 bp 的重复序列最多，有 6 个。该叶绿体基因组中共检测到长片段重复序列49 个，其中，21 个正向重复，8 个反向重复，19 个回文重复和 1 个互补重复（图 2）。重复长度范围为 19～50 bp，其中，长度为 20 bp 的重复最多（10 个），其次是 19 bp（9 个），而重复长度为 26、31、48、50 bp 的最少，均 1 个。在九叶青花椒叶绿体基因组中共预测到 70 个 SSR，包括 60 个单核苷酸，1 个二核苷酸，9 个复合型 SSR（表 1）。75.7% 的 SSR 位于 IGS 和内含子

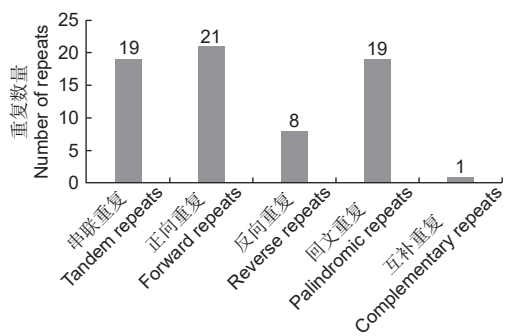


图 2 九叶青花椒叶绿体基因组重复序列类型

Fig. 2 Types of repeat sequences in the chloroplast genome of *Z. armatum*'Jiuyeqing'

等非编码区域且 A/T 碱基在 SSR 中出现频率较高，在 60 个单核苷酸 SSR 中，包含 A 碱基 SSR 有 26 个，包含 T 碱基 SSR 有 32 个。从表 1 可以看出，SSR 在叶绿体基因组中分布随机，表明 SSR 数量丰富，遗传信息含量高，可筛选具有多态性的引物，为未来研究花椒种群遗传多样性和遗传结构提供参考。

表 1　叶绿体基因组 SSR 预测

Table 1　Chloroplast genome SSR prediction

序号 Number	重复类型 Repeat type	简单重复序列 SSR	起始 Start	终止 End	所在位置 Position
1	p1	$(T)_{19}$	1 771	1 789	IGS
2	p1	$(T)_{11}$	2 418	2 428	*matK*
3	p1	$(A)_{10}$	3 085	3 094	*matK*
4	p1	$(A)_{10}$	4 684	4 693	IGS
5	p1	$(A)_{10}$	6 287	6 296	IGS
6	p1	$(T)_{13}$	6 873	6 885	IGS
7	c	$(A)_{12}$gacgtacattttatcaaaaattatacttaatagagttgctcaaa gggggaattc$(T)_{12}$	7 281	7 359	IGS
8	p1	$(A)_{14}$	7 770	7 783	IGS
9	p1	$(T)_{14}$	8 334	8 347	IGS
10	c	$(A)_{11}(AAT)_5$	8 717	8 740	IGS
11	p1	$(T)_{11}$	9 319	9 329	IGS
12	p1	$(A)_{13}$	9 834	9 846	intron
13	p1	$(A)_{15}$	13 273	13 287	intron
14	p1	$(A)_{11}$	14 122	14 132	IGS
15	p1	$(T)_{10}$	16 570	16 579	IGS
16	p1	$(T)_{11}$	19 493	19 503	*rpoC2*
17	p1	$(A)_{12}$	23 671	23 682	intron
18	p1	$(T)_{20}$	24 067	24 086	intron
19	p1	$(T)_{10}$	27 372	27 381	*rpoB*
20	p1	$(A)_{18}$	28 352	28 369	IGS
21	p1	$(C)_{10}$	28 617	28 626	IGS
22	c	$(T)_{12}$caatagaaaaaaagaaagagaaaagagagggattaagatattgtgtgtttt gagatactataaaatcaattgcagaaatggagttgctccaaaaaaag$(A)_{10}$	29 037	29 156	IGS
23	p1	$(A)_{10}$	29 784	29 793	IGS
24	p1	$(T)_{10}$	31 390	31 399	IGS
25	p1	$(T)_{10}$	31 563	31 572	IGS
26	p1	$(T)_{11}$	32 418	32 428	IGS
27	p1	$(T)_{13}$	33 368	33 380	IGS
28	p1	$(G)_{11}$	36 561	36 571	*psbC*
29	p1	$(T)_{10}$	45 777	45 786	intron
30	p1	$(A)_{12}$	47 071	47 082	IGS
31	p1	$(T)_{15}$	51 184	51 198	IGS
32	p1	$(T)_{10}$	52 626	52 635	IGS
33	p1	$(T)_{10}$	53 551	53 560	intron

续表 1

序号 Number	重复类型 Repeat type	简单重复序列 SSR	起始 Start	终止 End	所在位置 Position
34	p1	$(T)_{19}$	56 537	56 555	IGS
35	p1	$(A)_{12}$	57 034	57 045	IGS
36	p1	$(T)_{12}$	61 599	61 610	IGS
37	c	$(T)_{14}$catttttcaaaatattcgagagtttctttttaatag$(A)_{10}$	62 628	62 689	IGS
38	p1	$(A)_{10}$	65 486	65 495	IGS
39	p1	$(A)_{10}$	67 533	67 542	IGS
40	p1	$(A)_{10}$	68 248	68 257	IGS
41	p1	$(A)_{10}$	68 735	68 744	IGS
42	p1	$(A)_{11}$	69 053	69 063	IGS
43	p1	$(T)_{10}$	69 446	69 455	IGS
44	p1	$(A)_{10}$	72 760	72 769	*clpP*
45	p1	$(T)_{10}$	74 477	74 486	IGS
46	p1	$(T)_{10}$	76 170	76 179	IGS
47	p1	$(T)_{12}$	83 898	83 909	*rpl16*
48	p1	$(T)_{16}$	85 546	85 561	IGS
49	p1	$(T)_{11}$	101 363	101 373	IGS
50	p1	$(A)_{16}$	110 363	110 378	IGS
51	p1	$(T)_{10}$	110 534	110 543	IGS
52	c	$(T)_{11}$atttaccattaatggtaaat$(A)_{11}$	110 781	110 822	IGS
53	p1	$(A)_{12}$	112 609	112 620	*ycf1*
54	c	$(A)_{10}$ttcttttgctattcatttttcttttgctattcatagttattttgaatctttcccaac aactt $(A)_{10}$	115 675	115 756	IGS
55	c	$(A)_{17}$cgtac$(T)_{11}$	116 867	116 899	IGS
56	p1	$(A)_{11}$	117 100	117 110	IGS
57	p1	$(T)_{12}$	117 541	117 552	IGS
58	c	$(T)_{10}$atttaactgaaactgaagagaaggaaaaagacttccttgttcattggctaac gaac$(T)_{12}$	118 691	118 768	IGS
59	p1	$(A)_{10}$	120 421	120 430	IGS
60	p1	$(T)_{10}$	122 121	122 130	IGS
61	p2	$(TA)_{6}$	122 278	122 289	IGS
62	p1	$(A)_{13}$	124 680	124 692	intron
63	p1	$(T)_{11}$	126 547	126 557	IGS
64	p1	$(A)_{10}$	127 040	127 049	IGS
65	p1	$(T)_{11}$	129 898	129 908	*ycf1*
66	p1	$(T)_{12}$	131 740	131 751	*ycf1*
67	c	$(T)_{11}$atttaccattaatggtaaat$(A)_{11}$	133 538	133 579	IGS
68	p1	$(A)_{10}$	133 817	133 826	IGS
69	p1	$(T)_{16}$	133 982	133 997	IGS
70	p1	$(A)_{11}$	142 987	142 997	IGS

注：p单个SSR类型；p1/p2/p3中数字分别表示构成基序的碱基个数；c复合SSR类型；IGS代表基因内间隔区

Notes: p is a single SSR type; the numbers in p1/p2/p3 indicate the number of bases constituting the motif, respectively；c is the compound SSR type; IGS represents the intergenic spacer

2.3　密码子偏好性分析

在九叶青花椒叶绿体基因组中的蛋白编码基因共有 26 398 个密码子（不包括终止密码子），其中，亮氨酸 Leu 的密码子最丰富（2 821 个），占总数的 10.69%，编码 Leu 的 6 种同义密码子中 TTA 数量最多；编码半胱氨酸 Cys 的数量最少（310 个），仅占密码子总数的 1.17%。通过对密码子使用频率的计算结果表明：编码异亮氨酸 Ile 的 ATT（1 078）使用次数最多，编码甲硫氨酸 Met 的 ATT（1）和 GTG（2）的使用次数最少。相对同义密码子使用率值（RSCU）受影响密码子使用模式的突变因子的影响，因此，RSCU<1 表示密码子使用率较低，RSCU >1 表示密码子使用率较高，RSCU=1 表示无偏倚[25]。经计算分析，相对同义密码子使用值（RSCU）的范围为 0.005（ATT）到 2.986（ATG），分析还发现 RSCU > 1 的密码子有 31 种，说明这些密码子有一定的偏好性。在这些密码子中，29 个以 A 或 T 结尾，仅 2 个以 G 结尾，这说明九叶青花椒在密码子第三位碱基上有较强的 A/T 碱基偏好性。由图 3 可以看出：编码精氨酸 Arg、亮氨酸 Leu 和丝氨酸 Ser 密码子种类最多，有 6 种密码子编码。色氨酸 Trp 不存在密码子偏好性（RSCU=1），只有 1 个密码子（TGG）编码。

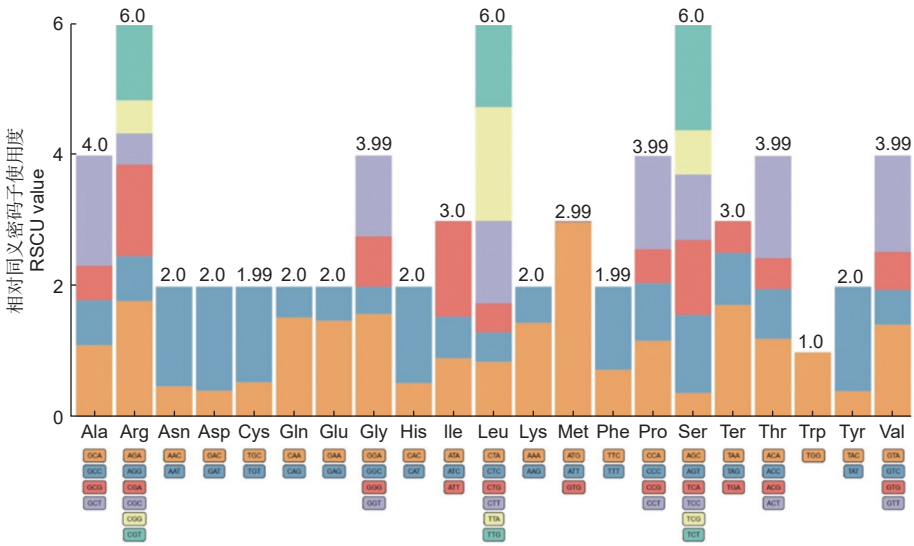


图 3　九叶青花椒叶绿体基因组氨基酸 RSCU 值

Fig. 3　Amino acid RSCU values of the chloroplast genome of *Z. armatum*'Jiuyeqing'

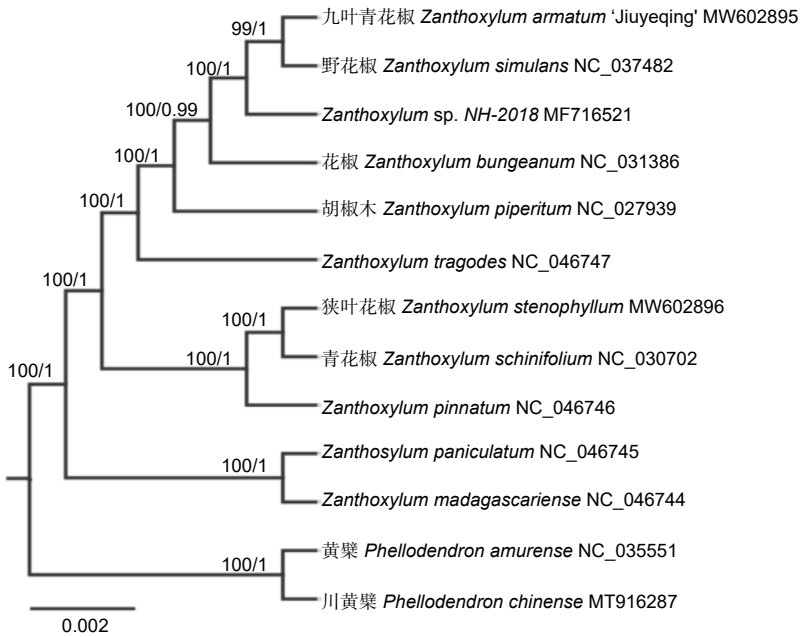
2.4　基于叶绿体基因组序列的系统进化分析

以黄檗和川黄檗 2 个黄檗属植物为外类群，连同花椒属 11 个植物的全叶绿体基因组序列构建 ML 和 BI 系统进化树，2 种方法得到的拓扑结构基本一致，仅在个别分支节点支持率上略有差别。从图 4 可看出：花椒属物种聚为一支，2 种方法在分支上的支持率均为 100%，为单系类群。*Z. paniculatum* 和 *Z. madagascariense* 位于分支基部与其他花椒物种构成姊妹群关系，且分支支持率为 100% 和 1（ML 和 BI）。狭叶花椒、青花椒和 *Z. pinnatum* 聚为一个分支与 *Z. tragodes*、胡椒木、花椒、*Z.* sp. *NH-2018*、野花椒和九叶青花椒形成的分支构成姊妹群关系，分支支持率均为 100%，其中，九叶青花椒与野花椒呈姊妹关系，亲缘关系最近，与花椒（*Z. bungeanum*）亲缘关系相对较远。

3　讨论

在大多数的被子植物中，叶绿体基因组具有典型的四分体结构[26]，序列长度一般为 120～160 kb。九叶青花椒叶绿体基因组同样具有相同的序列结构特征，其叶绿体基因组总长度为 158 579 bp，其中，LSC 长 85 780 bp，SSC 长 17 603 bp，IRs 长度为 27 598 bp，总的 GC 含量为 38.5%，IR 区域（42.5%）的 GC 含量明显高于 LSC 区域（36.8%）和 SSC 区域（33.6%），这主要是因为 IR 区域具有高 GC 含量的 rRNA 基因造成的。

九叶青花椒叶绿体基因组中的蛋白编码基因共有 26 398 个密码子，在 31 种 RSCU > 1 的密码子中，除了 TTG 和 ATG 外，其余密码子均以 A 或 T 结尾，这说明了九叶青花椒在密码子的使用偏向



注：系统发育树分支节点上的数值分别为最大似然法（ML）和贝叶斯推断法（BI）的支持率

Notes：Numbers above branches are bootstrap support values based on maximum likelihood and Bayesian inference, respectively

图 4 基于 13 种植物全叶绿体基因组序列构建的系统发育树

Fig. 4 Phylogenetic tree based on complete chloroplast genomes sequences of 13 species

于第三个密码子位置为 A 和 T，有较强的 A/T 碱基偏好性。研究结果与蒺藜苜蓿（*Medicago truncatula* Gaertn.）[27]、杓兰（*Cypripedium calceolus* L.）[28]、和马尾松（*Pinus massoniana* Lamb.）[29] 等植物叶绿体基因的第 3 位密码子偏好性趋势相吻合，但与酸枣（*Ziziphus jujuba* var. *Spinosa* (Bunge) Hu ex H. F. Chow）[30]、陆地棉（*Gossypium hirsutum* L.）[31] 和樟树（*Cinnamomum camphora* (L.) Presl）[32] 等植物的叶绿体基因分析结果不同。这表明密码子使用偏好性在不同物种中存在一定的差异。

叶绿体 SSR 是研究群体遗传学、生物地理学和系统进化等高效的分子标记[33]，在大量研究中广泛应用。在九叶青花椒叶绿体基因组 SSR 预测中，绝大多数为单碱基重复（60/70），这与肖蒲桃（*Syzygium acuminatissimum* (Blume) Candolle）[34]、胡桃（*Juglans regia* L.）[35] 和金钱槭（*Dipteronia sinensis* Oliv.）[36] 等大多数植物研究结果一致。重复序列的分析为将来青花椒品种及其他植物的物种鉴定和个体水平的遗传差异分析提供了丰富的遗传信息支撑。

我国花椒属物种繁多，品种较多，为该属的物种、品种鉴定带来极大困难。本研究基于叶绿体基因组序列得到的花椒属系统进化结果与 Appelhans 等[8] 利用叶绿体基因片段和核基因片段构建的进化树结果一致，花椒属物种均聚为一个分支，形成单系类群，这与 Feng 等[37] 的研究结果一致。花椒属物种形成 2 大分支，即崖椒亚属和花椒亚属 2 大分支，这与传统的分类学结果一致。本研究表明，九叶青花椒与野花椒亲缘关系较近，与北方主栽花椒亲缘关系相对较远，这与 Feng 等[37] 的研究结果一致。但由于花椒属物种较多，目前提交到数据库中的叶绿体基因组序列有限，基于叶绿体基因组序列构建花椒属的系统进化关系还不够全面，将来需获取更多花椒属物种的数据来全面分析物种间的系统进化关系。

4 结论

本研究利用高通量测序技术，组装和注释了九叶青花椒完整的叶绿体基因组，并解析了该基因组的结构特征和系统进化关系，结果表明，九叶青花椒叶绿体基因组的结构特征与其他花椒属植物相似，为典型的四分体组成结构，全长序列为 158 579 bp，编码 133 个基因，26 398 个密码子，检测到 19 个串联重复序列，49 个长片段重复、70 个简单重复序列。系统进化关系分析表明，花椒属为单系

类群，九叶青花椒与野花椒关系密切。研究结果可为未来研究花椒种群遗传多样性和遗传结构提供参考。

参考文献：

[1] 杨文英, 王　玲, 吕玉奎, 等. 荣昌无刺花椒与九叶青花椒主要品质特征比较分析[J]. 四川林业科技, 2019, 40（6）：60-64.

[2] 杨林生, 杨　敏, 彭　清, 等. 重庆市九叶青花椒施肥现状评价[J]. 西南大学学报:自然科学版, 2020, 42（3）：61-68.

[3] BARKATULLAH, IBRAR M, MUHAMMAD N, *et al*. Pharmacognostic and phytochemical studies of *Zanthoxylum armatum* DC[J]. Pakistan Journal of Pharmaceutical Sciences, 2017, 30(2): 429-438.

[4] JULISTIONO H, LESTARI F G, IRYANTO R, *et al*. Antimycobacterial activity of fruit of *Zanthoxylum acanthopodium* DC against *Mycobacterium smegmatis*[J]. Avicenna Journal of Phytomedicine, 2018, 8(5):432-438.

[5] ZHANG W J, WANG Y, GENG Z F, *et al*. Antifeedant activities of lignans from stem bark of *Zanthoxylum armatum* DC. against *Tribolium castaneum*[J]. Molecules, 2018, 23(3): E617.

[6] NANTONGO J S, ODOI J B, ABIGABA G, *et al*. Variability of phenolic and alkaloid content in different plant parts of *Carissa edulis* Vahl and *Zanthoxylum chalybeum* Engl[J]. BMC Research Notes, 2018, 11(1): 125.

[7] SUN Y L, PARK W G, KWON O W, *et al*. Molecular identification of *Zanthoxylum piperitum* by the internal transcribed spacers as targets using newly designed species-specific primers[J]. Journal of Biotechnology, 2010, 150: 505.

[8] APPELHANS M S, REICHELT N, GROPPO M, *et al*. Phylogeny and biogeography of the pantropical genus *Zanthoxylum* and its closest relatives in the proto-Rutaceae group (Rutaceae)[J]. Molecular Phylogenetics and Evolution, 2018, 126: 31-44.

[9] 郑海星, 李周岐, 薛惠丹, 等. 花椒种质资源的RAPD分析[J]. 西北林学院学报, 2011, 26（2）：96-100.

[10] FENG S J, YANG T X, LIU Z S, *et al*. Genetic diversity and relationships of wild and cultivated *Zanthoxylum germplasms* based on sequence-related amplified polymorphism (SRAP) markers[J]. Genetic Resources and Crop Evolution, 2015, 62(8): 1193-1204.

[11] CHO K S, YUN B K, YOON Y H, *et al*. Complete chloroplast genome sequence of tartary buckwheat (*Fagopyrum tataricum*) and comparative analysis with common buckwheat (*F. esculentum*)[J]. PloS ONE, 2015, 10(5): e0125332.

[12] KARLO K, PETR V, ALEXANDER B, *et al*. Allopolyploid origin of *Chenopodium album* s. str. (Chenopodiaceae): a molecular and cytogenetic insight[J]. PloS ONE, 2016, 11(8): e0161063.

[13] 张韵洁, 李德铢. 叶绿体系统发育基因组学的研究进展[J]. 植物分类与资源学报, 201, 33(4): 365-375.

[14] JOSE C C, ROBERTO A, VICTORIA I, *et al*. A phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus *Citrus*[J]. Molecular Biology and Evolution, 2015, 32(8): 2015-2035.

[15] CHEN Y X, CHEN Y S, SHI C M, *et al*. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data[J]. Giga Science, 2018, 7(1): 1-6.

[16] LI Y L, WEI A Z. The complete chloroplast genome sequence of an economically important plant, *Zanthoxylum bungeanum* (Rutaceae)[J]. Conservation Genetics Resources, 2017, 9: 25-27.

[17] BANKEVICH A, NURS S, ANTIPOV D, *et al*. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing[J]. Journal of Computational Biology, 2012, 19(5): 455-477.

[18] TILLICH M, LEHWARK P, PELLIZZER T, *et al*. GeSeq-versatile and accurate annotation of organelle genomes[J]. Nucleic Acids Research, 2017, 45(W1): W6-W11.

[19] BURLAND T G. DNASTAR's Lasergene sequence analysis software[J]. Methods in Molecular Biology (Clifton, N. J.), 2000, 132: 71-91.

[20] LOHSE M, DRECHSEL O, BOCK R. Organellar Genome DRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes[J]. Current Genetics, 2007, 52: 267-274.

[21] SHARP P M, LI W H. Codon usage in regulatory genes in *Escherichia coli* does not reflect selection for rare codons[J]. Nucleic Acids Research, 1986, 14(19): 7737-7749.

[22] KATOH K, TOH H. Parallelization of the MAFFT multiple sequence alignment program[J]. Bioinformatics, 2010, 26(15): 1899-1900.

[23] STAMATAKIS A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models[J]. Bioinformatics, 2006, 22(21): 2688-2690.

[24] RONQUIST F, HUELSENBECK J P. MrBayes 3: bayesian phylogenetic inference under mixed models[J]. Bioinformatics, 2003, 19(12): 1572-1574.

[25] BARBHUIYA P A, UDDIN A, CHAKRABORTY S. Analysis of compositional properties and codon usage bias of mitochondrial CYB gene in anura, urodela and gymnophiona[J]. Gene, 2020, 751: 144762.

[26] WICKE S, SCHNEEWEISS G M, DEPAMPHILIS C W, *et al*. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function[J]. Plant Molecular Biology, 2011, 76(3): 273-297.

[27] 杨国锋, 苏昆龙, 赵怡然, 等. 蒺藜苜蓿叶绿体密码子偏好性分

析[J]. 草业学报, 2015, 24（12）：171-179.
[28] 丁 锐, 胡 兵, 宗小雁, 等. 杓兰叶绿体基因组密码子偏好性分析[J]. 林业科学研究, 2021, 34（5）：177-185.
[29] 叶友菊, 倪州献, 白天道, 等. 马尾松叶绿体基因组密码子偏好性分析[J]. 基因组学与应用生物学, 2018, 37（10）：4464-4471.
[30] 胡晓艳, 许艳秋, 韩有志, 等. 酸枣叶绿体基因组密码子使用偏性分析[J]. 森林与环境学报, 2019, 39（6）：621-628.
[31] 尚明照, 刘 方, 华金平, 等. 陆地棉叶绿体基因组密码子使用偏性的分析[J]. 中国农业科学, 2011, 44（2）：245-253.
[32] 秦 政, 郑永杰, 桂丽静, 等. 樟树叶绿体基因组密码子偏好性分析[J]. 广西植物, 2018, 38（10）：1346-1355.
[33] ZENG J M, CHEN X J, WU X F, *et al*. Genetic diversity analysis of genus *Nicotiana* based on SSR markers in chloroplast genome and mitochondria genome[J]. Acta Tabacaria Sinica, 2016, 22(4): 89-97.
[34] 邓演文, 刘晓洲, 卓定龙, 等. 肖蒲桃叶绿体基因组结构特征及系统发育关系[J]. 广东农业科学, 2021, 48（6）：38-46.
[35] DONG W P, XU C, LI W Q, *et al*. Phylogenetic resolution in *Juglans* based on complete chloroplast genomes and nuclear DNA sequences[J]. Frontiers in Plant Science, 2017, 8: 1148.
[36] ZHOU T, CHEN C, WEI Y, *et al*. Comparative transcriptome and chloroplast genome analyses of two related *Dipteronia* species[J]. Frontiers in Plant Science, 2016, 7: 1512.
[37] FENG S J, LIU Z S, CHEN L, *et al*. Phylogenetic relationships among cultivated *Zanthoxylum* species in China based on cp-DNA markers[J]. Tree Genet Genomes, 2016, 12(1): 45-53.

Analysis of Complete Chloroplast Genome Sequences and Phylogenetic Evolution of *Zanthoxylum armatum* ‘Jiuyeqing’

LIU Xia[1], *SUN Chong*[1,2], *HUANG Qin-qin*[1,3], *XIE Run-lu*[1], *LIU Hao-wen*[1], *CHEN Ze-xiong*[1]

(1. College of Landscape Architecture and Life Science, Chongqing University of Arts and Sciences, Chongqing 402160, China;
2. Spice Crops Research Institute, College of Horticulture and Gardening, Yangtze University, Wuhan, 434025, Hubei, China;
3. College of Biology and Food Engineering, Chongqing Three Gorges University, Chongqing, 404000, China)

Abstract: [**Objective**] To reveal the structural characteristics and phylogenetic relationships of the chloroplast genomes of *Zanthoxylum armatum* ‘Jiuyeqing’, and provide reference for identifying germplasm resources, breeding new varieties and the genetic analysis among varieties of the genus *Zanthoxylum*. [**Method**] The total DNA of *Z. armatum* ‘Jiuyeqing’ was extracted using a modified CTAB method. High-throughput sequencing was performed using the BGISeq-500 platform, the chloroplast genome was assembled by SPAdes v3.13.0 software, the whole chloroplast genome information of *Z. armatum* ‘Jiuyeqing’ was annotated by GeSeq software, and the structural characteristics, repetitive sequences, codon preference and phylogenetic relationship were analyzed. [**Result**] *Z. armatum* ‘Jiuyeqing’ was a typical tetrad structure with a full-length sequence of 158 579 bp, encoding 133 genes. 19 tandem repeats were detected, 49 long repeat sequences, 70 simple sequence repeat (SSR) were detected. A total of 26 398 codons (excluding stop codons) were detected in the protein-coding gene of the chloroplast genome of *Z. armatum* ‘Jiuyeqing’, and there was a strong A/T base preference at the third base of the codon. [**Conclusion**] The complete chloroplast genome sequences of *Z. armatum* 'Jiuyeqing' was firstly assembled for the study. The results of phylogenetic analysis show that the genus *Zanthoxylum* is a monophyletic group, and the *Z. armatum* 'Jiuyeqing' is closely related to *Z. simulans*. The results of the study enrich the genetic information of *Zanthoxylum* and improve important references for *Zanthoxylum* germplasm resource evaluation, molecular breeding, development of cpSSR molecular markers, and genetic diversity research.

Keywords: High-throughput sequencing; *Zanthoxylum*; Chloroplast genome; *Zanthoxylum armatum* ‘Jiuyeqing’; Phylogeny

（责任编辑：张 研）