

• 药材与资源 •

黄连转录组 SSR 位点信息分析及分子标记开发

崔亚杰¹, 史荷凤¹, 聂广楼², 廖海迪³, 李学刚^{1*}, 庞 蕾^{1*}

1. 西南大学药学院, 重庆 400715

2. 石柱县农业委员会, 重庆 409100

3. 重庆旺隆黄连科技有限公司, 重庆 409100

摘要: 目的 探究黄连 *Coptis chinensis* 转录组中简单重复序列 (simple sequence repeat, SSR) 位点信息, 进而开发可用于黄连资源遗传多样性评价的 SSR 分子标记。方法 采用 MISA 软件对黄连叶片转录组获得的 93 629 条 Unigene 进行 SSR 位点数据挖掘, 利用 Primer 3.0 软件进行引物设计, 以 64 株不同形态的黄连作为分析样本, 通过聚丙烯酰胺凝胶电泳筛选多态性引物。结果 黄连转录组中的 12 848 个 SSR 位点分布于 10 016 条 Unigene 序列中, SSR 位点的出现频率为 13.72%。单核苷酸和三核苷酸重复是 SSR 序列中最主要的重复类型, 分别占 SSR 总数的 53.81% 和 25.01%。SSR 序列中共包含 115 种重复基元, 优势重复基元为 A/T、AG/CT、AAG/CTT。随机筛选的 180 对引物中, 94 对可扩增出清晰条带, 其中的 7 对引物具有明显多态性, 能够将 64 份黄连样本分成 4 个类群, 其生物性状与遗传多样性聚类结果基本吻合。结论 黄连转录组中 SSR 位点出现频率较高, 重复类型丰富, 多态性较高, 可为黄连资源研究提供候选分子标记。开发的 7 个 SSR 标记可用于黄连资源遗传多样性评价、种质鉴定、分子辅助育种等研究。

关键词: 黄连; 转录组; 简单重复序列; 分子标记; 多态性

中图分类号: R286.12 文献标志码: A 文章编号: 0253-2670(2024)19-6713-08

DOI: 10.7501/j.issn.0253-2670.2024.19.023

Site information analysis and molecular marker based development of SSR in *Coptis chinensis* transcriptomeCUI Yajie¹, SHI Hefeng¹, NIE Guanglou², LIAO Haidi³, LI Xuegang¹, PANG Lei¹

1. College of Pharmaceutical Sciences, Southwest University, Chongqing 400715, China

2. Shizhu County Agricultural Commission, Chongqing 409100, China

3. Chongqing Wang Long Hang Lian Technology Co., Ltd., Chongqing 409100, China

Abstract: Objective To explore the simple sequence repeat (SSR) site information in the transcriptome of *Coptis chinensis* and develop the SSR molecular markers which can be used for the genetic diversity evaluation of *C. chinensis* resource. **Methods** MISA software was used to mine the SSR site data of 93 629 Unigenes obtained from the transcriptome of *C. chinensis* leaves. SSR primers were designed through Primer 3.0 software. A total of 64 samples of *C. chinensis* with different morphological features were used to select the polymorphic primers through polyacrylamide gel electrophoresis (PAGE). **Results** A total of 12 848 SSR loci in the transcriptome of *C. chinensis* were distributed in 10 016 Unigenes, with a SSR loci frequency of 13.72%. Mono-nucleotide and tri-nucleotide repeats were the main repeat types in SSR sequences, accounting for 53.81% and 25.01% of the total SSRs, respectively. The SSR sequence contained 115 repeat motifs, with dominant motifs being A/T, AG/CT, and AAG/CTT. Among the 180 pairs of primers randomly selected, 94 pairs could be used to amplify expected bands and seven pairs exhibited significant polymorphism. A total of 64 samples of *C. chinensis* could be divided into four groups, and their biological characteristics were basically consistent with the clustering results of genetic diversity. **Conclusion** The SSR loci in the transcriptome of *C. chinensis* are high in frequency, rich in repetition type, high in polymorphism, which can provide candidate molecular markers for the research on *C. chinensis* resources. The seven SSR molecular markers screened in this study can be used for genetic diversity evaluation, variety identification, molecular assisted breeding of *C. chinensis*.

Key words: *Coptis chinensis* Franch.; transcriptome; simple sequence repeat (SSR); molecular marker; polymorphism

收稿日期: 2024-02-02

基金项目: 国家重点研发计划项目 (2022YFD1201600)

作者简介: 崔亚杰 (1998—), 女, 硕士研究生, 研究方向为中药资源可持续利用与保护。Tel: (023)68251225 E-mail: 1242119657@qq.com

*通信作者: 李学刚 (1964—), 男, 教授, 研究方向为药用资源综合开发利用。Tel: (023)68251225 E-mail: xuegangli@swu.edu.cn

庞 蕾 (1979—), 女, 讲师, 研究方向为中药资源可持续利用与保护。Tel: (023)68251225 E-mail: pangl@swu.edu.cn

黄连是我国最常用的中药材之一，具有清热燥湿、泻火解毒的功效，在临床上已有 2 000 多年的使用历史。《中国药典》2020 年版记载中药黄连来源于毛茛科植物黄连 *Coptis chinensis* Franch.、三角叶黄连 *C. deltoidea* C. Y. Cheng et Hsiao 或云南黄连 *C. teeta* Wall. 的干燥根茎，分别习称为“味连”“雅连”“云连”^[1]。目前全国市场上的主流黄连商品为味连^[2]，即黄连 *C. chinensis* Franch. 的干燥根茎，约占整个黄连销量的 90% 左右^[3]。黄连主要分布于重庆、湖北、四川、陕西等地，重庆石柱和湖北利川是味连的道地产区。重庆市石柱县人工栽培的味连产量约占全国总产量的 60% 左右，通过在石柱调研发现，不同类群黄连的亩产量和活性成分含量具有显著差异，这将严重影响中药黄连的质量。要使中药黄连质量稳定可控，使黄连产业可持续发展，必须加强黄连种质资源研究和保护，选育高产优质黄连品种。目前关于黄连的研究，主要集中在化学成分、药理活性、组方配伍方面，有学者就黄连属不同物种的遗传亲缘关系进行了比较研究^[4]，而黄连种内分类群的遗传多样性研究较为匮乏^[5-7]，亟需黄连种内不同变异类型的系统研究。

分子标记技术作为揭示物种遗传多样性最重要的工具之一，广泛应用于药用植物种质资源与道地性评价、亲缘关系研究以及品种鉴定等领域^[8]。其中的简单重复序列（simple sequence repeat, SSR）标记利用串联重复序列 2 侧的保守序列设计引物，获得扩增产物长度多态性分析。SSR 标记具有分布广泛、共显性遗传、信息含量丰富、染色体定位清楚、易于检测和重复性好等特点，已被广泛应用于药用植物的品种鉴定，种质资源保存与利用、遗传多样性分析、分子标记辅助育种等研究领域^[9-11]。本研究采用 Illumina 高通量测序技术获得了黄连叶片转录组数据，基于该转录组数据挖掘黄连 SSR 位点信息，分析其组成、分布等特征，并开发出 7 对多态性 SSR 引物，为鉴定黄连种质、评价黄连资源遗传多样性以及品种选育提供研究基础。

1 材料与仪器

材料采集自重庆市石柱土家族自治县黄水镇二架梁黄连种质资源圃，随机选取健康黄连植株共 64 份（编号 1~64），经西南大学药学院张迎春副教授鉴定为毛茛科植物黄连 *C. chinensis* Franch.。取较幼嫩叶片用清水洗净、晾干，液氮速冻后置 -80 °C 超低温冰箱保存。

2 方法

2.1 形态学指标检测

对 64 份黄连样本的花及叶的形态特征进行观察，测量和记录样本的叶片大小、表面光亮度、叶片分裂程度、叶片上花斑情况、花萼的颜色等特征。

2.2 转录组数据来源

从所采集的黄连植株中随机选择 2 株，其幼叶作为转录组测序样本，由上海美吉生物医药科技有限公司采用 Illumina NovaSeq6000 测序平台进行转录组测序、拼接和组装。

2.3 SSR 位点筛选

采用 MISA (<http://pgsc.ipk-gatersleben.de/misa/misa.html>) 软件对 Unigene 进行 SSR 检测。SSR 位点的筛选标准为一至六核苷酸的重复次数最少为 10、6、5、5、5 和 5 次。统计 SSR 出现频率、分布密度、重复类型等信息。

2.4 SSR 引物设计

根据 SSR 位点侧翼序列，用 Primer 3.0 软件设计引物，每个 SSR 位点分别设计 3 对引物。引物设置参数为：引物长度 18~27 bp，产物大小 100~280 bp，退火温度 57~63 °C。随机筛选出 180 对引物，由生工生物工程（上海）股份有限公司合成。

2.5 多态性引物的筛选与验证

黄连基因组 DNA 用 TIANGEN 植物基因组 DNA 提取试剂盒（DP305）进行提取，1% 琼脂糖凝胶电泳和核酸定量分析仪检测 DNA 的完整性及浓度，-20 °C 保存备用。

多态性引物的筛选分为 2 步。首先，任选 1 株黄连的 DNA 作为模板，以 1.5% 的琼脂糖凝胶电泳进行引物有效性初筛。扩增体系如下：Premix Taq（TaKaRa）10 μL，DNA 模板 40 ng，正、反向引物 0.4 μL（20 μmol/L），ddH₂O 补足 20 μL。反应程序为 94 °C 预变性 5 min；94 °C 变性 30 s，55 °C 退火 30 s，72 °C 延伸 30 s，32 个循环；72 °C 延伸 5 min。扩增产物进行琼脂糖凝胶电泳，选择可扩增出清晰目的条带的引物进行其多态性筛选。其次，随机选择 12 株黄连的 DNA 作为模板，以 8% 非变性聚丙烯酰胺凝胶电泳进行引物多态性复筛。PCR 反应程序同上，反应体系除 DNA 模板为 20 ng 外，其余同上。

复筛获得的具有多态性的引物用于 64 株黄连样本的遗传多样性分析，以验证多态性引物的可用性。PCR 反应体系及程序与复筛过程保持一致。

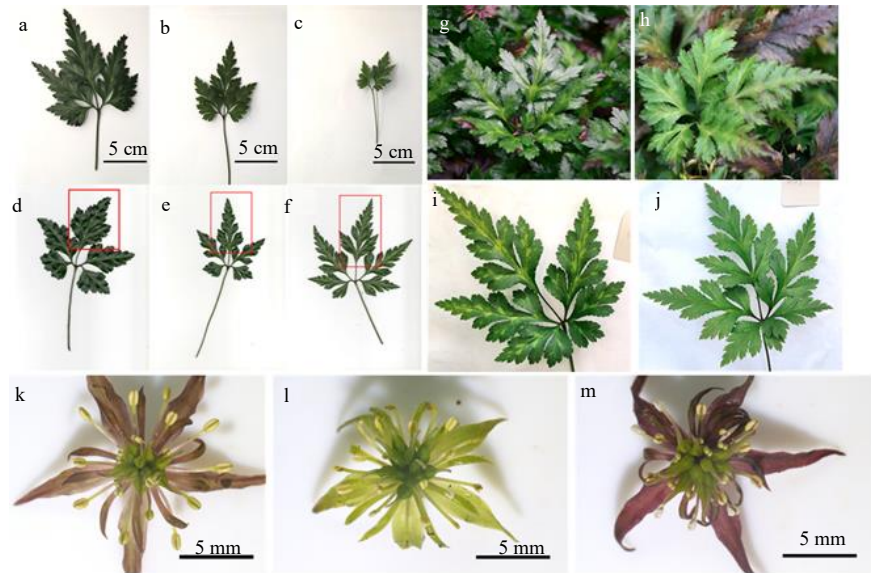
2.6 数据统计与分析

根据 DNA Marker 和引物信息中目标片段大小，人工判读 64 份黄连样本的聚丙烯酰胺凝胶电泳结果。不同等位基因，按照片段的大小依次标记为 A、B、C、D 等。采用 POPGENE 32 软件对扩增位点进行遗传多样性参数分析。采用 PowerMarker V3.25 软件计算多态信息量（polymorphism information content，PIC），用非加权组平均法（UPGMA）构建基于 Nei's 距离的系统发育树，对 64 份黄连样本进行聚类分析。

3 结果与分析

3.1 黄连形态特征

根据供试样本的形态特征，将叶片大小分为较大、中等和较小 3 个类型；将叶片光亮度分为光亮和不亮 2 类；将叶片分裂度分为羽片较宽、羽片中等宽度和羽片较窄 3 个类别；将叶片具花斑的情况分为有花斑和无花斑 2 类；将花萼颜色分为紫红、深紫红和黄绿 3 类。黄连叶和花的典型形态特征见图 1。



a~c-分别为较大叶片、中等叶片、较小叶片；d~f-分别为羽片较宽、羽片中等宽度、羽片较窄；g, h-分别为光亮叶片、不亮叶片；i, j-分别为具花斑叶、无花斑叶；k~m-分别为紫红花萼、黄绿花萼、深紫红花萼。
a, b, c-large leaves, medium leaves, small leaves; d, e, f-broad lobes, medium lobes, narrow lobes; g, h-lustrous leaf, unlustrous leaf; i, j-mottle leaf, mottle-free leaf; k, l, m-purplish red calyx, yellowish green calyx, dark purplish red calyx.

图 1 黄连叶和花的形态特征

Fig. 1 Morphological characteristics of leaves and flowers from *C. chinensis*

3.2 黄连转录组中 SSR 位点分布及特征

通过 Illumina NovaSeq6000 平台进行高通量测序，完成黄连叶片的转录组分析，共获得 93 629 个 Unigene 序列，从中检测到的 12 848 个 SSR 位点分布于 10 016 条序列中，SSR 位点出现频率为 13.72%，平均每 5.34 kb 出现 1 个 SSR 位点（表 1）。从黄连转录组中共检测到 6 种核苷酸重复类型，其中单核苷酸重复位点最多，为 6 914 个，占总 SSR 位点的 53.81%；其次是三核苷酸重复序列和二核苷酸重复序列，分别占总 SSR 位点的 25.01%和 19.32%；四核苷酸、五核苷酸、六核苷酸重复很少，总计不足 SSR 位点总数的 2%（表 2）。

3.3 黄连转录组 SSR 位点重复基元类型

对黄连转录组中 SSR 位点的不同重复基元进

表 1 黄连转录组中 SSR 位点信息

Table 1 SSR site information in transcriptome of *C. Chinensis*

项目	数值
Unigene 序列数目	93 629
检测序列总大小/bp	68 633 683
检测到的 SSR 数目	12 848
含 SSR 位点的 Unigene 序列数目	10 016
含多个 SSR 位点的 Unigene 序列数目	1 457
含复合型 SSR 的 Unigene 序列数目	895

行分析，共有 115 种重复基元类型。其中单核苷酸重复基元类型 2 种，占比最大的为 (A/T)_n 基元，形成的 SSR 位点共计 6 716 个，占总 SSR 位点的 52.27%，占单核苷酸 SSR 位点的 97.14%。二核苷

表 2 黄连转录组中 SSR 重复类型、数量及分布
Table 2 Repeat types, number and distribution of SSR in transcriptome of *C. Chinensis*

重复类型	SSR 数目/个	占比/%	频率/%	平均分布距离/kb
单核苷酸	6 914	53.81	7.38	9.93
二核苷酸	2 482	19.32	2.65	27.65
三核苷酸	3 213	25.01	3.43	21.36
四核苷酸	122	0.95	0.13	562.57
五核苷酸	31	0.24	0.03	2 213.99
六核苷酸	86	0.67	0.09	798.07

酸 4 种重复基元类型, (AG/CT)_n 基元类型占优势, 共形成 1 366 个 SSR 位点 (10.63%, 55.04%)。三核苷酸重复 10 种基元类型中, (AAG/CTT)_n 基元优势明显, 共形成 967 个 SSR 位点 (7.53%, 30.10%); 其次为(ATT/ATT)_n, 共形成 526 个 SSR 位点(4.09%, 16.37%)。四核苷酸、五核苷酸、六核苷酸重复基元类型较多, 分别有 20、21、58 种, 优势重复基元类型依次为 (AAAT/ATTT)_n、(AAAAC/GTTTT)_n、(AACAGC/CTGTTG)_n, 其重复基元数量在总 SSR 位点中占比很小, 分别占 SSR 总数的 0.23%、0.03%、0.04%。

3.4 黄连转录组 SSR 重复次数和基序长度

黄连转录组中 SSR 重复次数分布于 5~58 次, 5~15 次重复的 SSR 占总数的 90.87%。10 次重复最多, 有 3 228 个, 占比 25.12% (表 3)。单核苷酸重复次数集中于 10~11 次, 二核苷酸重复次数以 6~8 次为主, 三核苷酸重复次数多为 5~6 次, 高级重复类型 (四核苷酸至六核苷酸) 均以 5 次重复为主。SSR 长度是影响多态性高低的重要因素^[12], 本研究中 SSR 长度范围在 10~20 bp 的 SSR 有 9 539 个, 占 SSR 总数的 80.94%, 其次为 21~30 bp 的 SSR, 共计 1 249 个, 占 SSR 总数的 10.60%, 31 bp 及以上的 SSR 所占比例较小, 不到 10% (图 2)。

3.5 多态性 SSR 引物的筛选

根据转录组数据, 随机选取 180 对不同重复类型的引物, 通过琼脂糖凝胶电泳, 初步筛选出 94 对引物可扩增出明亮清晰的目的条带。以 12 份黄连植株的 DNA 为模板, 对 94 对引物进行二次筛选, 获得 21 对具有多态性的引物。以 64 份黄连样本的 DNA 为模板, 最终筛选出主带明显、多态性较好、可准确判读的 7 对引物, 多态性较好的引物 P4-11 的扩增结果见图 3, 7 对多态性引物信息见表 4。

表 3 黄连转录组中不同类型 SSR 的重复次数统计
Table 3 Repetition count of SSR with different type in transcriptome of *C. Chinensis*

重复次数	重复类型						频率/%
	单核苷酸	二核苷酸	三核苷酸	四核苷酸	五核苷酸	六核苷酸	
5	0	0	1 695	85	22	69	14.56
6	0	999	772	28	7	3	14.08
7	0	550	379	5	0	5	7.31
8	0	359	181	2	1	6	4.27
9	0	217	83	0	0	2	2.35
10	3 056	109	63	0	0	0	25.12
11	1 232	73	7	1	0	0	10.22
12	624	42	9	0	0	1	5.26
13	361	28	4	0	1	0	3.07
14	283	36	5	1	0	0	2.53
15	242	26	2	0	0	0	2.10
16	197	2	1	0	0	0	1.56
17	128	7	7	0	0	0	1.11
18	110	5	3	0	0	0	0.92
19	99	4	2	0	0	0	0.82
20	69	6	0	0	0	0	0.58
≥21	513	19	0	0	0	0	4.14

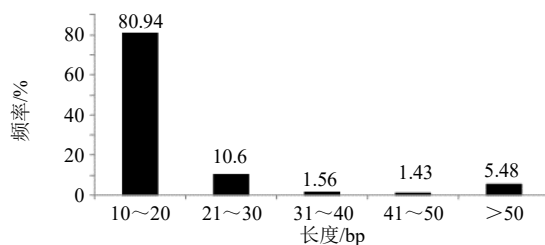


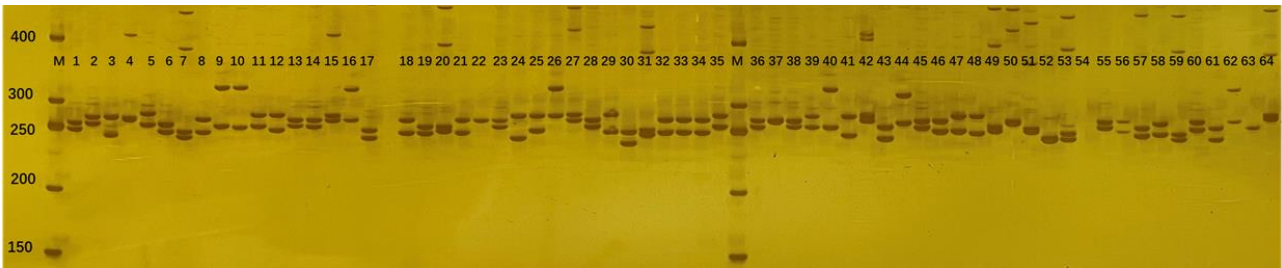
图 2 黄连转录组中 SSR 长度的频率
Fig. 2 Frequency of SSR with different length in transcriptome of *C. Chinensis*

3.6 SSR 标记的多态性分析

7 对多态性引物在 64 个黄连样本中共扩增得到 29 个多态性片段 (表 5), 平均每对引物扩增得到 4 个。根据 Nei's 算法获得的期望杂合度 (H_e) 为 0.334 1~0.774 9, 平均值为 0.561 1, 观测杂合度 (H_o) 为 0.349 2~0.761 9, 平均值为 0.567 1。PIC 从 0.300 8 到 0.726 1, 平均值为 0.493 8。引物 P6-5、P3-25 和 P4-11 为高度多态性位点 ($PIC > 0.5$), 其余 4 个均为中度多态性位点 ($0.25 < PIC < 0.5$)。Shannon 信息指数 (I) 范围为 0.681 5~1.525 6, 平均 1.006 6。以上结果表明基于黄连转录组开发的 7 对引物, 具有较高的多态性, 可用于黄连种质遗传多样性分析。

3.7 黄连遗传多样性分析

根据 64 个黄连样本间相应的 Nei's 遗传距离,



M-Marker; 1~64-供检测的黄连样本。
M-Marker; 1—64-samples of *C. chinensis*.

图 3 引物 P4~11 聚丙烯酰胺凝胶电泳结果
Fig. 3 Polyacrylamide gel electrophoresis (PAGE) result of primer P4—11

表 4 7 对多态性 SSR 引物信息

Table 4 Information of seven pairs of polymorphic SSR primers

引物编号	引物 (5'-3')	重复单元	退火温度/℃	产物大小/bp
P6-3	F: CAAAGAAGAGGAAGATGCGG R: ATCTTTTGCCTGCTGCTGTT	(CAACAG) ₆	53.5	221
P6-5	F: AACTTCTCCGCGTCAACATC R: GAAAGTGCCGCAGATTAGGA	(ACTATT) ₇	55.4	231
P5-10	F: AGGAGGAAAGCCACTCACAG R: AGCAAATCCACAGTCAACCC	(GAGAG) ₆	57.2	238
P4-11	F: AGTTCAGCACCATCCTCGTT R: CGTCCCTCATGCTCATGTAA	(TTCT) ₁₄	56.4	274
P4-12	F: TCCTCATCCTGATCCCTTTG R: CCTGCCGTCCTATTATCCAA	(TTTA) ₆	53.8	176
P3-25	F: AACCGGACAACCTCCTATC R: ACCGTAGCCCTCATAACACG	(CCA) ₆	56.7	254
P2-9	F: TACCAACTTATCCGCCTCCA R: TGTCTGTTGCAGTCTCGTCC	(AG) ₈	55.5	216

表 5 SSR 引物的多态性分析
Table 5 Polymorphism analysis of SSR primers

位点	等位基因数 (N_a)	有效等位基数 (N_e)	H_o	H_e	PIC	I
P6-3	4.000 0	1.495 8	0.349 2	0.334 1	0.300 8	0.687 0
P6-5	5.000 0	2.945 5	0.873 0	0.665 8	0.519 0	1.226 0
P5-10	4.000 0	2.062 4	0.460 3	0.519 2	0.421 1	0.835 2
P4-12	4.000 0	1.692 9	0.390 6	0.412 5	0.351 1	0.681 5
P4-11	5.000 0	4.323 5	0.761 9	0.774 9	0.726 1	1.525 6
P3-25	4.000 0	3.336 7	0.634 9	0.705 9	0.642 8	1.286 6
P2-9	3.000 0	2.045 4	0.500 0	0.515 1	0.424 0	0.804 1
平均值	4.142 9	2.557 4	0.567 1	0.561 1	0.493 8	1.006 6

在 PowerMarker V3.25 软件上进行 UPGMA 聚类(图 4)。结果显示 64 个黄连样本聚为 4 个大类,其中第 I 类中的黄连均具紫红花萼,叶片较大且有花斑;第 II 类包括的样本数量最多,包括 2 个小类, a 小类以大叶为主,叶片多光亮而具花斑,其分裂度较为混杂,花萼多呈紫红色,少数具黄绿花萼的样本较为

集中地聚在其中, b 小类叶片较大而羽片宽,多具花斑而不亮,花萼多紫红;第 III 类多为中等至较大叶片,羽片多为中等至较宽,花萼多呈紫红色;第 IV 类包括 c 和 d 2 个小类, c 小类羽片也多为中等至较宽,但叶片大小各异,叶片多不具光泽,花萼多紫红, d 小类叶片偏大,羽片宽度各异,叶片多



图 4 黄连遗传多样性的 UPGMA 聚类

Fig. 4 UPGMA cluster based on genetic diversity of *C. chinensis*

光亮而无花斑，花萼为紫红至深紫红色。聚类结果表明黄连遗传多样性丰富，亲缘关系较近的黄连在叶形和花色方面表现出较为一致的特征。

4 讨论

本研究从黄连转录组得到 93 629 条 Unigene，共检测到 12 848 个 SSR 位点，分布于 10 016 条 Unigene 序列中，出现频率为 13.72%，平均每 5.34 kb 序列出现 1 个 SSR 位点。与同科植物相比，其出现频率高于大叶铁线莲 (2.13%)^[13]，低于牡丹 (37.69%)^[14]。与其他药用植物相比，出现频率高于白木香 (7.27%)^[15]、川贝母 (8.49%)^[16]，低于防风 (18.51%)、川党参 (28.85%)^[17]、穿心莲 (32.36%)^[18]、冬虫夏草 (34.99%)^[19]等。SSR 位点的出现频率除了与物种本身有关，还与所检测的器官类型、转录组中 Unigene 长度、SSR 筛选标准等有关。即使在同一物种中，SSR 位点的出现频率也可能有差异，比如同样基于转录组分析，刘迪等^[20]报道的柴胡中 SSR 位点出现频率就明显低于王彬彬等^[21]的报道。同样本研究中黄连转录组中 SSR 位点的出现频率高于从琨^[4]报道的黄连转录组数据。总体来看黄连转录组中 SSR 位点出现频率较高，说明黄连转录组中 SSR 位点数量丰富，可为后期 SSR 引物开发及黄连遗传多样性分析提供资源。

黄连转录组中 SSR 基元类型丰富，包含了一至

六核苷酸重复，一至三核苷酸重复占 SSR 总数的 98.14%，而以单核苷酸重复和三核苷酸重复为主，分别占 53.81%和 25.01%。该结果与冬虫夏草^[19]SSR 主要基元类型一致，而与白木香^[15]、川贝母^[16]、穿心莲^[18]、柴胡^[19]等以二核苷酸和三核苷酸为主要重复基元有所差异。种子植物 SSR 重复类型以二核苷酸和三核苷酸重复为主，有学者认为早期微卫星的形成是通过碱基的替换或插入，再延伸^[22]或通过转座元件传播^[23-24]。早期微卫星更有可能在长度较短的基元中产生，这可能是二核苷酸重复数量众多的原因^[25]。而非 3 倍数的 SSR 基元在编码区会产生移码突变，这种突变压力限制了非 3 倍数的 SSR 基元的出现频率^[26]，从而使得三核苷酸重复在植物基因组中有很高的丰度^[27]。黄连转录组中的优势重复基元分别为 A/T、AG/CT、AAG/CTT，与白木香^[15]、川党参^[17]、密花香薷^[28]、苦参^[29]的优势重复基元一致。二核苷酸和三核苷酸重复中的优势基元也与从琨^[4]报道的黄连、云连一致。A/T、AG/CT、AT/AT、AAT/ATT、AAG/CTT 是植物 SSR 中极为普遍的组成方式，而 SSR 的重复基元类型和组成在不同的物种中表现出明显的差异，这可能是由于物种在分化过程中受到不同的选择压力，导致相应编码蛋白的使用频率不同所致^[25]。多数植物的 SSR 中存在 AT 优势，GC 基元很少，甚至没有^[25, 30]。GC 的含量是

基因结构和功能的重要影响因素,可能会使某些物种因氨基酸序列的变化而获得某些特定功能。冬虫夏草的转录组数据显示其 SSR 基元中, C/G、CCG/CGG 是优势重复单元^[19]。川贝母的三核苷酸重复中, CCG/CGG 也占有绝对优势^[16]。这些高 GC 含量的重复单元可能有助于 2 个物种适应高原环境。本研究从黄连转录组中检测到 GC 重复单元, 包括 198 个 C/G 重复, 11 个 CG/CG 重复和 79 个 CCG/CGG 重复。黄连多分布于海拔 1 200~1 800 m 的高山, 气候较为寒冷。这些 GC 重复单元是否在黄连适应其生境的过程中发挥正面作用还有待深入研究。

SSR 长度是评价 SSR 多态性的重要指标, 长度 ≥ 20 bp 的 SSR 具有较高的多态性, 长度位于 12~19 bp 的 SSR 多态性降低, 长度 < 12 bp 的 SSR 多态性最低^[12]。黄连转录组中 SSR 重复序列长度为 13~19 bp 的有 5 547 条, 占总量的 47.06%, 长度 ≥ 20 bp 的 SSR 有 2 465 条, 占总量的 20.91%, 这些多态性丰富的 SSR 序列可为后期的研究提供重要备选资源。本研究随机筛选设计的 180 对引物中, 有 94 对引物可扩增出清晰条带, 其中有 7 对引物在 64 份黄连样本中表现出良好的多态性。基于 UPGMA 的遗传多样性分析将 64 份黄连样本聚为 4 大类, 遗传关系较近的样本在叶形及花色方面表现出较为一致的特征, 然而各分支之间存在部分特征交差现象, 没有明显的形态特征的间断。黄连种内不同类型的划分, 目前多依据叶形差异^[6, 31]。黄连叶片三全裂, 中央裂片再羽状深裂, 侧裂片不等二深裂后, 小裂片再羽状分裂, 叶形复杂, 其分裂程度差异明显, 但不易量化。黄连花萼及花冠从深紫红至黄绿, 存在多个颜色差异类型, 同样不易量化。形态特征的统计可能存在误差。形态特征的量化分析及更多 SSR 多态性引物的开发将有助于更为客观地建立黄连形态特征与遗传信息之间的联系。

综上所述, 黄连转录组 SSR 位点出现频率较高, 重复类型丰富, 从中筛选的引物具有较高的多态性潜能。本研究表明基于转录组分析的黄连多态性 SSR 位点的开发具有可行性, 更多 SSR 标记的开发将为黄连种质的鉴定、黄连资源遗传多样性分析、分子辅助育种等研究工作提供研究基础。

利益冲突 所有作者均声明不存在利益冲突

参考文献

[1] 中国药典 [S]. 一部. 2020: 316.

- [2] 柳鑫, 黄河, 黄璐琦, 等. 黄连药材原植物资源和市场品种调查 [J]. 中国药师, 2014, 17(10): 1691-1695.
- [3] 崔学龙. 石柱黄连产业化发展的制约因素与对策研究 [J]. 产业与科技论坛, 2015, 14(6): 31-32.
- [4] 吴心悦, 丁可欣, 葛改变, 等. 胡黄连提取物化学指示性成分与感官指标变化规律研究 [J]. 中草药, 2023, 54(24): 8064-8076.
- [5] 黄正方, 杨美全, 孟忠贵, 等. 黄连生物学特性和主要栽培技术 [J]. 西南农业大学学报, 1994, 16 (3): 299-302.
- [6] 陈大霞, 王钰, 张雪, 等. ISSR 分析 3 种类型黄连的遗传多样性和遗传结构 [J]. 中草药, 2012, 43(8): 1595-1598.
- [7] 赵振宇. 黄连属植物物种鉴定及 mini-barcode 研究 [D]. 北京: 中国中医科学院, 2019.
- [8] 王刚, 曹佩, 韦学敏, 等. 分子标记技术在药用植物种质资源研究中的应用 [J]. 中国现代中药, 2019, 21(11): 1435-1444.
- [9] 宋芸, 张鑫瑞, 李政, 等. 基于 SSR 分子标记的柴胡遗传多样性与遗传结构分析 [J]. 药学学报, 2022, 57(4): 1193-1202.
- [10] 乔萍, 汪奕衡, 孙嘉惠, 等. 基于 SSR 的人参果色种质资源遗传多样性评估 [J]. 中国中药杂志, 2022, 47(8): 2158-2164.
- [11] 孙艺琦, 赵露颖, 朱波, 等. 基于 SSR 分子标记的蔓荆子基原植物的遗传多样性及遗传结构分析 [J]. 中国中药杂志, 2021, 46(15): 3824-3831.
- [12] Temnykh S, DeClerck G, Lukashova A, et al. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposon associations, and genetic marker potential [J]. *Genome Res*, 2001, 11(8): 1441-1452.
- [13] 宋琳琳, 王豆豆, 周晓君. 基于高通量测序的大叶铁线莲基因组微卫星特征分析 [J]. 分子植物育种, 2022, 20(4): 1087-1094.
- [14] 葛梅. 基于 SSR 标记与成分含量相结合的牡丹皮质量评价研究 [D]. 重庆: 西南大学, 2019: 8.
- [15] 向琳, 曾军, 王昊, 等. 基于白木香转录组的 SSR 序列特征分析 [J]. 分子植物育种, 2024, 22(9): 2971-2978.
- [16] 张婕, 李西文. 川贝母转录组中 SSR 位点信息分析 [J]. 中国实验方剂学杂志, 2018, 24(18): 30-35.
- [17] 蒋小刚, 周武先, 王华, 等. 川党参转录组 EST-SSR 位点分析 [J]. 西南农业学报, 2023, 36(6): 1132-1140.
- [18] 李俊仁, 陈秀珍, 汤小婷, 等. 穿心莲转录组 SSR 位点信息分析 [J]. 中国中药杂志, 2018, 43(12): 2503-2508.
- [19] 张涵, 王芳, 白静, 等. 冬虫夏草转录组 SSR 位点信息分析研究 [J]. 中国中药杂志, 2019, 44(21): 4605-4611.

- [20] 刘迪, 欧阳艳飞, 徐鹏, 等. 柴胡转录组 SSR 的分布及序列特征分析 [J]. 南方农业学报, 2022, 53(3): 676-683.
- [21] 王彬彬, 高妍夏, 郭欣慰, 等. 基于高通量测序的北柴胡根转录组 SSR 位点信息分析 [J]. 西北农业学报, 2024, 33(4): 707-717.
- [22] Messier W, Li S H, Stewart C B. The birth of microsatellites [J]. *Nature*, 1996, 381(6582): 483.
- [23] Wilder J, Hollocher H. Mobile elements and the genesis of microsatellites in dipterans [J]. *Mol Biol Evol*, 2001, 18(3): 384-392.
- [24] Buschiazzo E, Gemmell N J. The rise, fall and renaissance of microsatellites in eukaryotic genomes [J]. *Bioessays*, 2006, 28(10): 1040-1050.
- [25] Megléc E, Nève G, Biffin E, *et al.* Breakdown of phylogenetic signal: A survey of microsatellite densities in 454 shotgun sequences from 154 non model eukaryote species [J]. *PLoS One*, 2012, 7(7): e40861.
- [26] Metzgar D, Bytof J, Wills C. Selection against frameshift mutations limits microsatellite expansion in coding DNA [J]. *Genome Res*, 2000, 10(1): 72-80.
- [27] Haq S U, Jain R, Sharma M, *et al.* Identification and characterization of microsatellites in expressed sequence tags and their cross transferability in different plants [J]. *Int J Genomics*, 2014, 2014: 863948.
- [28] 富贵, 刘玉萍, 苏旭. 基于转录组数据的密花香薷 SSR 位点特征分析 [J]. 西北植物学报, 2021, 41(4): 654-663.
- [29] 张宁, 尹美强, 谭青青, 等. 苦参转录组 SSR 位点及基因功能注释分析 [J]. 江苏农业科学, 2019, 47(7): 41-44.
- [30] Sun L D, Yang W R, Zhang Q X, *et al.* Genome-wide characterization and linkage mapping of simple sequence repeats in Mei (*Prunus mume* Sieb. et Zucc.) [J]. *PLoS One*, 2013, 8(3): e59562.
- [31] 朱强, 王有为, 齐海涛, 等. 不同品系黄连产量和质量的研究 [J]. 中草药, 2006, 37(12): 1866-1869.

[责任编辑 时圣明]