

## 裸花紫珠群体遗传结构及其有效成分与 SSR 标记的关联分析

黄 梅<sup>1</sup>, 李 鑫<sup>1</sup>, 王 凯<sup>1</sup>, 江 芊<sup>1</sup>, 李 伟<sup>2</sup>, 陈振夏<sup>1\*</sup>, 于福来<sup>2\*</sup>

1. 中国热带农业科学院热带作物品种资源研究所 农业农村部中药材生物学与栽培重点实验室/海南省热带药用植物工程研究中心, 海南 海口 571101

2. 海南九芝堂药业有限公司, 海南 海口 570311

**摘要:** 目的 分析裸花紫珠 *Callicarpa nudiflora* 种质资源遗传变异情况, 发掘与有效成分含量相关联的分子标记, 以期为裸花紫珠新品种选育提供参考。方法 对 103 份裸花紫珠种质资源有效成分含量进行变异性分析, 并利用 14 对 SSR 标记对供试品进行遗传结构分析及连锁不平衡分析, 在此基础上采用 Talssel 2.1 软件的一般线性模型 (GLM) 和混合线性模型 (MLM) 进行 SSR 标记与有效成分含量的关联分析。结果 裸花紫珠中 7 种有效成分含量变异丰富, 可作为关联分析的理想材料。群体遗传结构分析结果表明,  $K=3$  时, 供试品分为 3 个亚群, 3 个亚群的遗传多样性均较为丰富, AMOVA 分析显示变异主要存在于个体内 (61%)。在 91 个 SSR 位点成对组合中, 不论共线性组合还是非共线性组合均存在一定程度的连锁不平衡 (LD)。D' 统计概率 ( $P<0.01$ ) 支持的 LD 成对位点有 21 个, 占全部位点组合的 23.08%。D' 值主要集中在 0.3~0.5 和 0.5~0.7, 占 SSR 标记全部对数的 84.62%, 表明所选种质间基因重组概率与突变概率较低。关联分析结果表明, 在 GLM 模型中, 发现有 11 对标记与 7 种有效成分含量呈显著相关 ( $P<0.05$ ) 或极显著相关 ( $P<0.01$ )。在 MLM 模型中, 6 对标记与 4 种有效成分含量呈显著相关 ( $P<0.05$ ) 或极显著相关 ( $P<0.01$ )。在 GLM 和 MLM 模型中, 均检测到 CnH305 与连翘酯苷 B 和木犀草苷相关联 ( $P<0.05$ ), CnH155 与木犀草苷和毛蕊花糖苷相关联 ( $P<0.01$ ), CnH156、CnH357 和 CnH205 与异毛蕊花糖苷相关联。结论 供试裸花紫珠种质存在丰富的遗传变异, 种质间存在一定程度的连锁不平衡, 5 对 SSR 标记能够很好的关联到药效成分含量, 可以为裸花紫珠分子标记辅助育种提供理论依据。

**关键词:** 裸花紫珠; SSR 标记; 群体结构; 连翘酯苷 B; 木犀草苷; 毛蕊花糖苷; 异毛蕊花糖苷; 关联分析

中图分类号: R286.12 文献标志码: A 文章编号: 0253-2670(2024)23-8162-09

DOI: 10.7501/j.issn.0253-2670.2024.23.022

## Analysis of population genetics and effective components associations with SSR markers in *Callicarpa nudiflora*

HUANG Mei<sup>1</sup>, LI Xin<sup>1</sup>, WANG Kai<sup>1</sup>, JIANG Qian<sup>1</sup>, LI Wei<sup>2</sup>, CHEN Zhenxia<sup>1</sup>, YU Fulai<sup>2</sup>

1. Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences/Key Laboratory of Biology and Cultivation of Herb Medicine (Haikou), Ministry of Agriculture and Rural Affairs, China/ Hainan Provincial Engineering Research Center for Tropical medicinal plants, Haikou 571101, China

2. Hainan Jiuzhitang Pharmaceuticals Co., Ltd., Haikou, 570311, China

**Abstract: Objective** The genetic variation of *Callicarpa nudiflora* germplasm resources was analyzed, and the molecular markers associated with the content of effective components were discovered, so as to provide references for the breeding of new *C. nudiflora* varieties. **Methods** Variability analysis was conducted on the content of effective components in 103 *C. nudiflora* germplasm resources, and 14 SSR markers were used for genetic structure and linkage disequilibrium analysis. Based on this, association analysis between SSR markers and the content of effective components was conducted using the general linear model (GLM) and mixed linear model (MLM) of TASSEL 2.1 software. **Results** The results showed rich variability in the content of seven effective components in *C. nudiflora*, making it an ideal material for further association analysis. Population genetic structure analysis indicated that when  $K=3$ , the test materials were divided into three subgroups, each

收稿日期: 2024-05-21

基金项目: 海南省重大科技计划项目 (ZDKJ2021001); 海南省自然科学基金青年基金 (322QN393); 海南省重点研发项目 (ZDYF2024XDNY216)

作者简介: 黄 梅 (1989—), 女, 助理研究员, 主要从事南药种质资源评价及品种培育研究。E-mail: huangmei122@126.com

\*通信作者: 于福来 (1983—), 男, 研究员, 主要从事南药资源定向培育研究。E-mail: fulai.yu@163.com

陈振夏 (1985—), 男, 农艺师, 主要从事南药资源收集、保存与评价研究。E-mail: hnchenzhenxia@126.com

with rich genetic diversity. AMOVA analysis showed that the variation mainly existed within individuals (61%). Among 91 paired SSR loci, both collinear and non-collinear combinations exhibited a certain degree of linkage disequilibrium (LD). Twenty-one paired loci supported by D' statistic probability ( $P < 0.01$ ) accounted for 23.08% of all loci combinations. D' values mainly ranged between 0.3—0.5 and 0.5—0.7, accounting for 84.62% of the total SSR marker pairs, indicating a lower probability of gene recombination and mutation among the selected germplasm. Association analysis revealed that in the GLM model, 11 markers were significantly or extremely significantly correlated with seven effective component contents ( $P < 0.05, 0.01$ ). In the MLM model, six markers were significantly or extremely significantly correlated with four effective component contents ( $P < 0.05, 0.01$ ). In GLM and MLM models, CnH305 was found to be associated with forsythoside B and luteoloside ( $P < 0.05$ ), CnH155 was associated with luteoloside and acteoside ( $P < 0.01$ ), and CnH156, CnH357, and CnH205 were associated with isoacteoside. **Conclusion** The tested *C. nudiflora* germplasm exhibits rich genetic variation and a certain degree of linkage disequilibrium. Five SSR markers were well associated with the content of effective components, providing a theoretical basis for molecular marker-assisted breeding of *C. nudiflora*.

**Keywords:** *Callicarpa nudiflora* Hook. et Arn; SSR markers; population structure; forsythoside B; luteoloside; acteoside; isoacteoside; association analysis

裸花紫珠 *Callicarpa nudiflora* Hook. et Arn 为唇形科紫珠属植物,主产于海南、广东、广西等省,收载于 1977 年和 2020 年版《中国药典》,也是 2015 年版《中国药典》第一增补本收录品种,主要以干燥叶片入药,是海南省道地药材,具有消炎、止血、抑菌等功效<sup>[1-2]</sup>。裸花紫珠在黎族、苗族、壮族有着悠久的用药历史,素有“海南黑药”之称,是“裸花紫珠片、裸花紫珠分散片、裸花紫珠胶囊”等多种中成药的主要原料,在医药领域有着广阔的应用前景<sup>[3]</sup>。药农目前种植的均为野生驯化种,生产上并无优势品种出现,导致药材质量参差不齐。优良的品种是质量稳定的基础,也是实现裸花紫珠规范化生产的保证。但目前品种选育多以常规方法为主,育种周期长,育种效率低,育成品种遗传基础较窄<sup>[4]</sup>。而分子辅助育种可大大加快育种进程,因此研究裸花紫珠主要有效成分含量的遗传规律,挖掘与之显著关联的分子标记,对裸花紫珠有效成分含量高的专用型品种的选育具有重要意义。

裸花紫珠含单体成分近 200 种,成分复杂,谭湘杰等<sup>[5]</sup>通过分析文献,以被定量过、具有止血、消炎活性且具有种属特异性为条件,筛选了毛蕊花糖苷、异毛蕊花糖苷、连翘酯苷 B 等 7 种可作为药材潜在的质量控制指标。关联分析是鉴定群体目标性状与遗传标记关系的方法,利用关联分析法挖掘植物的数量性状基因已经成为研究植物基因组学的常用方法。SSR 标记是关联分析常用的分子标记之一。吴怡超等<sup>[6]</sup>研究发现,12 对 SSR 标记与苦莽芦丁含量显著关联,其中 TatG0164 和 SWU\_Ft394 在两个模型中均与芦丁含量呈显著性关联。朱翔宇等<sup>[7]</sup>研究发现,共检测到 11 个标记与红三叶种质的

超氧化物歧化酶活性 (superoxide dismutase, SOD)、过氧化物酶活性 (peroxidase, POD)、过氧化氢酶活性 (catalase, CAT) 等 5 个与抗寒相关的相对生理生化指标相关联。于福来等<sup>[8]</sup>通过对裸花紫珠进行基因组大小、杂合率及 SSR 特征进行分析,发现裸花紫珠为高杂合、高重复的复杂基因组,且单、双、三核苷酸重复基序占比较高。张红瑞等<sup>[9]</sup>基于 SSR 分子标记对裸花紫珠种质资源进行遗传多样性分析,筛选了 14 对引物,并证实裸花紫珠资源具有丰富的遗传多样性且构建了 DNA 指纹图谱。裸花紫珠 SSR 标记已经开始陆续被开发,但主要用于遗传多样性分析及指纹图谱构建方面<sup>[9-10]</sup>,对于分子标记及有效成分含量关联分析研究目前并无报道。

基于此,本研究检测了来自我国裸花紫珠主要分布区的 103 份裸花紫珠种质资源的 7 种有效成分含量,利用 SSR 标记,分析裸花紫珠种质资源遗传变异情况,发掘与有效成分含量相关联的分子标记,以期分子标记辅助选育有效成分含量高的专用型品种提供理论依据。

## 1 材料与仪器

### 1.1 材料

从海南、广东、广西等地共收集到 103 份裸花紫珠种质资源,均保存于中国农业部(儋州)热带药用植物种质资源圃,由中国热带农业科学院热带作物品种资源研究所于福来研究员鉴定为唇形科紫珠属植物裸花紫珠 *C. nudiflora* Hook. et Arn, 信息见表 1。

### 1.2 仪器与试剂

1260 型高效液相色谱(美国安捷伦科技公司), Sartorius CPA225D 型电子分析天平(德国赛多利斯集团); KQ-500DB 型超声仪(昆山市超声仪器有限

表 1 不同来源裸花紫珠材料

Table 1 Different sources of *C. nudiflora* materials

来源地	份数	种质类型	来源地	份数	种质类型
海南省白沙县	12	野生资源、半野生资源	海南省三亚市天涯区	1	野生资源
海南省东方市	11	野生资源	海南省三亚市吉阳区	2	野生资源
海南省临高县	8	野生资源	海南省陵水县	3	野生资源、半野生资源
海南省五指山市	8	野生资源、半野生资源	海南省文昌市	1	野生资源
海南省海口市秀英区	5	野生资源	海南省琼海市	1	野生资源
海南省海口市龙华区	2	野生资源	海南省澄迈县	1	半野生资源
海南省乐东县	6	野生资源	海南省保亭县	1	半野生资源
海南省定安县	6	野生资源	广东省湛江市麻章区	1	野生资源
海南省琼中县	6	野生资源	广东省茂名市电白区	4	野生资源
海南省昌江县	5	野生资源	广西省北海市银海区	1	野生资源
海南省万宁市	5	野生资源、半野生资源	广西省南宁市兴宁区	1	半野生资源
海南省屯昌县	5	野生资源、半野生资源	广西省钦州市灵山县	1	野生资源
海南省儋州市	4	野生资源	广西省钦州市钦南区	2	野生资源

半野生资源指原来生长在野外，后被人发现并移植，对其进行了修剪、除草、施肥等相关培育处理的资源。

Semi-wild resources refer to resources that originally grew in the wild but were later discovered and transplanted by people, who then carried out related cultivation treatments such as pruning, weeding, and fertilizing.

公司), KingFisher™ Flex 型核酸提取仪(赛默飞有限公司), Mikro 120 型台式离心机(MIKRO 公司), VDY-6C 型电泳仪(北京六一生物科技有限公司), eriti 96 型 PCR 仪(美国应用生物系统公司), Gel Doc™ XR+型紫外分析仪(美国伯乐公司), 3730XL 型基因分析仪(美国应用生物系统公司)

对照品咖啡酸(批号 Y17D6C7672)、连翘酯苷 B(批号 20140120)、木犀草苷(批号 Y23S7H21785)、毛蕊花糖苷(批号 PMO321SA14)、异毛蕊花糖苷(批号 HA0820KA29)、芹菜素-7-O-β-D-葡萄糖苷(以下简称芹菜素糖苷,批号 P20N6F6297)、木犀草素(批号 YO1D6S6815)均采购于上海源叶生物制药公司,且质量分数均大于 98%;乙腈为色谱纯(安徽天地高纯溶剂有限公司),其余试剂均为分析纯,水为 Merck Millipore 纯水机制备的超纯水。2×Taq PCR Master Mix(美国基因泰克公司),琼脂糖(美国基因泰克公司), DL2000 DNA Marker(美国基因泰克公司),磁珠法基因组 DNA 提取试剂盒(天根生化有限公司),普通引物/荧光标记引物(天一辉远有限公司), GeneScan 500 LIZ Size Standard(美国应用生物系统公司), POP-7 Polymer(美国应用生物系统公司), Hi-Di Formamide(美国应用生物系统公司)。

## 2 方法

### 2.1 有效成分含量测定

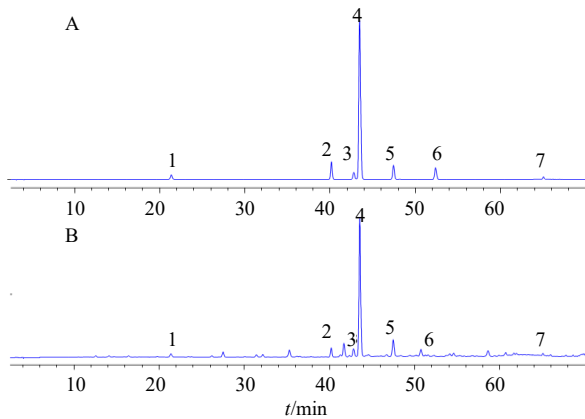
参考本课题组谭湘杰等<sup>[1]</sup>建立的裸花紫珠 7 种有效单体成分测定方法,对 103 份裸花紫珠种质资

源的 7 个有效成分含量进行测定,每份样品设 3 个重复,取平均值。

**2.1.1 混合对照品溶液的制备** 分别精密称取适量咖啡酸、木犀草素对照品,置于不同 10mL 容量瓶中,加入 70% 色谱甲醇溶解定容。分别精密称取适量连翘酯苷 B、木犀草苷、毛蕊花糖苷、异毛蕊花糖苷、芹菜素糖苷对照品,置于同一 10 mL 量瓶中,加入适量 70% 色谱甲醇溶解后再加入上述咖啡酸和木犀草素对照品溶液各 1 mL 定容。即配制成含咖啡酸 0.016 mg/mL、连翘酯苷 B 0.23 mg/mL、木犀草苷 0.062 mg/mL、毛蕊花糖苷 2.01 mg/mL、异毛蕊花糖苷 0.25 mg/mL、芹菜素糖苷 0.16 mg/mL、木犀草素 0.02 mg/mL 的混合对照品母液。

**2.1.2 供试品溶液的制备** 精密称取裸花紫珠粉末(过 20 目筛) 0.5 g,置于 100 mL 离心管中,加入 70% 甲醇 25 mL,在 40 kHz、400 W 条件下超声提取 20 min,吸取上层清液,0.22 μm 微孔滤膜过滤,取续滤液,即得供试品溶液。

**2.1.3 色谱条件<sup>[1]</sup>** Agilent-TC C<sub>18</sub> 色谱柱(250 mm×4.6 mm, 5 μm);流动相为乙腈(A)-0.4%磷酸(B),梯度洗脱为 0~30 min, 8%~16%A; 30~45 min, 16%~21%A; 45~55 min, 21%~24%A; 55~65 min, 24%~35%A; 65~70 min, 35%~100%A; 体积流量 1.0 mL/min; 检测波长 330 nm; 柱温 30 °C; 进样量 10 μL。色谱图见图 1。



1-咖啡酸; 2-连翘酯苷 B; 3-木犀草苷; 4-毛蕊花糖苷; 5-异毛蕊花糖苷; 6-芹菜素糖苷; 7-木犀草素。  
1-caffeic acid; 2-forsythiaside; 3-galuteolin; 4-acteoside; 5-isoacteoside; 6-apigenin-7-O-β-D-glucopyranoside; 7-luteolin.

图 1 混合对照品 (A) 和供试品 (B) 液相色谱图

Fig. 1 Liquid chromatogram of reference substances (A) and *C. nudiflora* sample (B)

表 2 7 种有效成分标准曲线方程及线性范围

Table 2 Standard curve equations and linear ranges of seven effective components

成分	回归方程	线性范围/(mg·mL <sup>-1</sup> )	R <sup>2</sup>
咖啡酸	$Y=44\ 983 X-1.702\ 7$	0.000 16~0.016	0.999 8
连翘酯苷 B	$Y=11\ 042 X-0.562$	0.002 3~0.23	0.999 9
木犀草苷	$Y=19\ 260 X+0.019\ 5$	0.000 62~0.062	0.999 5
毛蕊花糖苷	$Y=3\ 728 X+45.561$	0.020 1~2.01	0.999 8
异毛蕊花糖苷	$Y=13\ 202 X-3.223\ 3$	0.002 5~0.25	0.999 9
芹菜素	$Y=17\ 790 X-39.022$	0.001 6~0.16	0.999 5
木犀草素	$Y=30\ 845 X-2.370\ 3$	0.000 2~0.02	0.999 5

伸 30 s, 10 个循环, 每个循环下降 1 °C; 95 °C 变性 30 s, 52 °C 退火 30 s, 72 °C 延伸 30 s, 25 个循环; 72 °C 延伸 20 min, 4 °C 保存<sup>[9]</sup>, 基因分型数据来源于本课题组前期 103 份裸花紫珠种质资源 (即本试验使用的 103 份) 基因分型<sup>[9]</sup>, 用“0-1”读带法标记条带, 在相同水平线上的条带, 有条带记为“1”, 无条带记为“0”, 形成二元矩阵。

### 2.3 数据分析

采用 EXCEL 2010 和 SPSS 20.0 分析有效成分含量相关数据, 采用 GenALEx 6.502 将 SSR 分子标记的结果转换为 TXT 文本格式。使用 Structure 2.3.3<sup>[12]</sup> 对 103 份裸花紫珠种质资源进行群体结构分析, 将群体数量 (number of populations, *K* 值) 设置为 2~8, 每个 *K* 值重复运算 5 次, 将不作数迭代 LPB (length of burn in period) 和 MCMC (markov chain monte carlo) 均设为 100 000。并利用 Structure 2.3.3 程序

2.1.4 线性关系考察 取“2.1.1”项下混合对照品母液分别稀释 0、2.5、5、10、50、100 倍, 制成相应浓度梯度对照品溶液。以“2.1.3”项下色谱条件进样分析, 并以浓度为横坐标 (*X*), 峰面积为纵坐标 (*Y*), 进行线性回归, 得到相应的回归方程, 见表 2。

2.1.5 方法学考察 参考本课题组谭湘杰等<sup>[11]</sup>之前实验方法, 考察精密度、稳定性、重复性及加样回收率, 均符合实验要求。

### 2.2 SSR 标记分型

本实验所用 SSR 引物共 14 对, 由本课题组前期根据裸花紫珠转录组 Unigene 序列开发, PCR 反应体系 (10 μL): 2×Taq PCR Master Mix 5.0 μL, 基因组 DNA (约 20 ng) 1 μL, 上游引物 (浓度 10 pmol/μL) 0.5 μL, 下游引物 (浓度 10 pmol/μL) 0.5 μL, ddH<sub>2</sub>O 3.0 μL。反应程序为: 95 °C 预变性 5 min; 95 °C 变性 30 s, 62~52 °C 梯度退火 30 s, 72 °C 延

(<https://taylor0.biology.ucla.edu/structure/Harvester/>) 处理, 得到最佳 *K* 值和群体归属系数 (membership coefficient, *Q*) 值, 将 *Q* 值 ≥ 0.6 的种质划分到相应亚群中。使用 TASSEL 2.1 软件计算 SSR 位点组合间的连锁不平衡 (linkage disequilibrium, LD) 水平及支持概率, 绘制 LD 配对检测矩阵图。采用 TASSEL 2.1 软件对 103 份裸花紫珠种质的 7 个有效成分含量和 SSR 标记进行关联分析, 分别使用一般线性模型 (general linear model, GLM) 和混合线性模型 (mixed linear model, MLM), 计算 *P* 值和 *R*<sup>2</sup> 值, 如果 *P* < 0.05, 则说明该标记与该性状存在显著关联; 若 *P* < 0.01, 则说明该标记与该性状存在极显著关联<sup>[13]</sup>。

## 3 结果与分析

### 3.1 有效成分含量变异分析

对 103 份裸花紫珠种质资源的 7 种有效成分含量进行差异性分析 (表 3), 结果表明各裸花紫珠

表 3 7 种有效成分含量变异情况

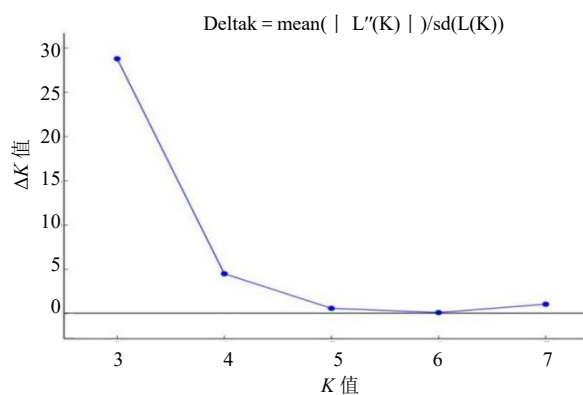
Table 3 Variation of content of seven effective components

有效成分	极小值/(mg·g <sup>-1</sup> )	极大值/(mg·g <sup>-1</sup> )	平均值/(mg·g <sup>-1</sup> )	中位数/(mg·g <sup>-1</sup> )	RSD/%	CV/%
咖啡酸	0.05	0.16	0.10	0.10	0.02	22.60
连翘酯苷 B	1.43	9.14	4.62	4.28	1.74	37.54
木犀草苷	0.51	3.79	2.17	2.09	0.68	31.19
毛蕊花糖苷	25.79	85.70	50.90	48.60	13.76	27.03
异毛蕊花糖苷	2.32	9.06	4.61	4.40	1.48	32.18
芹菜素糖苷	0.14	0.61	0.26	0.24	0.08	33.08
木犀草素	0.08	0.50	0.25	0.26	0.09	36.22

种质资源有效成分含量差异明显, 变异丰富, 可作为进一步关联分析的理想材料。连翘酯苷 B、木犀草苷、异毛蕊花糖苷、芹菜素糖苷和木犀草素的变异系数均大于 30%, 其中以连翘酯苷 B 的变异系数 (CV) 最高为 37.54%。咖啡酸的 CV 最小, 为 22.60%, 说明裸花紫珠种质资源的 7 个有效成分含量差异显著。

### 3.2 群体遗传结构分析

课题组前期对 103 份裸花紫珠种质资源进行 PCR 扩增及电泳检测, 获得 14 对多态性引物基因分型数据, 且研究结果显示 14 对引物在 103 份种质中共扩增出 92 个等位基因, 平均每对引物扩增 6.571 个, 变化范围为 3~13 个, 有效等位基因数量 (effective number of alleles,  $N_e$ ) 变化范围为 1.081 8~4.947 2, 平均多态性信息含量 (polymorphism information content, PIC) 平均值为 0.468 2, 高 PIC 值的引物占 42.9%, 表明 14 对引物可以从分子水平解释基因型差异, 具有丰富的遗传差异性<sup>[9]</sup>。本研究基于这 14 对引物的基因型数据, 采用 Structure 2.3.3 软件对 103 份裸花紫珠种质资源进行种群结构分析。结果表明,  $K=3$  时,  $\Delta K$  值为峰值 (图 2), 因此将 103 份裸花紫珠种质划分为 3 个亚群, 分别命名为 POP1、POP2 和 POP3 (图 3)。通过  $Q$  值分析发现, 其中 75 份种质的  $Q$  值  $\geq 0.6$ , 表明遗传背景相对单一, 可划分在 3 个亚群的某一亚群中。POP1 有 22 个种质 (红色部分); POP2 有 30 份种质 (绿色部分); POP3 有 23 个种质 (蓝色部分)。而其余 28 份种质  $Q$  值  $< 0.6$ , 说明这些种质的遗传背景比较复杂, 没有明确亚群归属特性, 形成一个混合群体。结果表明, 该供试种质群体结构较为单一, 可有效降低群体结构对关联分析产生的影响。从产地

图 2 基于 Structure 分析估算的  $\Delta K$  值

来看, 种群结构分析有一定的规律: POP1 将广东、广西 2 个省份以及海南定安的大部分种质聚在一起, 表明定安的种质与该 2 个省份的种质亲缘关系较近; POP2 来源较为丰富, 主要集中了来自白沙、五指山、东方和乐东的种质; POP3 主要集中了来自临高和海口的种质。种质分类结果表明种群结构的分类与地理来源具有一定相关性。

通过 GenALEx 软件评估了裸花紫珠种群结构的  $N_e$ 、Shannon 信息指数 (shannon's information index,  $I$ ) 等遗传多样性信息 (表 4)。结果表明 POP1、POP2 和 POP3 的遗传多样性均较为丰富, 其中 POP2 的等位基因数 (number of alleles,  $N_a$ )、 $N_e$ 、 $I$ 、观察杂合度 (observed heterozygosity,  $H_o$ ) 和期望杂合度 (expected heterozygosity,  $H_e$ ) 均最高, 分别为 5.500、3.191、1.191、0.514 和 0.588。3 个亚群的  $H_o$  均小于  $H_e$ , 说明杂合度不高。AMOVA 分析显示 (表 5), 遗传组间变异率为 16%, 个体间变异率为 23%, 个体内遗传变异率为 61%, 表明变异主要存在于个体内。

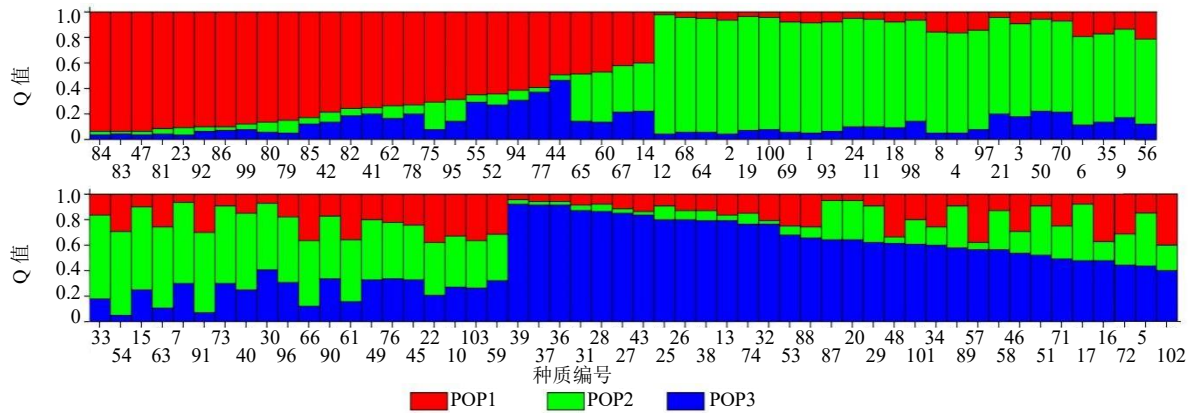


图3 103份种质的群体结构图

Fig. 3 Population structure map of 103 accessions

表4 各亚群遗传多样性参数

Table 4 Genetic diversity parameters of each subpopulation

亚群	<i>N</i>	<i>N<sub>a</sub></i>	<i>N<sub>e</sub></i>	<i>I</i>	<i>H<sub>o</sub></i>	<i>H<sub>e</sub></i>
POP1	22	3.000	1.733	0.599	0.190	0.331
POP2	30	5.500	3.191	1.191	0.514	0.588
POP3	23	3.500	1.797	0.656	0.248	0.353

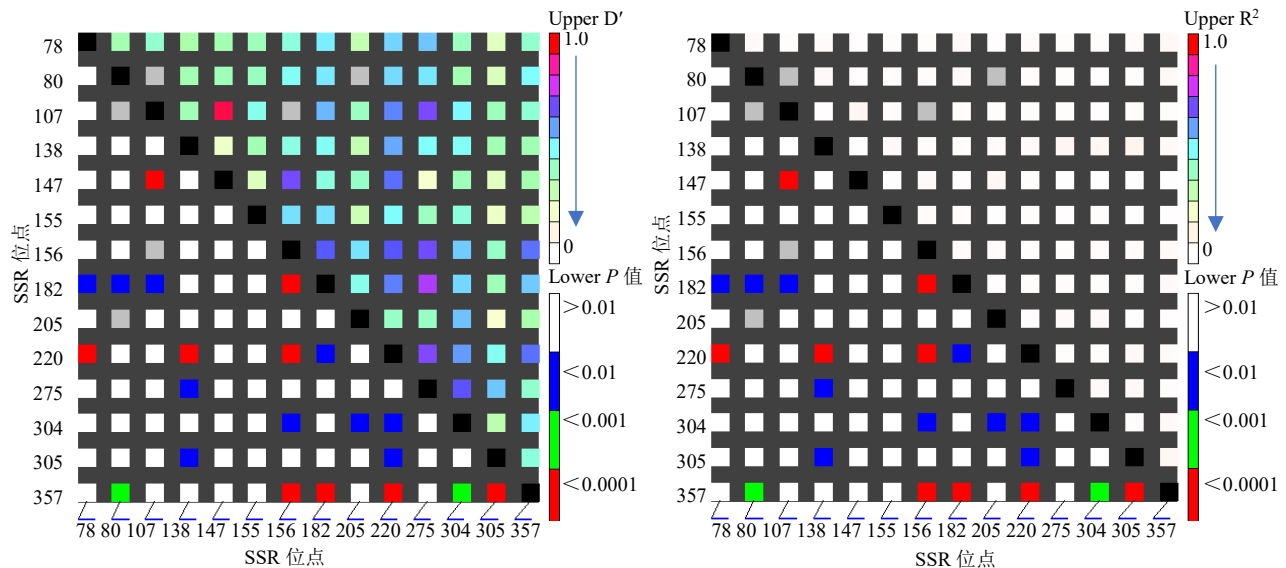
### 3.3 LD 分析

不同等位基因间的连锁不平衡分析是关联分析的基础，图4展示了使用2个不同的参数 *D'* 和 *R*<sup>2</sup> 来描述 LD 矩阵的结果，其中 *D'* 描述群体的重组历史，*R*<sup>2</sup> 反映群体的突变历史和重组历史。在 *C*<sub>14</sub><sup>2</sup>=91 个 SSR 位点成对组合中，不论共线性组合还是非共线性

表5 不同遗传群体间的分子变异分析

Table 5 Analysis of molecular variance among different genetic groups

来源	自由度	平方和	均方	估算的差异值	变异百分比/%
种群间	2	70.556	35.278	0.630	16
个体间	72	293.337	4.074	0.867	23
个体内	75	175.500	2.340	2.340	61
总和	149	539.393	—	3.838	100



X、Y 轴为 SSR 位点分布顺序，黑色对角线上方像素格分别表示成对位点间 *D'* 值和 *R*<sup>2</sup> 值，对角线下方像素格表示成对位点 LD 支持率。SSR markers are made along the X and Y axes. Each pixel above the black diagonal represent the *D'* and *R*<sup>2</sup> values of each marker pair, respectively, and each pixel below the diagonal represent the p-value of the testing LD of each marker pair.

图4 裸花紫珠 14 个 SSR 位点中的连锁不平衡分布

Fig. 4 Distribution of LD among 14 SSR locus on *C. nudiflora*

组合均存在一定程度的连锁不平衡。从图中可以看出, 2 个不同参数描述的 LD 矩阵的颜色较浅, LD 较低, 但  $D'$  值描述的 LD 比  $R^2$  描述的 LD 更明显。 $D'$  统计概率 ( $P < 0.01$ ) 支持的 LD 成对位点有 21 个, 占全部位点组合的 23.08%,  $D'$  值大于 0.5 表示较高的连锁不平衡程度, 本研究  $D'$  主要集中在 0.3~0.5 和 0.5~0.7, 占 SSR 标记全部对数的 84.62%, 表明供试种质间基因重组概率与突变概率较低, 这可能与所收集材料为野生或半野生材料有关。

### 3.4 SSR 标记与有效成分含量的关联分析

分别以群体结构分析的 Q 值与 (Q+K) 值作为协变量, 采用 Talssel 2.1 软件的 GLM 和 MLM 对裸花紫珠的 14 对 SSR 标记与 7 种有效成分含量进行关联分析, 挖掘与有效成分相关联的标记 (表 6)。结果表明, 在 GLM 模型中, 发现有 11 对标记与 7 种有效成分含量呈显著相关 ( $P < 0.05$ ) 或极显著相关 ( $P < 0.01$ ), 变异解释率范围为 6.31%~44.64%, 其中异毛蕊花糖苷和标记 CnH156 关联效果最佳, 变异解释率最高 (44.64%)。在 MLM 模型中, 6 对标记与 4 种有效成分含量呈显著相关 ( $P < 0.05$ ) 或极显著相关 ( $P < 0.01$ ), 变异解释率范围为 17.65%~36.1%, 其中异毛蕊花糖苷和引物 CnH156

的关联效果最佳, 变异解释率为 36.1% (相关引物信息见表 7)。MLM 模型较 GLM 得到的关联数据少, 且大部分对应的解释率有所下降, 说明与 GLM 模型相比, MLM 因同时考虑种群结构 (Q) 及亲缘关系 (K), 所以能避免一部分假阳性关联, 提高分析准确性。

在 GLM 和 MLM 2 种模型中, 均检测到 CnH305 与连翘酯苷 B 和木犀草苷相关联 ( $P < 0.05$ ), CnH155 与木犀草苷和毛蕊花糖苷相关联 ( $P < 0.01$ ), 说明标记位点可能存在“一因多效”, 或者该位点可能与多个功能基因存在连锁关系。同时, 2 种模型中, 均检测到 CnH156、CnH357 和 CnH205 与异毛蕊花糖苷相关联, 说明异毛蕊花糖苷含量受多个基因控制。

## 4 讨论

### 4.1 裸花紫珠种质资源 7 种药效物质含量的变异情况

裸花紫珠在临床上常作为单方入药, 以其水提取物为原料制成的裸花紫珠片、裸花紫珠分散片和裸花紫珠胶囊等常被用于治疗各种呼吸道、消化道出血、外伤出血及跌打肿痛等。其药效物质主要为苯乙醇苷类, 《中国药典》2020 年版中以毛蕊花糖苷

表 6 GLM 和 MLM 模型中与有效成分含量显著相关联的 SSR 标记

Table 6 SSR markers associated with effective component contents in GLM and MLM model

有效成分	GLM (Q)			有效成分	MLM (Q+K)		
	SSR 标记	P	R <sup>2</sup> /%		SSR 标记	P	R <sup>2</sup> /%
咖啡酸	CnH078	0.208 1	9.44*	连翘酯苷 B	CnH305	0.033 8	30.10*
连翘酯苷 B	CnH078	0.044 4	13.74*	木犀草苷	CnH155	1.80 × 10 <sup>-3</sup>	29.38**
	CnH305	0.011 5	10.51*		CnH305	0.016 7	35.76*
木犀草苷	CnH155	6.53 × 10 <sup>-4</sup>	31.55**	毛蕊花糖苷	CnH155	8.50 × 10 <sup>-3</sup>	24.42**
	CnH305	0.0167	9.78*		CnH304	0.045 1	17.65*
毛蕊花糖苷	CnH155	6.60 × 10 <sup>-3</sup>	6.47**	异毛蕊花糖苷	CnH156	3.48 × 10 <sup>-5</sup>	36.10**
异毛蕊花糖苷	CnH155	0.015 1	24.40*		CnH205	0.016	24.30*
	CnH156	4.42 × 10 <sup>-5</sup>	44.64**	CnH357	1.80 × 10 <sup>-3</sup>	32.40**	
	CnH182	5.40 × 10 <sup>-3</sup>	34.67**	—	—	—	
	CnH205	7.00 × 10 <sup>-3</sup>	21.53**	—	—	—	
	CnH357	4.85 × 10 <sup>-4</sup>	26.55**	—	—	—	
	芹菜素糖苷	CnH078	1.20 × 10 <sup>-4</sup>	26.09**	—	—	—
		CnH107	0.010 9	31.04*	—	—	—
CnH147		0.038 4	6.31*	—	—	—	
CnH182		0.043 4	28.88*	—	—	—	
CnH220		0.012 8	32.45*	—	—	—	
CnH305		0.017 4	9.69*	—	—	—	
木犀草素	CnH357	9.68 × 10 <sup>-4</sup>	25.18**	—	—	—	
	CnH080	4.26 × 10 <sup>-4</sup>	26.81**	—	—	—	

\*\*表示极显著相关 ( $P < 0.01$ ); \*表示显著相关 ( $P < 0.05$ )。

\*\*indicates extremely significant correlation ( $P < 0.01$ ); \*indicates significant correlation ( $P < 0.05$ ).

表 7 与有效成分显著相关的 SSR 引物序列信息

Table 7 Sequence information of SSR primers significantly correlated with active components

引物名称	重复基元	引物序列 (5'-3')	片段大小/bp	荧光基团
CnH078	(TCC) <sub>5</sub>	F: AGGCACCTAGCTAGACACCA R: ACGCTCAGGTAGTTGGTGTG	177~184	FAM
CnH080	(GAT) <sub>7</sub>	F: ACATCCGGTGCATCCTTTCT R: TTCCTTCAAGCACCTGTCCC	292~301	FAM
CnH107	(ATCA) <sub>5</sub>	F: ACGCCTTCCATTATGCACGA R: GCCCGAGTTTGTAGTGGTGA	137~148	HEX
CnH147	(TGA) <sub>5</sub>	F: CCGAGGCAAAGTTCCAACCA R: CCACTGATAAACCAACGAGCC	157~162	FAM
CnH155	(TCT) <sub>13</sub>	F: CACAGAAACAACACGGCCAG R: GCGTTTCGACTGCATTCGAT	259~277	HEX
CnH156	(AAG) <sub>12</sub>	F: GGCCCTGATGCTTCTTCTGT R: CCTTATAACGCACGCATGGC	277~301	TAMRA
CnH182	(TCT) <sub>15</sub>	F: ATCTCTGTCTCACAAAGCGCC R: CACCGGAACCCTGCATGTAA	251~286	FAM
CnH205	(AAT) <sub>13</sub>	F: GGTATGTCTCATGTCCCGA R: TTGCGTTCACGATGCTTTGG	250~269	HEX
CnH220	(TTC) <sub>14</sub>	F: CCCAAGAAGAGCTGTCCACA R: CCTCAGCATGCAGAACTGG	186~204	HEX
CnH304	(ATA) <sub>11</sub>	F: ACCAAGCTGCAATTCCTCAAG R: ACCAAATCTTAGGTCCACATCCT	276~298	HEX
CnH305	(CCA) <sub>8</sub>	F: CAGCGCAGGATTGTTGATCA R: GGGTAAACAGTGGGTGGAGG	264~270	FAM
CnH357	(AACC) <sub>5</sub>	F: CCCACCAGTGAAAGCTTCT R: CGATGTGGCTGTGATCGTA	210~222	HEX

和木犀草苷作为含量测定的指标物质，课题组前期对裸花紫珠化学成分进行了系统性分析，筛选出了包含这 2 种物质在内的 7 种有效成分<sup>[5]</sup>。本研究结果显示连翘酯苷 B、木犀草苷、异毛蕊花糖苷、芹菜素糖苷和木犀草素的 CV 均大于 30%，其中以连翘酯苷 B 的 CV 最高为 37.54%。咖啡酸的 CV 最小，为 22.60%，表明裸花紫珠种质资源有效成分含量差异明显，变异丰富。课题组前期对海南产的 29 份裸花紫珠毛蕊花糖苷与木犀草苷含量进行分析，结果发现不同产地间成分均具有显著差异<sup>[14]</sup>。本研究选取了裸花紫珠主要分布区，即海南、广东、广西 3 省的 103 份裸花紫珠，发现不同来源的裸花紫珠主要成分具有显著的差异。

#### 4.2 裸花紫珠种群遗传结构分析

课题组前期研究结果表明所选的 14 对引物可以从分子水平解释基因型差异，具有丰富的遗传多样性<sup>[9]</sup>。群体结构分析是进行关联分析的前提，过于复杂的群体结构会导致假阳性关联，因此简单的群体结构有助于获得可靠的关联结果<sup>[15]</sup>。本研究采用 Structure 2.3.3 软件对 103 份裸花紫珠种质资源进行种群结构分析，其中 75 份种质的 Q 值 ≥ 0.6，

遗传背景单一，可划分在 3 个亚群的某一亚群中。而 28 份种质的 Q 值 < 0.6，遗传背景较为复杂，归为混合群体<sup>[16]</sup>。整个群体结构相对单一，有利于提高关联分析结果的准确性。且 AMOVA 分析结果表明变异主要存在于个体内 (61%)，个体内变异是指由杂合的等位基因引起的遗传变异，变异百分比越大说明个体内各基因间存在高度遗传分化<sup>[17]</sup>。

#### 4.3 裸花紫珠 7 种有效成分含量与 SSR 标记的关联分析

目前中药育种以常规育种为主，育种效率低，且裸花紫珠 7 种有效成分含量为数量性状，受多基因控制，性状的最终表达由基因和环境互相影响。因此通过连锁不平衡分析及关联分析等遗传手段筛选与之显著关联的分子标记，进行分子辅助育种，可加快育种进程。本研究结果表明，在 GLM 模型中，发现有 11 对标记与有效成分含量呈显著相关 ( $P < 0.05$ ) 或极显著相关 ( $P < 0.01$ )。在 MLM 模型中，6 对标记与有效成分含量呈显著相关 ( $P < 0.05$ ) 或极显著相关 ( $P < 0.01$ )。MLM 模型较 GLM 得到的关联数据少，且大部分对应的解释率有所下降，说明与 GLM 模型相比，MLM 因同时考虑种群



结构 (Q) 及亲缘关系 (K), 所以能避免一部分假阳性关联, 更能保证关联分析结果的准确性。这与陈旭等<sup>[18]</sup>的研究结果一致, Zhang 等<sup>[19]</sup>也认为 MLM 比 GLM 更适合用于关联分析。在 GLM 和 MLM 2 种模型下, 均检测到 CnH305 与连翘酯苷 B 和木犀草苷相关联 ( $P < 0.05$ ), CnH155 与木犀草苷和毛蕊花糖苷相关联 ( $P < 0.01$ ), 说明标记位点可能存在“一因多效”, 或者该位点可能与多个功能基因存在连锁关系。同时, 2 种模型下, 均检测到 CnH156、CnH357 和 CnH205 与异毛蕊花糖苷相关联, 说明异毛蕊花糖苷含量受多个基因控制。

**利益冲突** 所有作者均声明不存在利益冲突

#### 参考文献

- [1] 刘明生. 黎药学概论 [M]. 北京: 人民卫生出版社, 2008: 263.
- [2] 冯世秀, 张旻, 易博, 等. 裸花紫珠化学成分与药理活性研究进展 [J]. 中草药, 2017, 48(5): 1015-1026.
- [3] 黄梅, 陈振夏, 于福来, 等. 海南岛裸花紫珠种质资源调查报告 [J]. 中国现代中药, 2017, 19(12): 1717-1721.
- [4] 陈于敏, 世荣, 刘吉新, 等. 基于 SSR 标记的高原粳稻品种遗传多样性分析 [J]. 分子植物育种, 2019, 17(12): 4012-4111.
- [5] 谭湘杰, 于福来, 庞玉新, 等. 裸花紫珠药材质量控制研究进展 [J]. 广东药科大学学报, 2017, 33(6): 803-810.
- [6] 吴怡超, 李晓瑜, 何星星, 等. 苦莽种质资源芦丁含量与 SSR 标记关联分析 [J]. 西南大学学报: 自然科学版, 2023, 45(9): 1-11.
- [7] 朱翔宇, 潘多锋, 王建丽, 等. 红三叶抗寒相关生理生化指标与 SSR 关联分析 [J]. 分子植物育种, 2023, 21(6): 56.
- [8] 于福来, 黄梅, 张影波, 等. 裸花紫珠基因组调研及 SSR 特征分析 [J]. 中国中药杂志, 2019, 44(18): 3974-3978.
- [9] 张红瑞, 李鑫, 陈振夏, 等. 基于 SSR 分子标记的裸花紫珠种质资源遗传多样性分析及 DNA 指纹图谱构建 [J]. 中草药, 2023, 54(12): 3971-3982.
- [10] 彭云露. 黎药裸花紫珠资源品质差异及转录组学研究 [D]. 海口: 海南医学院, 2021.
- [11] 谭湘杰, 于福来, 庞玉新, 等. 海南裸花紫珠药材中 7 种成分含量分析 [J]. 中国现代中药, 2018, 20(2): 173-178.
- [12] Pritchard J K, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data [J]. *Genetics*, 2000, 155(2): 945-959.
- [13] 孟亚雄, 孟祎林, 汪军成, 等. 青稞遗传多样性及其农艺性状与 SSR 标记的关联分析 [J]. 作物学报, 2016, 42(2): 180-189.
- [14] 于福来, 吴丽芬, 庞玉新, 等. 海南裸花紫珠中毛蕊花糖苷和木犀草苷含量分析 [J]. 中国现代中药, 2016, 18(8): 996-1000.
- [15] Liu J Y, Zhang Y W, Han X, *et al.* An evolutionary population structure model reveals pleiotropic effects of GmPDAT for traits related to seed size and oil content in soybean [J]. *J Exp Bot*, 2020, 71(22): 6988-7002.
- [16] 刘玉玲, 张红岩, 滕长才, 等. 蚕豆 SSR 标记遗传多样性及与淀粉含量的关联分析 [J]. 作物学报, 2022, 48(11): 2786-2796.
- [17] 刘珊廷, 易显荣, 周民武, 等. 广西梨种质资源遗传多样性和群体结构分析 [J]. 果树学报, 2024, 41(3): 379-391.
- [18] 陈旭, 郝雅琼, 聂兴华, 等. 板栗总苞和坚果主要性状与 SSR 标记的关联分析 [J]. 中国农业科学, 2022, 55(13): 2613-2628.
- [19] Zhang Z W, Ersoz E, Lai C Q, *et al.* Mixed linear model approach adapted for genome-wide association studies [J]. *Nat Genet*, 2010, 42(4): 355-360.

[责任编辑 时圣明]