



基于RAD高通量测序的蓝花楹群体遗传分析

刘学锋, 李小梅, 张国武, 张沛健, 刘果, 黄敏, 陈银霞

引用本文:

刘学锋, 李小梅, 张国武, 张沛健, 刘果, 黄敏, 陈银霞. 基于RAD高通量测序的蓝花楹群体遗传分析[J]. 热带亚热带植物学报, 2022, 30(5): 613–622.

在线阅读 View online: <https://doi.org/10.11926/jtsb.4517>

您可能感兴趣的其他文章

Articles you may be interested in

基于RAD-SNPs分析的四川核桃良种资源的遗传多样性研究

Studies on Genetic Diversity of Juglans Cultivar Germplasm in Sichuan Based on RAD-SNPs Analysis
热带亚热带植物学报. 2019, 27(1): 19–28 <https://doi.org/10.11926/jtsb.3906>

基于SSR标记的广东栽培益智群体遗传多样性分析

Genetic Diversity of Cultivated Populations of *Alpinia oxyphylla* in Guangdong Based on SSR Marker
热带亚热带植物学报. 2022, 30(3): 329–335 <https://doi.org/10.11926/jtsb.4433>

极小种群野生植物海南风吹楠的遗传多样性研究

Genetic Diversity of *Horsfieldia hainanensis*: An Endangered Species with Extremely Small Populations
热带亚热带植物学报. 2021, 29(5): 547–555 <https://doi.org/10.11926/jtsb.4364>

广东省猴耳环遗传多样性研究

Genetic Diversity of *Archidendron clypearia* in Guangdong Province
热带亚热带植物学报. 2021, 29(5): 539–546 <https://doi.org/10.11926/jtsb.4337>

海南岛油茶种质资源遗传多样性的SRAP分析

Genetic Diversity Analysis of *Camellia oleifera* Resources Based on SRAP Markers in Hainan Island
热带亚热带植物学报. 2019, 27(6): 659–668 <https://doi.org/10.11926/jtsb.4027>

向下翻页, 浏览PDF全文

基于 RAD 高通量测序的蓝花楹群体遗传分析

刘学锋¹, 李小梅², 张国武^{1*}, 张沛健¹, 刘果¹, 黄敏¹, 陈银霞³

(1. 中国林业科学研究院速生树木研究所, 广东 湛江 524022; 2. 湛江科技学院, 广东 湛江 524000; 3. 江西省林业科技推广和宣传教育中心, 南昌 330033)

摘要: 为了解蓝花楹(*Jacaranda mimosifolia*)种质资源的遗传多样性和群体遗传结构, 对 168 份种质材料进行 RAD-seq 测序, 构建了系统进化树并进行主成分、群体结构和遗传多样性分析。结果表明, 比对参考基因组平均比率为 81.02%, 平均测序深度 23.18 \times , 最终获得 45 552 个高质量的 SNPs。群体遗传结构分析表明, 供试蓝花楹可划分为 2 个大的类群, 来自川、渝地区的种质材料基本归为一类; 其余地区归为另一类。19 个地区的蓝花楹在 SNP 水平上的遗传多样性较高, 云南昆明(YNKM)居群的核苷酸多样性(π)和期望杂合度(He)最大, 表现出最高的遗传多样性。因此, 来自川、渝地区的蓝花楹具有相对较近的亲缘关系, 推断来自同一祖先, 而其余地区的种质可能是随机引种栽培。

关键词: 蓝花楹; RAD-seq; 遗传多样性; 群体遗传

doi: 10.11926/jtsb.4517

Population Genetic Analysis of *Jacaranda mimosifolia* by RAD High Throughput Sequencing Technique

LIU Xuefeng¹, LI Xiaomei², ZHANG Guowu^{1*}, ZHANG Peijian¹, LIU Guo¹, HUANG Min¹, CHEN Yinxia³

(1. China Eucalypt Research Centre, Zhanjiang 524022, Guangdong, China; 2. Zhanjiang University of Science and Technology, Zhanjiang 524000, Guangdong, China; 3. Jiangxi Forestry Sci-tech Promotion and Propaganda Education Center, Nanchang 330033, China)

Abstract: In order to understand the genetic diversity of *Jacaranda mimosifolia* germplasm resources, 168 germplasms of *Jacaranda mimosifolia* were restriction-site associated DNA sequencing (RAD-seq), and then the phylogenetic tree was constructed, principal component (PCA), population structure and genetic diversity were analyzed. The results showed that the mean alignment rates to the reference genomes were 81.02% with an average sequencing depth of 23.18 \times . After cleaned and filtered, a total of 45 552 SNPs were obtained. All tested germplasms could be clustered into two groups, those from Sichuan-Chongqing region were in one group, and the rest were in another group. *Jacaranda mimosifolia* has a high genetic diversity at the SNP level. Among them, nucleotide diversity (π) and expected heterozygosity (He) in YNKM population is the largest, showing the highest genetic diversity. Therefore, *Jacaranda mimosifolia* from Sichuan-Chongqing region showed relatively close genetic relationship, suggesting that they might from the same ancestor, and those from other regions might be randomly introduced and cultivated.

Key words: *Jacaranda mimosifolia*; RAD-seq; Genetic diversity; Population genetic

蓝花楹(*Jacaranda mimosifolia*)是紫葳科(Bigno-
niaceae)蓝花楹属植物, 落叶乔木, 原产南美洲, 广

泛分布于澳大利亚、南非、秘鲁、墨西哥、巴西、
阿根廷等国, 我国引种栽培于广东、广西、四川、

收稿日期: 2021-09-02

接受日期: 2021-12-13

基金项目: 中央级公益科研院所专项资金项目(CAFYBB2019MB004); 广东省林业科技创新项目(2018KJCX024)资助

This work was supported by the Project for Central Public Research Institutes (Grant No. CAFYBB2019MB004), and the Project for Forestry Science and Technology Innovation in Guangdong (Grant No. 2018KJCX024).

作者简介: 刘学锋(1985 生), 工程师, 从事森林培育及园林植物学研究。Email: cercliuxf@caf.ac.cn

* 通信作者 Corresponding author. E-mail: fzygwu@163.com

福建和云南等地^[1]。蓝花楹树形优美, 蓝紫色花, 可作行道树, 具有观花、观叶、观形等观赏价值, 是一种值得推广应用的园林树种^[2-3]。

对蓝花楹的研究主要集中在潜在栽培区预测、与木霉真菌的相互作用^[4]、药用成分分析^[5]和土壤重金属及化合物吸附^[6-7]等方面。迄今为止, 蓝花楹在分子水平上的研究仅限于染色体^[8-9]、质体序列^[10]、ISSR 分子标记^[11-12]等。因蓝花楹早期引种栽培的品种来源未知, 种源关系混乱, 严重影响了蓝花楹的种质资源保护、良种选育及创新利用, 因此亟待新的方法来揭示国内蓝花楹的遗传多样性和群体遗传结构。近年来, 随着基因组测序技术的飞速发展, 为种质资源鉴定及遗传多样性研究提供了新的方法。限制性酶切位点关联 DNA 测序技术 (restriction-site-associated DNA sequencing, RAD-seq) 是在二代测序基础上发展起来的一种基于全基因组酶切位点的简化基因组测序技术。随着高通量测序技术和生物信息技术的发展, RAD-seq 分析也日趋完善, 并在很多生物上得以应用^[13-15]。RAD-seq 不仅可以得到可靠的高密度遗传标记, 而且适用于没有参考基因组的物种研究^[16]。Premarathne 等^[17]利用 RAD-seq 技术对来自不同地区的 93 份日本胡椒 (*Zanthoxylum piperitum*) 种质资源进行了分析。Feng

等^[18]利用 RAD-seq 技术对甘薯 (*Ipomoea batatas*) 的全基因组遗传多样性进行检测并分析群体结构, 开发了相应的 SSR 标记。Fukuda 等^[19]利用 RAD-seq 技术生成的 SNP 标记构建了枇杷 (*Eriobotrya deflexa*) 高密度遗传连锁图谱, 可用于农艺性状的 QTLs 研究。黄承玲等^[20]基于 RAD-seq 测序技术对贵州百里杜鹃保护区杜鹃花属 (*Rhododendron*) 植物进行了分类, 证实 RAD-seq 技术在复杂植物类群的物种分类方面比传统分子标记具有明显优势。

群体遗传学是研究植物群体遗传结构及其变化规律的遗传学重要学科。因此, 本研究以在国内收集的 168 份蓝花楹种质为材料, 利用 RAD-seq 技术进行蓝花楹的群体遗传分析, 从基因组水平上揭示不同种质间的遗传分化关系, 为蓝花楹种质资源的保存利用、育种策略的实施、分子标记辅助育种和关联图谱的构建提供理论依据和实践参考。

1 材料和方法

1.1 试验材料

分别从四川、重庆、福建、江西、云南、广东、广西 7 省(区、直辖市)的 19 个地区早期引种的蓝花楹 (*Jacaranda mimosifolia*) 树上采集种子(表 1), 每个自

表 1 蓝花楹试验材料及其来源

Table 1 Origin of test varieties of *Jacaranda mimosifolia*

编号 No.	产地来源 Origin	简称 Code	纬度 Latitude (N)	经度 Longitude (E)	海拔 Altitude (m)
1	四川成都 Chengdu, Sichuan	SCCD	30.62'	104.07'	493
2	四川内江 Neijiang, Sichuan	SCNJ	29.58'	105.06'	329
3	重庆北碚 Beibei, Chongqing	CQBB	29.48'	106.23'	248
4	四川宜宾 Yibin, Sichuan	SCYB	28.45'	104.38'	318
5	四川西昌 Xichang, Sichuan	SCXC	27.89'	102.25'	1 513
6	四川攀枝花 Panzhihua, Sichuan	SCPZH	26.57'	101.69'	1 144
7	福建福州 Fuzhou, Fujian	FJFZ	26.08'	119.30'	15
8	江西赣州 Ganzhou, Jiangxi	JXGZ	25.85'	114.93'	115
9	云南大理 Dali, Yunnan	YN DL	25.61'	100.27'	2 007
10	福建莆田(1) Putian, Fujian (1)	FJPT1	25.38'	119.09'	18
11	福建莆田(2) Putian, Fujian (2)	FJPT2	25.38'	119.09'	18
12	云南昆明 Kunming, Yunnan	YN KM	25.08'	102.70'	1 908
13	广西融水 Rongshui, Guangxi	GXRS	25.07'	109.26'	99
14	福建泉州 Quanzhou, Fujian	FJQZ	24.91'	118.59'	11
15	福建厦门 Xiamen, Fujian	FJXM	24.48'	118.09'	0
16	云南开远 Kaiyuan, Yunnan	YNKY	23.71'	103.16'	1 55
17	云南文山(1) Wenshan, Yunnan (1)	YNWS1	23.38'	104.24'	1 265
18	云南文山(2) Wenshan, Yunnan (2)	YNWS2	23.38'	104.24'	1 265
19	云南蒙自 Mengzi, Yunnan	YNMZ	23.36'	103.40'	1303
20	广东广州 Guangzhou, Guangdong	GDGZ	23.13'	113.35'	14
21	广西南宁 Nanning, Guangxi	GXNN	22.82'	108.40'	116

然居群选择 9~16 棵植株, 相邻植株间隔约 50 m, 个体在 10 株以下的居群全部采样。福建莆田(1)和云南文山(1)居群为大树, 福建莆田(2)和云南文山(2)居群为小苗, 且相隔 50 km 以上。种子在位于广东湛江的南方国家级林木种苗示范基地分产地播种培育成 1 a 生小苗, 于 2018 年 6 月从每个产地小苗中随机选取 8 株, 剪取其新鲜健康的幼叶适量, 分别装入 5 mL 离心管中, 再放入液氮速冻后带回, 编号并保存在-80 °C 冰箱备用。

1.2 基因组 DNA 提取

采用 CTAB 法提取样本基因组 DNA 并进行质量检测, 对质量合格的 DNA 进行下一步上机测序。

1.3 RAD 文库构建

首先应用限制性内切酶 *EcoR* I 对基因组进行酶切, 然后每个样本分别进行物理打碎, 选取 200~400 bp 插入片段文库。利用 Illumina HiSeq2000 测序平台进行双末端(Paired End-150)测序, 所有的文库构建及上机测序都由北京诺禾致源科技股份有限公司完成。

1.4 测序 reads 过滤

对原始数据进行过滤, 首先去除接头(adapter)的 reads pair; 当单端测序 read 中含有的 N 含量超过该条 read 长度的 10%时, 需要过滤 paired reads; 当单端测序 read 中含有的低质量($Q \leq 5$)碱基数超过该条 read 长度的 50%时, 去除此对 paired reads^[21]。

1.5 比对和变异检测

将组装样本带有酶切识别序列的 reads 进行聚类, 并按照深度进行降序排序。将深度高的 reads 作为种子进行聚类。根据深度信息对聚类后的 reads 进行纠错、过滤重复区域等。根据聚类的结果, 将一端的 reads 进行 contig 拼接, 结合 insert 的大小和 overlap 关系, 将拼接后的 contig 与聚集在另一端的 reads 连接起来, 组装成最终的 contig 序列。使用 Velvet Optimiser 软件对个体 RAD 局部组装, 得到最后的组装序列, 并对 100 bp 以下的 contig 进行过滤得到基因组^[22]。基因组大小为 82 851 747 bp, 群体样本比率为 78.21%~83.46%, 基因组平均测序深度为 23.18 \times , 平均覆盖度为 92.64%。有效的高质量测序数据通过 BWA 软件^[23](参数: mem-t4-k32-M)比对到蓝花楹组

装后的参考基因组, 并过滤掉无法匹配和匹配到多处位点的序列; 比对结果经 SAM-TOOLS^[24]去除重复(参数: rmdup)并归类^[25]用于变异检测。经 HWB (Hardy-Weinberg equilibrium)过滤后进行 HWB 检验, $P < 0.05$ 为遗传不平衡, 需要过滤掉, 软件: 无, 脚本执行。再进行 LD (linkage disequilibrium)过滤, 连锁强度 $R^2 > 0.8$, 软件: plink1.9-allow-extra-chr-indep-pairwise 50 5 0.8。

1.6 数据的统计分析

群体遗传结构分析 经检测和过滤后的高质量 SNP 可用于计算种群间的距离。运用 Treebest-1.9.2 软件计算距离矩阵, 然后通过邻接法(neighbor-joining method)构建系统进化树判断样本间的亲缘关系, 使用 Bootstrap 为置信度检验统计方法, 重复抽样次数设为 1 000 次。运用 GCTA 软件进行群体主成分分析。利用 ADMIXTURE 软件进行群体遗传结构分析^[26]。群体进化树分析、主成分分析和群体遗传结构分析是群体遗传学的常用分析方法。其中系统进化树是描述群体间分化顺序的分支图或树, 可用来揭示群体间的进化关系, 并根据群体遗传特征推断它们的亲缘关系^[27]。主成分分析是一种研究群体遗传结构的多元统计方法。它主要基于个体基因组中 SNP 差异的程度, 根据不同的性状将个体聚类为不同的亚群, 是确定物种亚群数量的有效方法^[28]。为了进一步明晰 168 份种质间的遗传关系, 对其遗传结构进行解析。采用 PLINK 软件进行群体结构分析。首先创建 PLINK 的输入文件-Ped 文件, 然后利用 ADMIXTURE 软件构建群体遗传结构和群体世系信息。

群体遗传多样性 类群的遗传多样性包括核苷酸多样性、观测杂合度(H_o)、期望杂合度(H_e)和群体间遗传分化指数(F_{st})。运用 vcftools V 0.1.14 软件计算单位点核苷酸多样性 $[P_i(\pi)]$ ^[29-30]。

2 结果和分析

2.1 测序基本数据分析

对 168 份蓝花楹种质进行了 RAD 简化基因组测序, 共获得测序数据量 355.51 Gb, 平均 2.12 Gb。经过对测序数据的严格过滤, 得到高质量的 clean data 为 353.28 Gb, 平均 2.10 Gb, 平均 GC 含量为 35.14%, 碱基质量 Q30 达到 92.10%, 表明测序质

量较高(表 2)。每份种质资源平均获得高质量序列 13 920 867 条, 平均有 11 290 873 的序列条数可以比对到参考基因组, 比对率为 81.02%, 平均测序深度为 23.18 \times 。平均比对率和平均测序深度均满足分析要求, 可以进行后续分析。

进行群体 SNP 的检测, 共获得 573 704 个原始

SNP。通过 Q20 质量控制、SNP 的支持数(覆盖深度)在 4 以上、miss 小于 20%、maf 大于 5%过滤和筛选, 得到 185 054 个 SNPs; 经 HWB 过滤, 共保留 60 285 个 SNP; 经 LD 过滤, 共保留 45 552 个 SNP, 后续分析均基于 HWB 和 LD 过滤后高质量 SNP 进行分析。

表 2 文库构建测序信息

Table 2 Information of sequencing library

地区 Area	原始数据 Raw data (Gb)	有效数据 Clean data (Gb)	Q20 /%	Q30 /%	GC /%	有效片段 Clean reads	可比对片段 Alignment reads	比对效率 % Alignment rate
SCCD	2.53	2.51	96.67	91.23	35.26	16 355 662	13 423 057	81.91
SCNJ	2.22	2.21	96.58	91.06	35.18	13 670 770	11 143 421	81.73
CQBB	2.22	2.21	96.47	90.92	35.00	14 704 911	12 055 841	82.02
SCYB	1.84	1.83	97.23	92.54	34.95	12 181 733	9 762 404	80.05
SCXC	1.96	1.95	97.29	92.65	34.89	13 011 060	10 436 806	80.20
SCPZH	2.07	2.05	97.14	92.40	35.29	13 662 805	10 840 236	79.33
FJFZ	2.31	2.31	96.82	91.65	35.13	15 365 621	12 302 930	80.09
JXGZ	1.87	1.86	96.50	91.16	35.21	12 389 627	9 917 918	80.08
YNDL	1.88	1.87	96.69	91.57	35.00	12 449 208	9 993 992	80.21
FJPT1	1.97	1.96	97.17	92.37	34.83	13 052 439	10 529 356	80.73
FJPT2	1.89	1.88	97.23	92.50	35.06	12 513 897	10 201 631	81.56
YNKM	1.85	1.83	97.15	92.34	35.18	12 183 692	9 813 166	80.54
GXRS	1.99	1.96	97.26	92.70	35.15	13 103 354	10 642 116	81.37
FJQZ	1.92	1.90	97.22	92.64	35.04	12 668 537	10 346 675	81.67
FJXM	1.93	1.92	97.15	92.50	35.13	12 796 933	10 472 656	81.80
YNKY	2.27	2.26	97.48	93.00	35.55	15 048 761	12 369 318	82.03
YNWS1	2.54	2.52	96.82	91.71	35.28	16 594 812	13 565 875	81.46
YNWS2	1.95	1.94	96.38	90.83	35.15	12 711 106	10 270 057	80.71
YNMZ	2.21	2.20	97.58	93.08	35.14	14 688 920	11 921 052	81.18
GDGZ	2.47	2.45	97.41	92.80	35.26	16 352 840	13 271 906	81.08
GXNN	2.56	2.55	97.18	92.42	35.37	16 831 527	13 827 913	81.71

2.2 群体进化树和主成分分析

利用鉴定到的高质量 SNP 对这 168 份蓝花楹种质资源构建进化树(图 1), 可见, 供试蓝花楹可分为 2 大类群, 其中 SCNJ、CQBB、SCYB、SCXC 和 SCPZH 共 5 个地区的种质材料归为类群 I。其余 14 个地区的种质材料归为类群 II, 包括 SCCD、YNKM、FJQZ、YNDL、FJPT1、FJPT2、JXGZ、YNKY、GDGZ、GXRS、FJXM、YNMZ、YNWS1、YNWS2、GXNN 和 FJFZ。从 SNP 中提取关键信息, 采用 GCTA 软件对 168 份蓝花楹材料进行 PCA 主成分分析, 图 2 为基于 3 个主成分的 PCA 聚类图, 可以清楚地反映个体的聚类情况, 群体之间的距离能够反映亲缘关系的远近。PCA 分析中 PCA1、PCA2 和 PCA3 的解释度分别为: 0.182 476 058 936 713、0.046 452 359 970 902 和 0.040 372 436 296 260 5。从 3 个主成分 PC1、PC2

和 PC3 可知: 来自 SCNJ、SCYB、SCXC、SCPZH、SCCD、CQBB 等 6 个地区的种质材料被归为同一类群; 其余 13 个地区的种质材料归为另一类群, 与群体进化聚类分析结果基本一致。

2.3 群体遗传结构分析

为了确定合适的分群数(K 值), 用 ADMIXTURE 软件分析样品的群体结构, 预先假定 K 值为 2~15, 分别进行运算, 并使用交叉验证确定 K 值, 结果表明当 K=2 时为最优分群, 即分为 2 大类群。结果表明(图 3), 群体 I 为橙色, 种质材料分别来自 SCCD、SCNJ、CQBB、SCYB、SCXC 和 SCPZH 等 6 个地区; 群体 II 为蓝色, 种质材料来自其余 13 个地区。这与群体进化分析、主成分分析结果基本一致。为了了解蓝花楹的遗传差异, 计算了蓝花楹群体间的遗

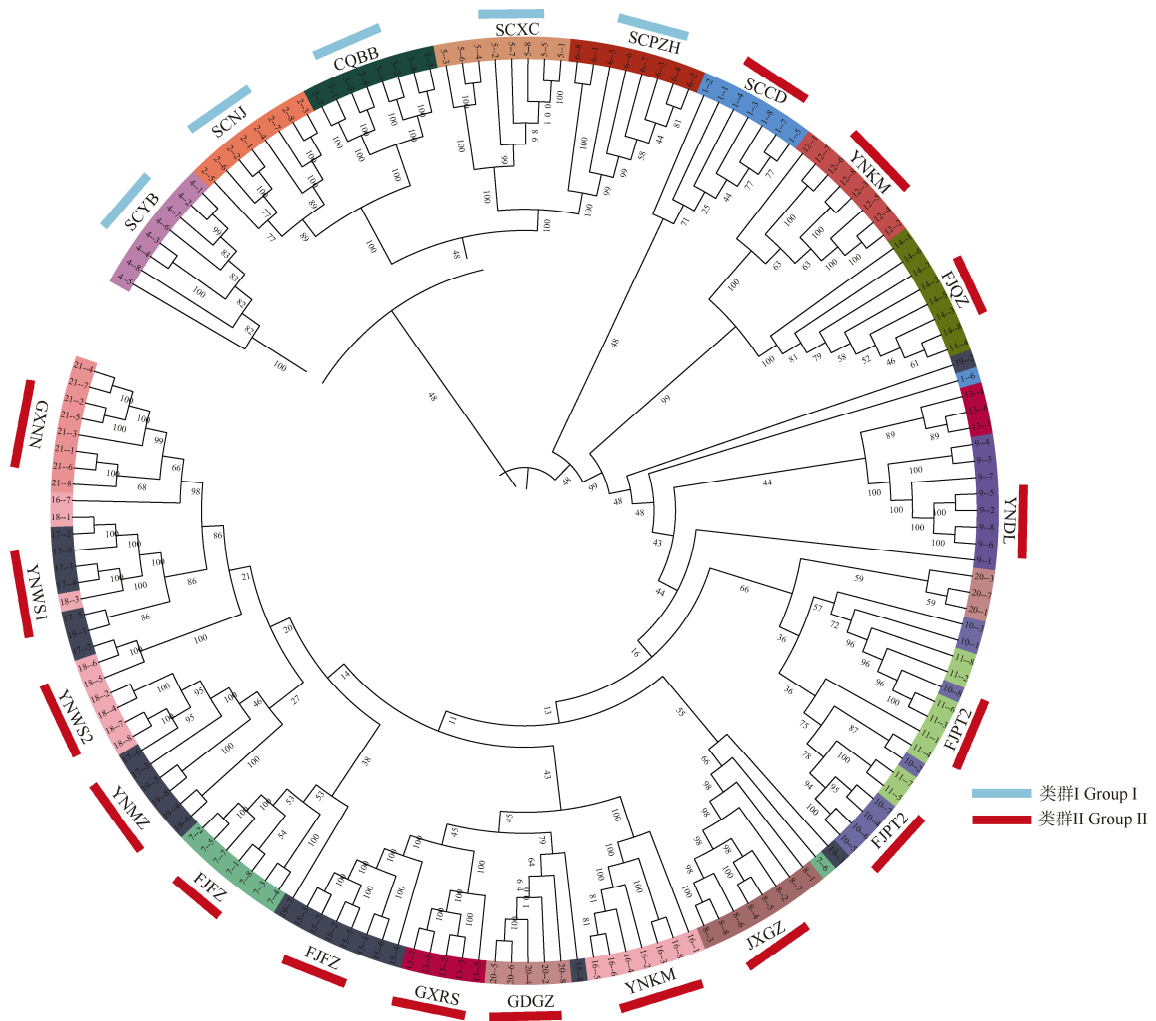


图 1 168 份蓝花楹资源的群体系统进化树

Fig. 1 The phylogenetic tree of the 168 *Jacaranda mimosifolia* samples

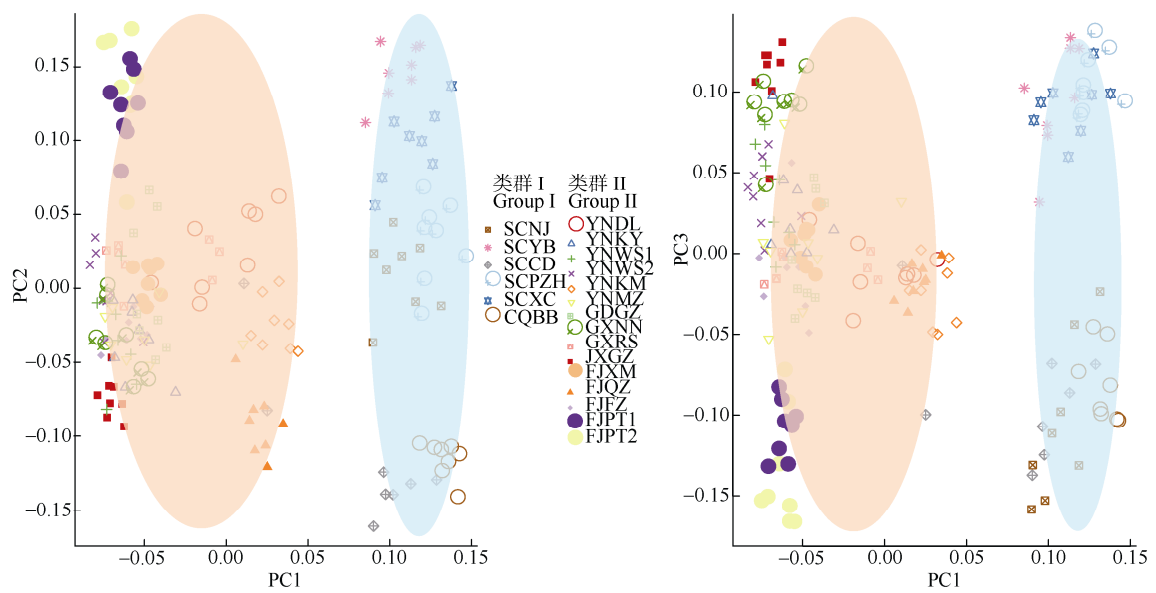


图 2 168 份蓝花楹种质资源的群体主成分分析

Fig. 2 Principal component analysis of the 168 germplasms of *Jacaranda mimosifolia*

传分化系数(F_{st}) (表 3)。可见, 21 个蓝花楹群体间的 F_{st} 为 0.016~0.178, 其中 CQBB 与 FJPT1、CQBB 与 FJPT2、CQBB 与 JXGZ 的 F_{st} 分别为 0.178、0.173 和 0.174; SCPZH 与 JXGZ 和 SCPZH 与 FJPT1 分

别为 0.175 和 0.173, F_{st} 均大于 0.15, 说明遗传分化较大。FJPT1 与 FJPT2、YNWS1 与 YNWS2 的 $F_{st} < 0.05$, 说明遗传分化很小, 采样时分别对大树和小苗取样的做法无意义。

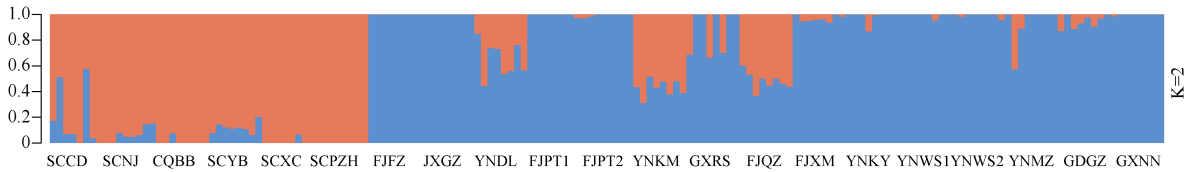


图 3 168 份蓝花楹种质资源的遗传结构关系

Fig. 3 Genetic structure of 168 *Jacaranda mimosifolia* germplasms

表 3 蓝花楹群体间遗传分化系数(F_{st})

Table 3 Genetic differentiation coefficient (F_{st}) among *Jacaranda mimosifolia* populations

编号 No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	0.090																			
3	0.078	0.069																		
4	0.111	0.095	0.105																	
5	0.096	0.087	0.095	0.081																
6	0.085	0.094	0.090	0.086	0.059															
7	0.113	0.141	0.153	0.151	0.144	0.154														
8	0.140	0.165	0.174	0.162	0.169	0.175	0.070													
9	0.096	0.109	0.119	0.104	0.111	0.114	0.062	0.084												
10	0.134	0.151	0.178	0.156	0.155	0.173	0.066	0.093	0.080											
11	0.130	0.149	0.173	0.150	0.149	0.165	0.061	0.087	0.074	0.016										
12	0.088	0.089	0.101	0.104	0.105	0.100	0.092	0.109	0.050	0.108	0.102									
13	0.128	0.144	0.154	0.140	0.147	0.150	0.075	0.090	0.067	0.090	0.083	0.097								
14	0.096	0.103	0.105	0.128	0.107	0.111	0.090	0.104	0.078	0.115	0.110	0.065	0.113							
15	0.113	0.138	0.158	0.142	0.144	0.143	0.066	0.082	0.078	0.075	0.074	0.097	0.058	0.101						
16	0.111	0.145	0.152	0.150	0.139	0.146	0.045	0.070	0.075	0.072	0.069	0.097	0.070	0.083	0.057					
17	0.127	0.149	0.165	0.147	0.158	0.162	0.053	0.065	0.072	0.066	0.060	0.098	0.073	0.105	0.057	0.049				
18	0.120	0.143	0.155	0.145	0.152	0.155	0.049	0.059	0.070	0.069	0.066	0.096	0.068	0.096	0.057	0.045	0.025			
19	0.102	0.124	0.141	0.136	0.134	0.141	0.039	0.060	0.061	0.064	0.062	0.078	0.070	0.082	0.053	0.042	0.045	0.039		
20	0.114	0.127	0.143	0.140	0.137	0.142	0.058	0.071	0.069	0.066	0.059	0.088	0.071	0.089	0.058	0.058	0.058	0.057	0.045	
21	0.134	0.156	0.169	0.159	0.160	0.163	0.065	0.069	0.089	0.087	0.080	0.106	0.102	0.106	0.072	0.059	0.046	0.055	0.059	0.070

1~21 见表 1。

1-21 see Table 1.

2.4 群体遗传多样性分析

为了解蓝花楹群体遗传多样性水平, 进行了核苷酸多样性和杂合度分析, 核苷酸多样性是指样本中所有可能匹配成对的序列间核苷酸位点差异百分比的平均值, 用 π 表示。杂合度反映了群体中的遗传变异程度, 杂合度越高, 表明群体内遗传多样性越高, 包括观测杂合度(H_o)和期望杂合度(H_e)。从表 4 可知, 居群水平上, 蓝花楹的 π 为 0.293~0.347, 平均为 0.320, H_o 为 0.406~0.474, 平均为 0.433, H_e

为 0.273~0.323, 平均为 0.298。其中, YNKM 居群的 π 和 H_e 最大, 分别为 0.347 和 0.323, 表现出最高的遗传多样性; 其次为 YNDL, π 、 H_o 和 H_e 分别为 0.345、0.474 和 0.321; 遗传多样性最小的是 GXRS, π 、 H_o 和 H_e 分别为 0.293、0.406 和 0.273。

3 结论和讨论

蓝花楹为国外引进的具有高景观价值的观花

表 4 21 个蓝花楹群体的遗传多样性参数

Table 4 Genetic diversity parameters of 21 *Jacaranda mimosifolia* populations

地区 Area	核苷酸多样性 (π) Nucleotide diversity	观测杂合度 (H_o) Observed heterozygosity	期望杂合度 (H_e) Expected heterozygosity
SCCD	0.331	0.428	0.309
SCNJ	0.346	0.438	0.322
CQBB	0.325	0.438	0.303
SCYB	0.319	0.429	0.297
SCXC	0.323	0.443	0.301
SCPZH	0.326	0.461	0.304
FJFZ	0.317	0.440	0.296
JXGZ	0.301	0.425	0.280
YNDL	0.345	0.474	0.321
FJPT1	0.299	0.418	0.279
FJPT2	0.303	0.428	0.283
YNKM	0.347	0.450	0.323
GXRS	0.293	0.406	0.273
FJQZ	0.315	0.445	0.294
FJXM	0.304	0.420	0.283
YNKY	0.313	0.422	0.292
YNWS1	0.315	0.420	0.294
YNWS2	0.332	0.442	0.309
YNMZ	0.338	0.435	0.315
GDGZ	0.321	0.410	0.300
GXNN	0.301	0.410	0.280
平均 Mean	0.320	0.433	0.298

乔木, 其种质资源利用还处在起步阶段, 进行遗传多样性分析和背景研究是十分必要的。本研究首次通过 RAD-seq 测序技术在全基因组水平上分析蓝花楹群体的遗传结构和遗传多样性, 共获得测序数据 355.51 Gb, 过滤后的高质量 clean data 为 353.28 Gb。测序质量较高, 平均 GC 含量为 35.14%, 碱基质量 Q20 达到 97.02%, Q30 达到 92.10%, 说明建库测序成功。共检测获得原始 SNP 573 704 个, 鉴定出 185 054 个高质量 SNP 标记, 经 HWB 和 LD 过滤保留 45 552 个 SNP。这些标记信息将为蓝花楹的遗传育种、保护研究提供了基础。

对 168 份蓝花楹种质进行群体进化树构建、主成分分析和群体结构分析, 其结果基本一致。群体 I 的种质材料基本来自川、渝的 6 个地区, 包括 SCNJ、SCYB、SCXC、SCPZH、SCCD 和 CQBB, 显示亲缘关系较近, 这些地区的行政权属早期均归属四川省管辖, 引种栽培时有条件共享同一批种质材料, 说明行政权属归属或地理区域近的材料遗传距离比较接近, 这与屈洋等^[31]的研究结果一致。由系统进化树的支持率可知, SCNJ、CQBB、SCYB、SCXC 和 SCPZH 的支持率为 48%, YNDL 和 GXRS 为 44%,

YNKY、GDGZ、GXRS 和 FJXM 为 43%, YNMZ、YNWS1、YNWS2、GXNN 和 FJFZ 为 14%, 这些种质无法通过系统进化树分开, 可能与种源引入我国后是以种子实生苗繁殖, 多年的混乱杂交造成的; 也可能是在来源地已经有充分的基因流, 没有形成显著分化。GXRS 和 YNDL、FJXM, FJFZ 和 JXGZ 以及 GDGZ 和 FJPT 群体间呈现出少数的个体混杂现象, 表明群体间存在基因交流。

遗传多样性分析基于 SNP 位点的核苷酸多样性(π)、观测杂合度(H_o)、期望杂合度(H_e)^[32]。本研究结果表明, 21 个蓝花楹群体的 π 、 H_o 和 H_e 均较高, 意味着蓝花楹在 SNP 水平上的遗传多样性较高, 与海南风吹楠(*Horsfieldia hainanensis*)^[33]的研究结果相似。

F_{st} 是分析群体遗传分化程度的重要指标之一^[30], Wright^[34]提出, 当 F_{st} 为 0~0.05 时群体间遗传分化很小, 可以忽略不计; 当 F_{st} 为 0.05~0.15 时群体间的遗传分化表现中等; 当 F_{st} 为 0.15~0.25 时群体间遗传分化较大; 当 F_{st} 大于 0.25 时, 群体间有很大的遗传分化。本研究中群体间的 F_{st} 为 0.016~0.178, 表明各群体间的遗传分化程度较高。其中, 来自川、

渝地区的群体 I (SCNJ、SCYB、SCXC、SCPZH、SCCD 和 CQBB) 与群体 II (FJFZ、JXGZ、YNDL、FJPT1、FJPT2、YNKM、GXRS、FJQZ、FJXM、YNKY、YNWS1、YNWS2、YNMZ、GDGZ 和 GXNN) 的遗传分化程度较高, 这与群体系统进化树、主成分分析与群体遗传结构分析结果一致。即两大群体间遗传分化比较大, 群体内部居群间遗传分化较小。FJPT1 与 FJPT2 居群和 YNWS1 与 YNWS2 居群间的遗传分化最小, 说明同一地域大树和小苗采样对遗传多样性和遗传分化无差异, 说明该地区蓝花楹种苗无迁移。

尽管蓝花楹最早于 20 世纪 20 年代就引入中国, 主要作为城市景观树种, 但早期保存下来的数量并不多, 90 年代后期蓝花楹在中国得到了较快推广, 遗憾的是, 早期引种的蓝花楹的来源地、具体栽培时间等遗传背景均不清晰, 后期栽培的蓝花楹基本来自国内种子繁殖培育, 或许正是由于这种不清晰的遗传背景, 从而难于寻找出类群 II 中来自其余 13 个地区种质材料彼此间亲缘关系的远近规律。少数几个种质材料在不同的研究中被划分到不同类群或亚群, 呈现出较细微差别, 这可能是计算依据不同而导致的, 这与王小柯等^[35]的研究结果一致。

尽管蓝花楹引种栽培到中国的遗传背景不清, 影响了对本研究部分结果的深入研判, 但本研究通过对蓝花楹的遗传多样性和群体结构分析, 构建起了蓝花楹的遗传图谱, 为蓝花楹育种、系统发育研究、遗传保护、分子标记开发以及产业开发应用等提供了理论基础。同时较系统完整的保存了一批早期引种到中国的蓝花楹遗传材料, 为后续的深入研究奠定了基础。对于蓝花楹种质资源遗传背景的判定, 相比于性状鉴别法和 ISSR 分子标记^[36]来说, RAD-seq 技术更具优势, 更能准确地反映种质资源间的亲缘关系。

参考文献

- [1] Delectis Florae Reipublicae Popularis Sinicae, Agendae Academiae Sinicae Editta. Florae Reipublicae Popularis Sinicae, Tomus 69 [M]. Beijing: Science Press, 1981.
中国植物志编辑委员会. 中国植物志, 第 69 卷 [M]. 北京: 科学出版社, 1981.
- [2] LIU X F, LI X M, ZHANG P J, et al. Research progress of *Jacaranda mimosifolia* and the development prospects [J]. Eucalypt Sci Technol, 2015, 32(4): 51–55. doi: 10.13987/j.cnki.askj.2015.04.012.
- [3] LI Q. Study on distribution and landscape application of *Jacaranda mimosifolia* D. Don in Sichuan region [D]. Chengdu: Sichuan Agricultural University, 2015.
李青. 蓝花楹在四川地区的分布及园林应用研究 [D]. 成都: 四川农业大学, 2015.
- [4] FERREIRA N C D F, ROCHA E C, RODRIGUES F, et al. *Trichoderma* spp. in growth promotion of *Jacaranda mimosifolia* D. Don [J]. J Agric Stud, 2021, 9(2): 335–346. doi: 10.5296/jas.v9i2.18410.
- [5] NAZ R, ROBERTS T H, BANO A, et al. GC-MS analysis, antimicrobial, antioxidant, antilipoxygenase and cytotoxic activities of *Jacaranda mimosifolia* methanol leaf extracts and fractions [J]. PLoS One, 2020, 15(7): e0236319. doi: 10.1371/journal.pone.0236319.
- [6] GEORGIN J, SALOMÓN Y L D O, FRANCO D S P, et al. Development of highly porous activated carbon from *Jacaranda mimosifolia* seed pods for remarkable removal of aqueous-phase ketoprofen [J]. J Environ Chem Eng, 2021, 9(4): 105676. doi: 10.1016/j.jece.2021.105676.
- [7] FARIAS C P, ALVES G S, OLIVEIRA D C, et al. A consortium of fungal isolates and biochar improved the phytoremediation potential of *Jacaranda mimosifolia* D. Don and reduced copper, manganese, and zinc leaching [J]. J Soils Sediments, 2020, 20(1): 260–271. doi: 10.1007/s11368-019-02414-3.
- [8] WANG M C, ZHANG L, WANG Z Q. Chromosomal-level reference genome of the neotropical tree *Jacaranda mimosifolia* D. Don [J]. Genome Biol Evol, 2021, 13(6): evab094. doi: 10.1093/gbe/evab094.
- [9] CORDEIRO J M P, LIMA S A A, PAZ S N, et al. Karyotype evolution in the genus *Jacaranda* Juss. (Jacarandaceae, Bignoniaceae): Chromosome numbers and heterochromatin [J]. Genet Mol Res, 2016, 15(4): gmr15048973. doi: 10.4238/gmr15048973.
- [10] ZHAO K K, CHEN R L, WANG J H, et al. Complete plastome sequence of *Jacaranda mimosifolia* D. Don (Bignoniaceae): A beautiful landscaping tree species [J]. Mitochondr DNA B Resour, 2019, 4(2): 4111–4112. doi: 10.1080/23802359.2019.1692702.
- [11] ESCANDÓN A, DE LA TORRE M P, ACEVEDO A, et al. Anchored ISSR as molecular marker to characterize accessions of *Jacaranda mimosifolia* L. Don [J]. Acta Hort, 2005, 683: 121–127. doi: 10.17660/ActaHortic.2005.683.11.
- [12] LIU X F, LIU G, LI X M, et al. Analysis of genetic diversity and relationship of *Jacaranda mimosifolias* based on ISSR molecular markers technique [J]. Mol Plant Breed, 2021, 19(21): 7146–7153. doi: 10.13271/j.mpb.019.007146.

- 刘学锋, 刘果, 李小梅, 等. 基于 ISSR 分子标记技术的蓝花楹遗传多样性和亲缘关系分析 [J]. 分子植物育种, 2021, 19(21): 7146–7153. doi: 10.13271/j.mpb.019.007146.
- [13] WANG Y K, HU Y, ZHANG T Z. Current status and perspective of RAD-seq in genomic research [J]. Hereditas, 2014, 36(1): 41–49. doi: 10.3724/SP.J.1005.2014.0050.
- 王洋坤, 胡艳, 张天真. RAD-seq 技术在基因组研究中的现状及展望 [J]. 遗传, 2014, 36(1): 41–49. doi: 10.3724/SP.J.1005.2014.0050.
- [14] ZHANG Y, ZHOU W Y, SUN W. The two major technologies of sequencing based on simplified genome by restriction enzyme digestion [J]. Mol Plant Breed, 2020, 18(11): 3562–3570. doi: 10.13271/j.mpb.018.003562.
- 张羽, 周婉莹, 孙旺. 基于限制性内切酶简化基因组测序的两种主要技术 [J]. 分子植物育种, 2020, 18(11): 3562–3570. doi: 10.13271/j.mpb.018.003562.
- [15] XIONG Y, ZHANG J Z, DONG J, et al. A review of reduced-representation genome sequencing technique and its applications in ornamental plants [J]. Acta Hort Sin, 2020, 47(6): 1194–1202. doi: 10.16420/j.issn.0513-353x.2019-0716.
- 熊燕, 张金柱, 董婕, 等. 简化基因组测序技术在观赏植物中的应用研究进展 [J]. 园艺学报, 2020, 47(6): 1194–1202. doi: 10.16420/j.issn.0513-353x.2019-0716.
- [16] ZHAI Z X. SNP mapping and population genetic analyses for 13 Chinese indigenous chicken breeds using red sequencing [D]. Shanghai: Shanghai Jiao Tong University, 2014.
- 翟正晓. 基于 RAD 简化基因组测序技术的 13 种中国地方优良鸡品种 SNPs 多态性图谱构建及群体遗传学分析 [D]. 上海: 上海交通大学, 2014.
- [17] PREMARATHNE M D G P, FUKUTOME N, YAMASAKI K, et al. Elucidation of Japanese pepper (*Zanthoxylum piperitum* De Candolle) domestication using RAD-Seq [J]. Sci Rep, 2021, 11: 6464. doi: 10.1038/s41598-021-85909-9.
- [18] FENG J Y, ZHAO S, LI M, et al. Genome-wide genetic diversity detection and population structure analysis in sweetpotato (*Ipomoea batatas*) using RAD-seq [J]. Genomics, 2020, 112(2): 1978–1987. doi: 10.1016/j.ygeno.2019.11.010.
- [19] FUKUDA S, NAGANO Y, MATSUGUMA K, et al. Construction of a high-density linkage map for bronze loquat using RAD-Seq [J]. Sci Hort, 2019, 251: 59–64. doi: 10.1016/j.scienta.2019.02.065.
- [20] HUANG C L, YAO G, TIAN X L, et al. Phylogenomic analysis of *Rhododendron* species in Guizhou Baili *Rhododendron* Reserve Based on RAD sequencing [J]. Sci Silv Sin, 2021, 57(2): 72–81. doi: 10.11707/j.1001-7488.20210208.
- 黄承玲, 姚刚, 田晓玲, 等. 基于 RAD 高通量测序的贵州百里杜鹃保护区杜鹃花属分类 [J]. 林业科学, 2021, 57(2): 72–81. doi: 10.11707/j.1001-7488.20210208.
- [21] ZHANG S D, YUAN D Y, LU H F, et al. The results of rice germplasm EDV test by genomic analysis and related discussions [J]. Sci Sin Vitae, 2020, 50(6): 633–649. doi: 10.1360/SSV-2020-0068.
- 张上都, 袁定阳, 路洪凤, 等. 基因组学方法用于水稻种质资源实质派生的检测结果和应用讨论 [J]. 中国科学 生命科学, 2020, 50(6): 633–649. doi: 10.1360/SSV-2020-0068.
- [22] ZHANG B D, XUE D X, WANG J, et al. Development and preliminary evaluation of a genomewide single nucleotide polymorphisms resource generated by RAD-seq for the small yellow croaker (*Larimichthys polyactis*) [J]. Mol Ecol Resour, 2016, 16(3): 755–768. doi: 10.1111/1755-0998.12476.
- [23] LI H, DURBIN R. Fast and accurate long-read alignment with Burrows-Wheeler transform [J]. Bioinformatics, 2010, 26(5): 589–595. doi: 10.1093/bioinformatics/btp698.
- [24] LI H, HANDSAKER B, WYSOKER A, et al. The sequence alignment/map format and SAMtools [J]. Bioinformatics, 2009, 25(16): 2078–2079. doi: 10.1093/bioinformatics/btp352.
- [25] GUI L L. Study on phylogeny of *Vitis flexuosa* clade in the genus of *Vitis* distributed in China based on RAD-seq [D]. Wuhan: Huazhong Agricultural University, 2017.
- 桂柳柳. 基于 RAD-seq 技术的中国葡萄属葛藟葡萄支系的系统发育研究 [D]. 武汉: 华中农业大学, 2017.
- [26] ALEXANDER D H, NOVEMBRE J, LANGE K. Fast model-based estimation of ancestry in unrelated individuals [J]. Genome Res, 2009, 19(9): 1655–1664. doi: 10.1101/gr.094052.109.
- [27] WANG X, GAO M, WU L W, et al. A study on population genetic diversity among *Litsea populifolia* (Hemsl.) Gamble in Mount Emei Area [J]. J Plant Genet Resour, 2019, 20(2): 359–369. doi: 10.13430/j.cnki.jpgr.20180812001.
- 王雪, 高暝, 吴立文, 等. 峨眉山地区杨叶木姜子群体遗传多样性研究 [J]. 植物遗传资源学报, 2019, 20(2): 359–369. doi: 10.13430/j.cnki.jpgr.20180812001.
- [28] LI Y F, LI S M, JIN X, et al. Phylogenomic analysis of 85 *Rhododendron* species in China based on RAD sequencing [J]. For Res, 2019, 32(3): 1–8. doi: 10.13275/j.cnki.lykxyj.2019.03.001.
- 李云飞, 李世明, 金鑫, 等. 基于 RAD 高通量测序探讨中国 85 种杜鹃花属植物的分类 [J]. 林业科学研究, 2019, 32(3): 1–8. doi: 10.13275/j.cnki.lykxyj.2019.03.001.
- [29] YAN S Y, ZHU P, GONG W, et al. Studies on genetic diversity of *Juglans* cultivar germplasms in Sichuan based on RAD-SNPs analysis

- [J]. *J Trop Subtrop Bot*, 2019, 27(1): 19–28. doi: 10.11926/jtsb.3906.
- 闫思宇, 朱鹏, 龚伟, 等. 基于 RAD-SNPs 分析的四川核桃良种资源的遗传多样性研究 [J]. *热带亚热带植物学报*, 2019, 27(1): 19–28. doi: 10.11926/jtsb.3906.
- [30] CAO Y Y, DIAO Q N, CHEN Y Y, et al. Analysis of genetic diversity of melon based on 2b-RAD simplified genome sequencing [J]. *Acta Bot Boreali-Occid Sin*, 2021, 41(1): 96–106. doi: 10.7606/j.issn.1000-4025.2021.01.0096.
- 曹燕燕, 刁倩楠, 陈幼源, 等. 基于 2b-RAD 简化基因组测序的甜瓜遗传多样性分析 [J]. *西北植物学报*, 2021, 41(1): 96–106. doi: 10.7606/j.issn.1000-4025.2021.01.0096.
- [31] QU Y, ZHOU Y, WANG Z, et al. Analysis of genetic diversity and structure of tartary buckwheat resources from production regions [J]. *Sci Agric Sin*, 2016, 49(11): 2049–2062. doi: 10.3864/j.issn.0578-1752.2016.11.002.
- 屈洋, 周瑜, 王钊, 等. 苦荞产区种质资源遗传多样性和遗传结构分析 [J]. *中国农业科学*, 2016, 49(11): 2049–2062. doi: 10.3864/j.issn.0578-1752.2016.11.002.
- [32] ZHANG S S, KANG H M, YANG W Z. Population genetic analysis of *Nyssa yunnanensis* by reduced-representation sequencing technique [J]. *Bull Bot Res*, 2019, 39(6): 899–907. doi: 10.7525/j.issn.1673-5102.2019.06.013.
- 张珊珊, 康洪梅, 杨文忠. 基于简化基因组技术的云南蓝果树群体遗传分析 [J]. *植物研究*, 2019, 39(6): 899–907. doi: 10.7525/j.issn.1673-5102.2019.06.013.
- [33] CAI C N, HOU Q X, CI X Q, et al. Genetic diversity of *Horsfieldia hainanensis*: An endangered species with extremely small populations [J]. *J Trop Subtrop Bot*, 2021, 29(5): 547–555. doi: 10.11926/jtsb.4364.
- 蔡超男, 侯勤曦, 慈秀芹, 等. 极小种群野生植物海南风吹楠的遗传多样性研究 [J]. *热带亚热带植物学报*, 2021, 29(5): 547–555. doi: 10.11926/jtsb.4364.
- [34] WRIGHT S. The interpretation of population structure by F-statistics with special regard to systems of mating [J]. *Evolution*, 1965, 19(3): 395–420. doi: 10.2307/2406450.
- [35] WANG X K, JIANG D, SUN Z Z. Study on phylogeny of 240 mandarin accessions with genotyping-by-sequencing technology [J]. *Sci Agric Sin*, 2017, 50(9): 1666–1673. doi: 10.3864/j.issn.0578-1752.2017.09.012.
- 王小柯, 江东, 孙珍珠. 利用 GBS 技术研究 240 份宽皮柑橘的系统演化 [J]. *中国农业科学*, 2017, 50(9): 1666–1673. doi: 10.3864/j.issn.0578-1752.2017.09.012.
- [36] YANG Y L, MA X Q, ZHANG M Q. ISSR molecular marker and its application to the research on genetic breeding of trees [J]. *Subtrop Agric Res*, 2006, 2(1): 18–24. doi: 10.3969/j.issn.1673-0925.2006.01.006.
- 杨玉玲, 马祥庆, 张木清. ISSR 分子标记及其在树木遗传育种研究中的应用 [J]. *亚热带农业研究*, 2006, 2(1): 18–24. doi: 10.3969/j.issn.1673-0925.2006.01.006.