

基于 SNP 标记的辣木群体遗传分析

普天磊¹, 韩学琴¹, 罗会英¹, 邓红山¹, 邹枚伶², 金杰^{1*}, 夏志强^{2*}, 王文泉²

1. 云南省农业科学院热区生态农业研究所/元谋干热河谷植物园, 云南元谋 651300; 2. 海南大学, 海南海口 570228

摘要: 基于 SNP 标记对辣木亲本及其子代群体的杂合度、遗传结构及遗传多样性进行分析, 研究其遗传变异特征。利用 AFSM 技术对 96 份辣木材料进行简化基因组测序, 将获得的测序过滤数据比对至参考基因组, 使用 VCFtools 和 BCFtools 软件检测并统计 SNP 和 Indel 位点信息。利用 AWK 语言分析杂合位点, 并对比出子代与亲本的差异位点, 以分析辣木的繁育类型。利用 Plink 软件对变异位点进行过滤, 保留高质量的变异位点, 再通过 ADMIXTURE 软件进行群体结构分析, 根据交叉验证错误率确定最佳 K 值, 同时采用 GCTA 软件进行主成分分析, 构建系统进化树, 解析辣木材料的群体结构。采用 VCFtools 计算遗传多样性指数及群体分化指数, 分析该群体的遗传多样性。通过 LDBlockShow 软件进行连锁不平衡分析, 得出位点间的连锁不平衡程度。结果显示, 共检测出 1 187 831 个 SNP 位点, 150 861 个 Indel 位点。将辣木子代基因与亲本比对后, 发现辣木子代杂合的基因中, 约有 4.89% 的基因为自身杂合基因, 19.96% 为外来遗传物质导致杂合的基因, 基本表明辣木可通过自花和异花 2 种授粉方式繁衍后代。群体结构和主成分分析将辣木样品划分为 3 个亚群, 与聚类分析的结果大概一致, 即各亚群大致能聚在一起, 且样品间有一定的交叉。辣木不同群体间的遗传分化指数 (0.0049–0.0110) 和遗传多样性指数 (0.001) 低, 表明群体遗传分化弱, 遗传多样性水平低。对检测到的 136 个 scaffold 的 SNP 进行统计并进行连锁不平衡分析后发现, scaffold 1 的 SNP 数目最多, 为 62 225 个, 且其 6 748 044–6 748 185 位点之间具有强连锁不平衡关系。本研究通过对辣木及其子代的遗传分析, 为辣木的杂交选育提供遗传学依据。

关键词: 辣木; SNP; 杂合度; 群体结构; 遗传多样性

中图分类号: S792.99 文献标识码: A

Analysis of Genetic of *Moringa oleifera* Population Based on SNP Markers

PU Tianlei¹, HAN Xueqin¹, LUO Huiying¹, DENG Hongshan¹, ZOU Meiling², JIN Jie^{1*}, XIA Zhiqiang^{2*}, WANG Wenquan²

1. Institute of Tropical Eco-agriculture, Yunnan Academy of Agricultural Sciences / Yuanmou Dry-hot Valley Botanical Garden, Yuanmou, Yunnan 651300, China; 2. Hainan University, Haikou, Hainan 570228, China

Abstract: The heterozygosity, population structure and genetic diversity of *Moringa oleifera* and its progeny were analyzed using the SNP markers. The characteristics of genetic variation of *M. oleifera* and its progeny were studied. Simplified genome sequencing of 96 *M. oleifera* materials was performed by AFSM technology. The obtained sequencing filtering data were compared to the reference genome. SNP and Indel loci were detected and counted by using VCFtools and BCFtools. In order to analyze the reproduction types of *M. oleifera*, AWK language was used to analyze heterozygous loci, and compare the different loci of the progeny with the parents. The mutation sites were filtered by Plink software, and the high-quality mutation sites were reserved. The population structure was analyzed by ADMIXTURE software, and the optimal K value was determined according to the cross validation error rate. At the same time, GCTA

收稿日期 2021-12-09; 修回日期 2022-04-08

基金项目 国家现代农业产业技术体系项目 (No. CARS-11-YNJJ); 云南省重大科技专项“生物资源数字化开发应用”项目 (No. 202002AA100007); 海南大学科研启动基金项目 (No. KYQD(ZR)-20101)。

作者简介 普天磊 (1992—), 女, 硕士, 研究实习员, 研究方向: 功能植物资源与利用。*通信作者 (Corresponding author): 金杰 (JIN Jie), E-mail: 276361917@qq.com; 夏志强 (XIA Zhiqiang), E-mail: xiazhiqiang@itbb.org.cn。

software was used for principal component analysis, and the phylogenetic tree was constructed to analyze the population structure of the materials. VCFtools was used to calculate the genetic diversity index and population differentiation index for analyzing the genetic polymorphism of the population. LDBlockShow software was used for linkage imbalance analysis, and the linkage imbalance degree between loci was obtained. The result showed that a total of 1 187 831 SNP sites and 150 861 Indel sites were detected in this paper. We compared *M. oleifera* progeny genes with parents, which found that among the heterozygous genes in *Moringa* progeny, about 4.89% of the genes are self-heterozygous genes, and 19.96% of the heterozygous genes caused by foreign genetic materia. It indicates that *M. oleifera* produce offspring by the ways of self-pollination and cross-flowering. *M. oleifera* samples were divided into 3 subgroups by population structure and principal component analysis. The cluster analysis is roughly consistent with the above results. The subgroups can be clustered together, and there is a little crossover between the samples. The low genetic differentiation index (0.0049–0.0110) and genetic diversity index (0.001) among different populations of *M. oleifera* indicated that the level of genetic diversity was low and genetic differentiation was weak. The SNPS of 136 scaffolds which detected were counted and linkage imbalance analysis was performed, we found that scaffold 1 had the most SNPS (62 225), there was strong linkage imbalance between 6 748 044 and 6 748 185 loci. This study analyzed the genetic diversity of *M. oleifera* and its progeny. It would provide genetic basis for the cross breeding of *M. oleifera*.

Keywords: *Moringa oleifera* Lam.; SNP; heterozygosity; population structure; genetic diversity

DOI: 10.3969/j.issn.1000-2561.2023.09.006

辣木 (*Moringa oleifera* Lam.) 属于辣木科辣木属的多年生落叶乔木, 又被称为鼓槌树, 辣木原产于印度, 是埃塞俄比亚、尼日利亚、菲律宾和苏丹的重要农作物, 在非洲、美洲等热带亚热带地区均有分布, 在我国的云南、海南、福建、广州等地均有种植^[1]。辣木有 13 个种, 其叶片中约含有 20%~30% 的蛋白质, 叶片、花、果实含有丰富的维生素 A、维生素 B、维生素 C 和钙、镁等矿物质, 种子含高油酸, 可用作化妆品、烹饪和机械润滑油, 种子榨油后的剩余物可用于净化污水、饲喂动物^[2-4]。同时辣木含有丰富的皂苷、生物碱、黄酮、酚类等次生代谢产物, 具有抗氧化、抗炎、细胞保护、神经保护、抗癌等药理作用^[5-6]。

SNP 标记相比较于 RFLP、SSR 等传统分子标记而言, 可检测单个碱基的插入、缺失、转换和颠换, 具有变异数量多, 分布广, 遗传稳定性高, 检测快、通量高的优点^[7]。基于 SNP 标记进行的遗传分析在植物学领域应用较多, 例如, 高嵩等^[8]利用 SNP 芯片进行玉米遗传多样性、群体遗传结构和类群间遗传关系分析, 选育并审定了玉米新品种; 韩志刚等^[9]基于 SNP 标记对 148 份马铃薯种质遗传多样性进行分析, 认为马铃薯绝大部分栽培种遗传相似性高, 遗传背景不够丰富。目前, 国内并没有利用 SNP 分子标记对辣木种群的遗传学进行分析的报道。AFSM 技术为简化基因组测序技术, 该法分别利用 *EcoR* I-*Msp* I 和 *EcoR* I-*Hpa* II 两种酶对基因组 DNA 进行双酶切,

并在两端加上区分不同样本的标签和接头, 样品混合后进行双端测序, 测序后获得的 SNP 标记数量多, 比传统分子标记更好地代表全基因组的遗传信息, 具有成本低、准确性和稳定性高、易于操作的优点^[10]。

杂合度分析有助于深入了解辣木的遗传组成情况, 确定繁育类型, 合理规划育种, 传统研究繁育类型的方法主要是基于对花器官的形态学分析, 传粉媒介的观察以及温室杂交试验展开, 主要通过表型性状进行评估, 易受环境、气候、栽培措施等因素影响, 不能准确地反映植物基因型^[11-13]。辣木群体结构的研究对于辣木种质资源的挖掘、利用和保护具有重要的理论和实践意义, 遗传多样性及群体分化分析是遗传学研究的核心内容, 亲本的遗传关系很大程度决定子代种子的质量, 亲本间存在差异的遗传信息会随着杂交或自交过程传递给子代, 使之在单核苷酸水平上呈现出来。目前辣木的繁育类型颇受争议, 还没有学者基于 SNP 对辣木的繁育类型进行研究, 国内辣木育种工作进展缓慢, 没有自主知识产权的辣木品种, 存在种子繁殖会发生性状分离及种子管理不规范等问题, 造成辣木优良品种缺乏、品种混乱的现象^[14], 辣木亲本和子代群体的遗传分析对于确定辣木繁育类型、分析亲缘关系及选育优良品种具有重要的意义。

本研究以 96 份辣木为研究材料, 结合基因组 AFSM 高通量测序技术, 与参考基因组进行比对后, 进行基因型分析、杂合度分析、群体结构、

遗传多样性、群体分化及连锁不平衡分析,以揭示辣木亲本与子代间的遗传关系,为辣木繁育类型和种质亲缘关系提供理论指导,以及为发掘控制辣木种质优良性状的优异等位基因提供理论依据。

1 材料与方法

1.1 材料

选取来源于同一母本通过自然授粉得到的 YMLM002 辣木种子 94 粒,该种质是经过连续 3a 的跟踪观测筛选出的果用型优良单株材料,具有产量高、果型好、种子饱满的特点^[15]。辣木种子先用清水浸泡 10 h,软化种子硬壳,再用 100 mg/L 高锰酸钾溶液浸泡 0.5 h 消毒,清水洗净后点播于穴盘中(红土:蛭石=1:1),适时补充水分保证湿润,待苗长至 15 cm 左右,收集 94 份子代样品、1 份母株和 1 份扦插苗样品备用。

1.2 方法

1.2.1 辣木基因组 DNA 提取及建库 采用 CTAB 法提取辣木样品 DNA,用 Nano Drop ND-1000 对 DNA 样品浓度进行检测,并调节样品浓度至 100 ng/ μ L,置于-20 °C 保存。采用 AFSM 技术进行建库,利用 *EcoR* I-*Msp* I 和 *EcoR* I-*Hpa* II 两种酶对 96 份辣木 DNA 样品进行混合双酶切,再将酶切产物连接加上用于区分不同样品的接头标签,纯化后进行 PCR 扩增,样品混合后再用高通量测序平台 Illumina 进行双端测序,并计算 GC 含量和 Q30 评估测序数据质量。

1.2.2 辣木群体基因型分析 利用 Perl 脚本对原始测序数据进行过滤,使用 Bowtie 软件将过滤数据比对到辣木参考基因组 ASM980114v1,再使用 VCFtools 和 BCFtools 软件检测并统计 SNP 和 Indel 位点信息。

1.2.3 辣木基因杂合度分析 使用 AWK 语言分析 96 份样品的杂合位点,并计算个体内基因的杂合位点比率即为个体内杂合度;同时通过将子代数据分别与亲本进行比对,找出差异位点,统计差异位点概率即为子代与亲本比对杂合度,分别生成个体内杂合度及子代与亲本比对杂合度统计图。

1.2.4 群体结构分析、遗传多样性及群体分化分析 利用 Plink 对变异位点进行过滤,过滤掉最小等位基因频率低于 0.05 及基因型缺失率小于 5% 的位点,哈迪-温伯格检验显著性 $P > 0.0001$,保留高质量的变异位点,再使用 ADMIXTURE 软件进行群体结构分析,将亚群数 K 值范围设置为

1~10,根据得到的交叉验证错误率(cross-validation error, CV error)值选择合适的亚群数 K 值,以个体占亚群的遗传成分系数确定个体归属的类群,用 R 软件绘制群体遗传结构矩阵图。采用 GCTA 软件对过滤得到的高质量文件进行主成分分析,并用 R 软件绘图;采用 VCFtools 软件计算辣木群体的遗传多样性指数(π)及群体分化指数(F_{st})。

1.2.5 连锁不平衡分析 使用 LDBlockShow 软件,计算不同标记距离下的 D' 值,并生成单体型块图以展示位点间的连锁不平衡程度。

2 结果与分析

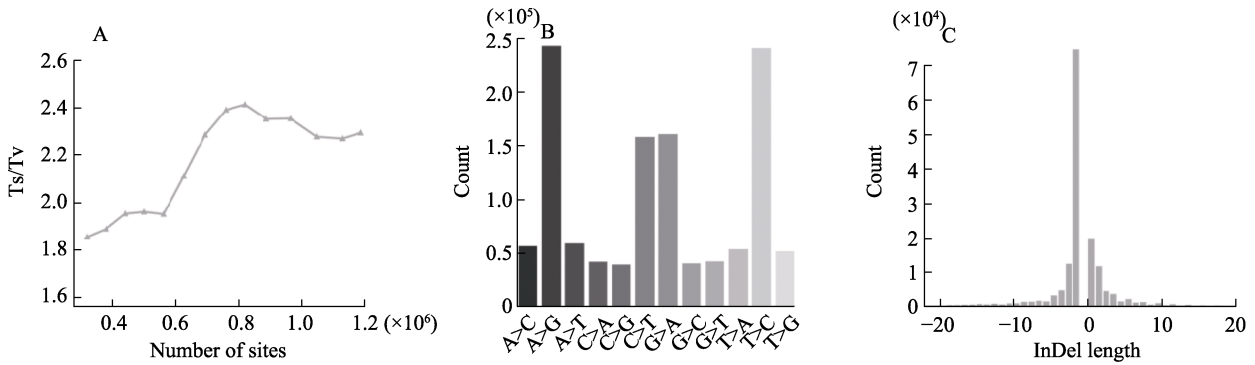
2.1 辣木群体基因型分析

96 份辣木样品基因组 DNA 经过 AFSM 技术建库、测序,将数据过滤并比对至辣木参考基因组 ASM980114v1,参考基因组信息来源于 NCBI 数据库(<https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/3735/>),基因组大小为 253.9 Mb,测序得到 1.8 G 数据文件,346 615 757 条 reads,测序长度为 150 bp,平均 GC 值为 50.53%,平均 Q_{30} 为 94.49%。采用 VCFtools 和 BCFtools 软件处理样品基因组数据后,得到 1 187 831 个 SNP 和 150 861 个 Indel 位点,以及 11 158 个多等位基因位点,4930 个多等位基因 SNP 位点。SNP 同类型碱基之间的突变为转换,不同类型碱基之间的突变为颠换,SNPs 发生转换概率与颠换概率的比值为 2.08,单一序列发生转换次数为 804 031,单一序列发生颠换次数为 383 471(图 1)。

SNPs 发生碱基转换和颠换概率随着位点的增大呈现先逐渐增加后缓慢降低的趋势(图 1)。辣木不同类型的突变位点中,碱基转换的变异数量显著大于颠换的数量,其中碱基 G/A 和 C/T 的替换都较高,分别为 243 672 和 241 616 次;碱基 A/G 和 T/C 的替换次之,分别为 161 258 和 158 648 次,碱基插入/缺失发生的次数随着碱基插入/缺失长度的增加而呈现出迅速下降的趋势。

2.2 辣木自然结实子代基因杂合度分析

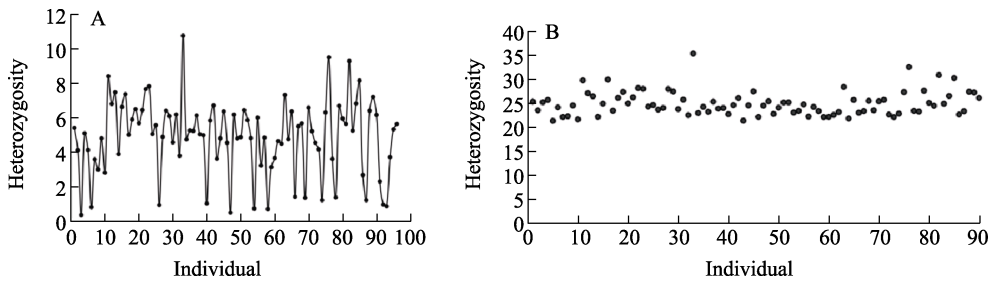
采用 1 187 831 个 SNP 位点和 150 861 个 Indel 位点分析 96 份辣木样品的杂合度(图 2)。辣木同源染色体上的 SNP 位点为同一类型碱基,则该 SNP 位点称为纯合 SNP 位点,若为不同类型碱基,则为 SNP 杂合位点。由图 2A 可知,辣木个



A: 碱基转换与颠换概率分布; B: 碱基转换类型分布; C: 插入/缺失长度分布。
 A: Probability distribution of DNA base transition and transition; B: Distribution of DNA base conversion;
 C: Distribution of inset/deletion length.

图 1 辣木基因型分析

Fig. 1 Genotype analysis of *M. oleifera*



A: 个体内杂合度统计图; B: 子代与亲本比对杂合度统计图。
 A: Statistics of heterozygosity within individuals; B: Statistics of heterozygosity between individuals.

图 2 辣木基因杂合度分析

Fig. 2 Analysis of gene heterozygosity of *M. oleifera*

体内杂合度在 10.79%~0.36%之间, 个体内平均杂合度为 4.89%, 其中, 母株杂合度为 5.65%, 扦插苗杂合度为 5.34%。由图 2B 可知, 子代与亲本比对杂合度在 21.22%~35.33%之间, 子代与亲本的比对平均杂合度为 24.85%。由此可知, 导致辣木子代杂合的基因中, 约有 4.89%的基因为自身杂合基因, 19.96%为外来遗传物质导致杂合的基因, 基本表明辣木通过自花和异花 2 种授粉方式繁衍后代。

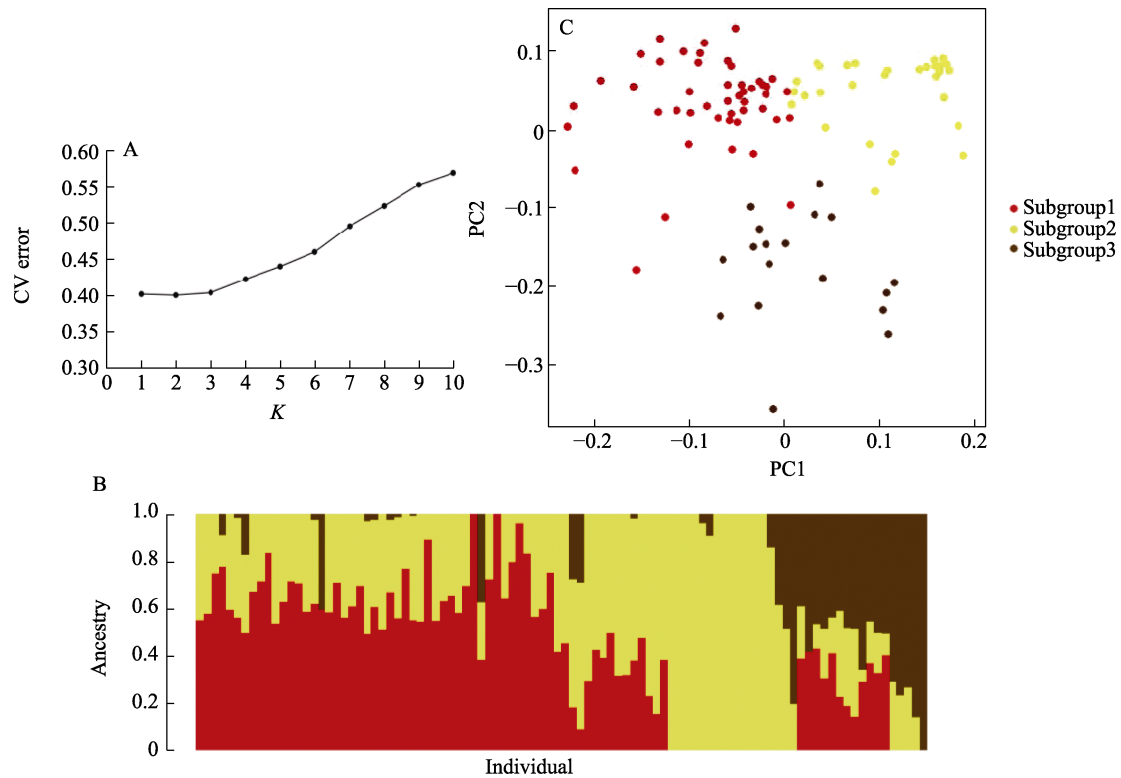
2.3 辣木群体结构分析

采用 Plink 对变异位点进行过滤后, 得到 141 323 个 SNP 位点, 再通过软件利用所有 SNP 和 Indels 分子标记对 96 份辣木样品进行群体遗传结构分析, 由图 3A 可知, 当 K 值为 3 时, 随着 K 值的增大, CV error 逐渐增大。由于 K 值为 2 和 3 时, CV error 值均较小且较为接近, 分别为 0.401 和 0.404, 但当 K 值为 2, 即将 96 份样品分为 2 个亚群时, 各亚群的个体呈现分布不集中的现象, 故将 96 份辣木样品分为 3 个亚群 (subgroup1-3)。根据个体在 3 个亚群的 Q 值, 将个体

归类到 Q 值占比最大的亚群 (图 3B), 发现 3 个亚群中分别有 47、31、18 份材料, 其中母株和扦插苗属于亚群 1, 亚群 2、亚群 3 均为子代样品。

主成分分析发现 (图 3C), 亚群 1 和亚群 2 在 PC1 轴上有分布差距, 而亚群 3 与亚群 1、2 在 PC2 轴上有一定的分布差距。大部分亚群可以聚类在一起, 表明聚类结果与群体结构的划分具有一致性。同时, 上述结果 (辣木亲本与子代样本聚类为 3 个亚群) 再次论证了杂合度分析结果, 即在生殖遗传的过程中, 辣木并非以自花授粉的方式繁衍后代, 在一定程度上接受了外来的花粉, 导致后代在不包含母株的另外 2 个群体中有分布。

由图 4 可知, 群体进化树的聚类结果与群体结构的划分相一致, 各亚群大致能聚在一起, 且样品间有一定的交叉。相比较而言, 亚群 1 的分枝长度较短, 有 4 个个体分散在亚群 2 中; 亚群 2 的分布总体集中, 有 7 个个体与亚群 3 有交叉; 同时, 亚群 3 有 3 个个体与亚群 1 有交叉。



A: CV 值随 K 值的变化分析; B: 辣木群体结构; C: 主成分分析。

A: Analysis of the variation of CV value with the K value; B: Group structure of moringa ; C: Analysis of principal component.

图 3 辣木的群体遗传结构

Fig. 3 Population genetic structure of *M. oleifera*

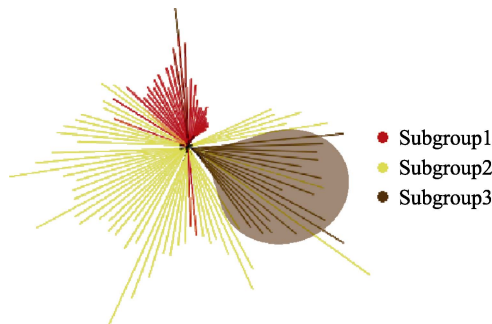


图 4 辣木的系统进化树

Fig. 4 Phylogenetic tree of *M. oleifera*

2.4 辣木群体遗传多样性及分化分析

3 个亚群的 π 值差距较小且均较低, 平均 π 值也较低, 为 0.0010, 表明 96 个辣木群体的遗传多样性水平低。各亚群的 F_{st} 在 0.0049~0.0110 之间, 其中亚群 1 和亚群 2 间的 F_{st} 值最小, 亚群 2 和亚群 3 间的 F_{st} 值最大, 各亚群间的 F_{st} 值均小于 0.05, 表明各样本之间存在较弱的遗传分化(当 F_{st} 等于 0 或 1 时, 分别表明亚群间没有分化或完全分化; 当 F_{st} 为 0~0.05 时, 表明亚群间的分化较弱; 当 F_{st} 为 0.05~0.15 时, 表明亚群间为中度分化; 当 F_{st} 为 0.15~0.25 时, 表明亚群间的分化

较强^[16]), 各亚群间的亲缘关系相对较近。

2.5 连锁不平衡分析

结合多态性核心 SNP 位点在辣木基因组上对应位置分析, 发现共有 136 个 Scaffold, 主要 Scaffold 统计情况见表 1 (SNP 数量前十的 Scaffold)。其中, Scaffold 1 的 SNP 数目最多, 为 62 225 个, Scaffold 122 的 SNP 数目最少, 为 288 个。通过 LDBlockShow 软件对 Scaffold 1 在 6.748~6.749 Mb 区域内的变异信息进行连锁不平衡分析, 发现 6 748 044~6 748 185 位点之间具有强连锁不平衡关系, 而 6 748 040 与 6 748 041、6 748 041 与 6 748 044 等位点间的连锁关系弱 (图 5)。

3 讨论

AFSM 技术采用 *EcoR* I-*Hpa* II 和 *EcoR* I-*Msp* I 两组双酶切体系简化基因组 DNA 的复杂度, 目前已发展得较为成熟, 已用于检测巴西木薯、澳洲坚果、麻疯树等植物的 SNP、Indel 及甲基化位点^[17-19], 该技术 DNA 处理步骤和数据分析步骤相对简单, 效率高, 测定的位点稳定, 无需进行超声剪切或荧光标记, 试验成本低, 适用于大量

表 1 主要支架数目统计

Tab. 1 Statistics of the number of main Scaffold

项目 Item	Scaffold 1	Scaffold 25	Scaffold 2	Scaffold 16	Scaffold 9	Scaffold 48	Scaffold 15	Scaffold 30	Scaffold 32	Scaffold 8
SNP 数目	62 225	45 379	39 993	37 312	36 782	29 386	28 269	28 125	26 778	26 230

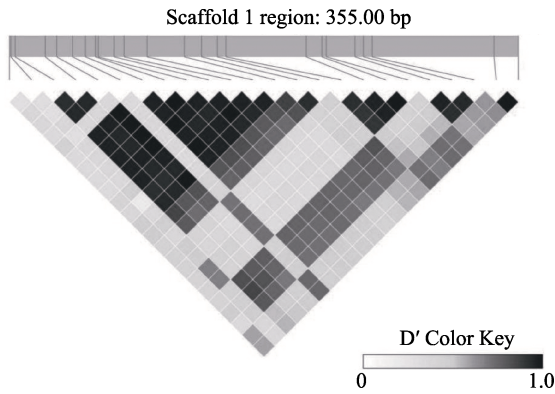


图 5 Scaffold 1 连锁不平衡分析图

Fig. 5 Analysis linkage disequilibrium of Scaffold 1

非模型物种的基因分型。本研究利用该技术得到 1 187 831 个 SNP 和 150 861 个 Indel 位点，可实现辣木亲本及子代遗传分析的目的。

国内外相关学者从不同的角度对辣木的繁育系统进行研究，吕亚等^[20]发现狭瓣辣木在开花第一天就有花粉活力和微弱的柱头可授性，且开花之初柱头高于雄蕊，之后逐渐低于雄蕊。MULUVI 等^[21]利用 AFLP 分子标记研究肯尼亚种源辣木的繁育系统，表明该种源辣木种子是自交和异交的混合产物。起国海^[22]研究辣木对于干热河谷传粉网络的影响，并表明辣木单花能提供 5~30 μL ，含糖量高达 60.5% 的花蜜报酬物，属于昆虫传粉植物，主要传粉者为蜂类。本研究中辣木个体内平均杂合度 4.89%，子代与亲本的比对平均杂合度为 24.85%，表明辣木繁殖方式为自交与异交同时存在。因此，在进行辣木杂种优势利用时，需要关注相关个体间的随机化分布和最小距离，以最大限度地增加差异品种/系间的杂交受精，并尽量减少品种内部的自交。

植物的繁育系统、选择、遗传漂移、突变和迁移是影响植物群体遗传结构的进化因子^[23]，本研究利用 ADMIXTURE 软件对辣木的群体结构进行分析，将 96 个辣木群体划分成 3 个亚群，该结果与聚类分析和主成分分析的结果类似，3 种群体结构分析方法相互补充印证，表明辣木群体的遗传结构划分可靠。在群体结构划分中，大部分亚群可以聚类在一起，其中，1 亚群有亲本及

子代样品，而 2、3 亚群均为子代样品，该结果表明辣木自然结实子代群体除了携带亲本的遗传信息外，还携带有外来的遗传信息，即辣木繁衍后代的方式不仅为自花授粉，而且还存在异花授粉。子代样品在 3 个亚群中均有分布，可能是由于亲本植株种植于保存有不同辣木种质的资源圃内，不同来源材料的花粉传播至亲本植株所导致的。

群体的 π 值和 F_{st} 值是衡量群体遗传分化程度的重要参数，RAJALAKSHMI 等^[24]使用 ISSR、SRAP 标记研究印度种源辣木的遗传多样性，表明辣木的平均遗传分化系数为 0.15，总遗传多样性指数为 0.17。本研究发现辣木群体的 π 值为 0.0010， F_{st} 值在 0.0049~0.0110 之间，表明本研究所用辣木群体遗传分化弱，遗传多样性较低，该现象应该这是由于选取的辣木群体是亲本与子代亲缘关系较近造成的。同时，也表明引起子代杂合的外来基因与亲本的基因型差异不大，这可能是由于本研究文采用的是栽培种辣木资源，经历了多次的人工选择育种，资源间丰度低造成的，后续可引进印度、非洲种源的优良辣木种质，以丰富资源圃内的辣木种质^[25]。

当 2 个距离较近的等位基因在同一单体型上同时出现的频率高于随机出现的频率时，表明它们处于连锁不平衡状态。在定位克隆中，通过连锁可检测到产生连锁信号的变异，在关联分析中，利用邻近位点形成的强连锁不平衡，有助于找到与性状相关的位点^[26-27]。本文对辣木的 SNP 位点进行了连锁不平衡分析，并在单体型块图上发现了连锁不平衡关系强的基因区域，可为研究多个处于连锁不平衡的位点与重要性状的关联性提供参考依据。

参考文献

- [1] ALNIDAWI A, ALI F, ABDELGAYED S, AHMED F, FARID M. *Moringa oleifera* leaves in broiler diets: effect on chicken performance and health[J]. Food Science and Quality Management, 2016, 58: 40-48.
- [2] SHAD M, XIANG S P. Application of *Moringa (Moringa oleifera)* as natural feed supplement in poultry diets[J]. Ani-

- mals (Basel), 2019, 9(7): 431.
- [3] SPADA A, BATTEZZATI A, SCHIRALDI A, ARISTIL J, BERTOLI S. Cultivation, genetic, ethnopharmacology, phytochemistry and pharmacology of *Moringa oleifera* leaves: an overview[J]. International Journal of Molecular Sciences, 2015, 16(6): 12791-12835.
- [4] 普天磊, 韩学琴, 廖承飞, 邓红山, 罗会英, 金杰. 辣木抗逆相关研究现状及应用[J]. 生物技术通报, 2020, 36(11): 133-140.
PU T L, HAN X Q, LIAO C F, DENG H S, LUO H Y, JIN J. Research status on the stress resistance of *Moringa oleifera* and its application[J]. Biotechnology Bulletin, 2020, 36(11): 133-140. (in Chinese)
- [5] FAHEY J W, OLSON M E, STEPHENSON K K, WADE K L, DAVID O. The diversity of chemoprotective glucosinolates in *Moringaceae* (*Moringa* spp.)[J]. Scientific Reports, 2018, 8(1): 7994.
- [6] NUR ZAHIRAH A R, KHAIRANA H, ENDANG K. *Moringa* genus: a review of phytochemistry and pharmacology[J]. Frontiers in Pharmacology, 2018, 9: 108.
- [7] 王富强, 樊秀彩, 张颖, 刘崇怀, 姜建福. SNP 分子标记在作物品种鉴定中的应用和展望[J]. 植物遗传资源学报, 2020, 21(5): 1308-1320.
WANG F Q, FAN X C, ZHANG Y, LIU C H, JIANG J F. Application and prospect of snp molecular markers in crop variety identification[J]. Journal of Plant Genetic Resources, 2020, 21(5): 1308-1320. (in Chinese)
- [8] 高嵩, 刘宏伟, 何欢, 吕庆雪, 周德龙, 邢跃先, 赵兴彦, 张志军, 张建新, 仲义, 夏远峰, 宋广树. 利用 SNP 芯片进行玉米遗传多样性和群体遗传结构分析及新品种选育[J]. 玉米科学, 2021, 29(1): 39-45.
GAO S, LIU H W, HE H, LYU Q X, ZHOU D L, XING Y X, ZHAO X Y, ZHANG Z J, ZHANG J X, ZHONG Y, XIA Y F, SONG G S. Analysis on maize genetic diversity and population genetic structure by SNP chip & breeding of new maize varieties[J]. Journal of Maize Sciences, 2021, 29(1): 39-45. (in Chinese)
- [9] 韩志刚, 郝文胜, 谢锐, 郭景山, 伊六喜, 侯建华. 基于全基因组重测序 SNP 标记的 148 份马铃薯种质遗传多样性分析[J]. 西北植物学报, 2021, 41(8): 1302-1314.
HAN Z G, HAO W S, XIE R, GUO J S, YI L X, HOU J H. Analysis of genetic diversity of 148 potato germplasm based on SNP markers from whole genome resequencing[J]. Acta Botanica Boreali-Occidentalia Sinica, 2021, 41(8): 1302-1314. (in Chinese)
- [10] XIA Z, ZOU M, ZHANG S, FENG B, WANG W. AFSSM sequencing approach: a simple and rapid method for genome-wide SNP and methylation site discovery and genetic mapping[J]. Scientific Reports, 2014, 4: 1-8.
- [11] 尹子丽, 冯泽辉, 谭慧, 谭文红. 药用植物大理藜芦的开花特性及其繁育系统研究[J]. 中药材, 2021, 44(2): 273-277.
YIN Z L, FENG Z H, TAN H, TAN W H. Study on flowering characteristics and breeding system of medicinal plant *Veratrum taliense*[J]. Journal of Chinese Medicinal Materials, 2021, 44(2): 273-277. (in Chinese)
- [12] 李海燕, 张午曲, 李宏博. 朝鲜白头翁开花与繁育特性研究[J]. 时珍国医国药, 2020, 31(4): 942-944.
LI H Y, ZHANG W Q, LI H B. Study on floral biology and breeding system of *Pulsatilla koreana*[J]. Lishizhen Medicine and Madica Research, 2020, 31(4): 942-944. (in Chinese)
- [13] 曾建军, 王路平, 武忆雯. 井冈山地区铁皮石斛的繁殖生物学研究[J]. 井冈山大学学报(自然科学版), 2019, 40(5): 21-25.
ZENG J J, WANG L P, WU Y W. Reproductive biology of *Dendrobium officinale* in mount Jinggangshan region[J]. Journal of Jinggangshan University (Natural Science), 2019, 40(5): 21-25. (in Chinese)
- [14] 林宗铿, 张天翔, 杨俊杰. 基于 SRAP 的辣木种质资源遗传多样性和亲缘关系分析[J]. 热带作物学报, 2021, 42(4): 945-950.
LIN Z K, ZHANG T X, YANG J J. Assessment of genetic diversity and relationship among 18 *Moringa* species based on SRAP marker[J]. Chinese Journal of Tropical Crops, 2021, 42(4): 945-950. (in Chinese)
- [15] 罗会英, 普天磊, 邓红山, 赵琼玲, 韩学琴, 范建成, 廖承飞, 金杰. 辣木优良单株筛选比较[J]. 中国农学通报, 2020, 36(34): 32-37.
LUO H Y, PU T L, DENG H S, ZHAO Q L, HAN X Q, FAN J C, LIAO C F, JIN J. Selection and comparison of excellent individual of *Moringa oleifera*[J]. Chinese Agricultural Science Bulletin, 2020, 36(34): 32-37. (in Chinese)
- [16] WRIGHT S. Evolution and the genetics of population 4: variability within and among natural populations[M]. Chicago & London: the University of Chicago Press, 1978: 157.
- [17] 孙倩, 邹枚伶, 张辰笈, 江思容, DE OLIVEIRA E J, 张圣奎, 夏志强, 王文泉, 李有志. 基于 SNP 和 InDel 标记的巴西木薯遗传多样性与群体遗传结构分析[J]. 作物学报, 2021, 47(1): 42-49.
SUN Q, ZOU M L, ZHANG C J, JIANG S R, DE OLIVEIRA E J, ZHANG S K, XIA Z Q, WANG W Q, LI Y Z. Genetic diversity and population structure analysis by SNP and InDel markers of cassava in Brazil[J]. Acta Agronomica Sinica, 2021, 47(1): 42-49. (in Chinese)
- [18] 谭秋锦, 王文林, 陈海生, 韦媛荣, 郑树芳, 黄锡云, 贺鹏, 夏志强. 基于 SNP 分子标记的澳洲坚果种质遗传多样性

- 分析[J]. 分子植物育种, 2020, 18(21): 7246-7253.
- TAN Q J, WANG W L, CHEN H S, WEI Y R, ZHENG S F, HUANG X Y, HE P, XIA Z Q. Genetic diversity analysis of macadamia varieties based on single-nucleotide polymorphism[J]. *Molecular Plant Breeding*, 2020, 18(21): 7246-7253. (in Chinese)
- [19] 张辰笈, 邹枚伶, 夏志强, 孙倩, 江思容, 王文泉. 麻疯树杂交子代与亲本间基因组 DNA 甲基化分析[J]. 热带作物学报, 2021, 42(3): 620-628.
- ZHANG C J, ZOU M L, XIA Z Q, SUN Q, JIANG S R, WANG W Q. Analysis of genomic DNA methylation between progenies and parents for *Jatropha curcas*[J]. *Chinese Journal of Tropical Crops*, 2021, 42(3): 620-628. (in Chinese)
- [20] 吕亚, 王楞, 马志亮, 杨焱, 李海泉, 张祖兵. 狭瓣辣木的开花动态及花粉活力和柱头可授性[J]. 分子植物育种, 2020(7): 2371-2375.
- LYU Y, WANG L, MA Z L, YANG Y, LI H Q, ZHANG Z B. Floral dynamics, pollen viability and stigma receptivity of *Moringa stenopetala*[J]. *Molecular Plant Breeding*, 2020, 18(7): 2371-2375. (in Chinese)
- [21] MULUVI G, SPRENT J, ODEE D, POWELL W. Estimates of outcrossing rates in *Moringa oleifera* using amplified fragment length polymorphism (AFLP)[J]. *African Journal of Biotechnology*, 2004, 3(2): 146-151.
- [22] 起国海. 外来引种植物印楝和辣木对红河干热河谷传粉网络的影响[D]. 北京: 中国林业科学研究院, 2018.
- QI G H. Effect of exotic introduced plants *Azadirachta indica* and *Moringa oleifera* on pollination network of Dry-hot Valley in Honghe Area[D]. Beijing: Chinese Academy of Forestry, 2018. (in Chinese)
- [23] 徐刚标. 植物群体遗传学[M]. 北京: 科学出版社, 2009.
- XU G B. *Plant population genetics*[M]. Beijing: Science Press, 2009. (in Chinese)
- [24] RAJALAKSHMI R, RAJALAKSHMI S, PARIDA A. Genetic diversity, population structure and correlation study in *Moringa oleifera* Lam. using ISSR and SRAP markers[J]. *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences*, 2019, 89(4): 1361-1371.
- [25] VARSHNEY R K, THUDI M, ROORKIWAL M, HE W, UPADHYAYA H D. Resequencing of 429 chickpea accessions from 45 countries provides insights into genome diversity, domestication and agronomic traits[J]. *Nature Genetics*, 2019, 51(5): 857-864.
- [26] 梁晓宇. 利用连锁不平衡发掘大豆灰斑病 10 号小种抗性位点[D]. 哈尔滨: 黑龙江大学, 2019.
- LIANG X Y. Identification of the resistance locus of soybean gray spot 10 based on linkage disequilibrium[D]. Harbin: Heilongjiang University, 2019. (in Chinese)
- [27] 万何平, 陈禅友, 陈高, 曹新华, 夏明. 全基因组关联分析在大豆遗传学上的研究进展[J]. 江汉大学学报(自然科学版), 2019, 47(3): 197-203.
- WAN H P, CHEN C Y, CHEN G, CAO X H, XIA M. Research status of genome-wide association study in soybean[J]. *Journal of Jiangnan University (Natural Science)*, 2019, 47(3): 197-203. (in Chinese)