

单建伟, 索海翠, 王丽, 安康, 刘计涛, 李成晨, 白建明, 李小波. 基于 ddRADseq 的马铃薯品种遗传多样性分析 [J]. 广东农业科学, 2021, 48(12): 120-128.

基于 ddRADseq 的马铃薯品种遗传多样性分析

单建伟¹, 索海翠¹, 王丽¹, 安康¹, 刘计涛¹, 李成晨¹, 白建明², 李小波¹

(1. 广东省农业科学院作物研究所 / 广东省农作物遗传改良重点实验室, 广东 广州 510640;

2. 云南省农业科学院经济作物研究所, 云南 昆明 650000)

摘要: 【目的】收集国内外马铃薯品种(系)资源, 研究马铃薯种质群体结构和遗传多样性, 开发SNP分子标记, 为马铃薯品种鉴定、分子标记辅助选择育种奠定基础。【方法】提取马铃薯嫩芽基因组DNA, 利用限制性内切酶MseI和SacI进行双酶切建库, 质量检测合格后用Illumina HiSeq平台进行双末端测序, 用BWA软件将测序数据与马铃薯参考基因组进行比对, 采用GATK软件进行SNP和InDel变异位点检测, 利用ANNOVAR软件进行变异注释, 并基于上述变异信息对群体结构进行Structure分析、主成分分析(PCA)和遗传多样性分析。

【结果】通过简化基因组测序共获得 7.50×10^8 条测序reads和 2.04×10^{11} 个碱基, 共检测到39 038个变异位点, 其中SNP位点36 267个, InDel位点2 771个; 基于上述变异位点的Structure和PCA分析均将研究群体划分为两个亚群, 群体系统发生分析表明, G2亚群中的个体聚类在一起, 亲缘关系较近, 但与其他个体相比G2亚群中的分枝距离中心较远, 积累了更多的变异量; 群体的平均多态性信息含量 $PIC=0.3107$, 期望杂合度 $He=0.3932$, 观测杂合度 $Ho=0.1852$, 群体自交系数 $Fis=0.553$, 表明马铃薯品种(系)间遗传多样性较低; 开发了覆盖马铃薯12条染色体且包含120个SNP标记的SNP-Panel, 并选取其中的60个标记在46个样品中进行了验证。【结论】研究结果为马铃薯遗传多样性研究、分子标记辅助选择育种、分子身份证开发、遗传图谱构建奠定了一定基础。

关键词: 马铃薯; 遗传多样性; 群体结构; SNP 标记; ddRADseq

中图分类号: S532.032

文献标志码: A

文章编号: 1004-874X(2021)12-0120-09

Analysis on Genetic Diversity of Potato Varieties Based on ddRADseq

SHAN Jianwei¹, SUO Haicui¹, WANG Li¹, AN Kang¹, LIU Jitao¹,

LI Chengchen¹, BAI Jianming², LI Xiaobo¹

(1. Crops Research Institute, Guangdong Academy of Agricultural Sciences/

Guangdong Key Laboratory of Crops Genetics & Improvement, Guangzhou 510640, China;

2. Industrial Crops Institute, Yunnan Academy of Agricultural Sciences, Kunming 650000, China)

Abstract: 【Objective】The present research aimed to collect potato variety (strain) resources, study the population structure and genetic diversity of potato germplasm and develop SNP molecular markers, with a view to providing theoretical foundation for potato variety identification and molecular marker-assisted selection breeding.

【Method】Genome DNA of potato tender shoots was extracted and digested with restriction enzymes MseI and SacI. After quality verification, Paired-end sequencing was carried out on Illumina HiSeq platform, sequencing data was

收稿日期: 2021-11-07

基金项目: 广东省农业科学院作物研究所所长基金(201904); 广东省农作物遗传改良重点实验室开放基金(201904); 广东省科技计划项目(2017B020232002); 广东省现代农业产业技术体系创新团队项目(2021KJ111)

作者简介: 单建伟(1980—), 男, 博士, 助理研究员, 研究方向为马铃薯分子生物学及遗传育种, E-mail: chanyujianwei@126.com

通信作者: 李小波(1981—), 男, 博士, 研究员, 研究方向为马铃薯遗传育种及示范推广, E-mail: Lixiaobo1981@163.com

aligned to potato reference genome with BWA software, SNP and InDel variation sites were detected and annotated with GATK and ANNOVAR respectively, and population structure analysis, principal component analysis and genetic diversity analysis were conducted based on the SNPs. 【Result】 By simplifying genome sequencing, 7.50×10^8 sequencing reads and 2.04×10^{11} bases were obtained. A total of 39038 variation sites were detected, of which 36 267 were SNP sites and the other 2 771 sites were InDel. Based on the Structure analysis and PCA analysis, the 185 samples were divided into 2 subgroups. Population phylogenetic analysis showed that the individuals in G2 subgroup were clustered together with close genetic relationship, but the branches in G2 subgroup were far from the center and accumulated more variations than other individuals. *PIC*, *He*, *Ho* and *Fis* were 0.3107, 0.3932, 0.1852 and 0.553 respectively, which indicated that the genetic diversity of potato varieties (strains) was poor, SNP-Panel with 120 SNPs distributing on all the 12 potato chromosomes was developed and 60 SNPs of 46 samples were verified. 【Conclusion】 The research results laid a foundation for the study of potato genetic diversity, molecular marker-assisted selection breeding, development of molecular identity card and construction of genetic map.

Key words: potato; genetic diversity; population structure; SNP marker; ddRADseq

【研究意义】马铃薯为茄科 (Solanaceae) 茄属 (*Solanum*) 马铃薯组 (*Petota*) 草本植物, 原产于南美洲安第斯高原^[1]。马铃薯的植物分类学是一个不断更新和完善的过程, Spooner 等综合前人有关马铃薯表型、分子标记、系统发生、生殖生物学、倍性、核型等方面的研究将马铃薯组分类为 107 个野生种和 4 个栽培种^[2]。对于栽培作物而言, 通常在其起源地具有最丰富的遗传资源和遗传多样性, 而在其传播过程中遗传多样性逐渐降低^[3]。20 世纪 20 年代, 前苏联、英国、德国和美国学者先后赴南美洲考察和收集马铃薯野生种及栽培种种质资源, 为马铃薯的育种和遗传学研究开创了新纪元^[4]。当今世界范围内广泛种植的现代马铃薯品种 (modern cultivar) 是以马铃薯普通栽培种为基础, 通过与其他地方栽培种或野生种杂交利用了有利基因的产物, 使其能够适应更宽泛的生态条件和地理区域, 但目前广泛种植的马铃薯品种遗传多样性依然很低^[5]。广泛收集国内外马铃薯品种 (系) 资源, 对其群体结构和遗传多样性进行分析评价, 将为我国马铃薯遗传育种提供理论指导。

【前人研究进展】我国马铃薯的栽培历史并不长^[6], 对其育种起步也较晚, 始于 20 世纪 30 年代中末期, 早期的马铃薯育种工作以引种为主, 先后从英国、美国和前苏联等引进一批育成品种和杂交组合。由于马铃薯适应性广、产量高, 新中国成立以后为了尽快恢复农业生产, 国家有计划地组织调运种薯, 扩大马铃薯种植面积, 全国开展了马铃薯育种协作, 真正开始马铃薯育种工作, 促进了马铃薯在我国的迅猛发展, 并相继育

成了一批有代表性的马铃薯品种, 如东农 303、克新系列、陇薯系列等^[7-8]。当前, 我国马铃薯种植面积和总产量均超过全球的 1/4, 居世界首位。由于我国马铃薯育种工作起步较晚, 所用亲本多引自欧洲和北美等国, 对种质资源特别是野生种和地方栽培种的收集、研究、创新不够, 因此我国栽培的马铃薯品种遗传多样性更低, 许多品种之间同质性高, 部分品种存在同种异名现象, 仅靠表型差异难以分辨。

【本研究切入点】为了指导马铃薯遗传育种, 有必要收集国内外马铃薯品种 (系) 资源, 并对其遗传多样性进行评价。【拟解决的关键问题】本研究对收集的国内外马铃薯品种 (系) 资源, 利用限制性酶切位点关联 DNA 测序 (Restriction-site Associated DNA sequencing, RADseq) 中的 ddRADseq 技术对其进行简化基因组测序, 结合群体结构分析和遗传多样性分析对其进行群体结构和遗传多样性评价, 开发马铃薯 SNP-Panel, 为马铃薯育种、品种鉴定提供理论依据。

1 材料与方法

1.1 试验材料

供试材料为广东省农业科学院作物研究所马铃薯研究室收集的 185 份马铃薯品种 (系), 包含育成品种以及一些高代品系 (表 1)。这些材料分别来自我国的华南、华北、西南、华中地区以及国外的秘鲁和意大利。

1.2 试验方法

1.2.1 基因组 DNA 制备 取每个马铃薯品种 (系) 的嫩叶, 利用液氮速冻后储存于 -80°C 超

表 1 马铃薯样品来源
Table1 Sources of potato samples

| 序号 No. | 来源 Source | 品种 Variety |
|--------|-------------|---|
| 1 | 广东省农业科学院 | 1606-1、1613-1、1613-2、1613-28-1、1613-29、1619-1、BY3、Y16-1 Y16-11、Y16-12、Y16-13、Y16-22、Y16-24 Y16-25、Y16-3、Y16-4、Y16-6、Y17-10、Y17-11、Y17-12、Y17-13、Y17-14、Y17-15、Y17-16、Y17-17、Y17-18、Y17-19、Y17-20、Y17-9、大西洋 (DXY)、咖喱薯 (GaLi)、粤红一号 (YH1)、粤引 85-38 (YY85-38) |
| 2 | 云南农业科学院 | S17-166-11、S16-1342、S16-1346、S16-1539、S16-156、S16-1658、S16-2069、S16-407、S16-411、S16-425、S16-829、S17-1-1、S17-1-11、S17-114-10、S17-114-2、S17-114-4、S17-1-15、S17-11-51、S17-115-4、S17-115-6、S17-119-1、S17-119-2、S17-129-3、S17-147-10、S17-147-12、S17-147-13、S17-147-23、S17-147-25、S17-150-1、S17-150-10、S17-150-13、S17-150-17、S17-150-2、S17-150-3、S17-150-6、S17-150-8、S17-1-6、S17-172-2、S17-173-30、S17-173-4、S17-173-97、S17-176-1、S17-179-1、S17-179-10、S17-179-11、S17-179-2、S17-179-5、S17-179-9、S17-181-1、S17-181-12、S17-187-4、S17-188-13、S17-188-2、S17-188-4、S17-188-5、S17-1-9、S17-192-12、S17-201-1、S17-2-5、S17-27-1、S17-4-13、S17-4-21、S17-5-47、S17-5-67、S17-8-1、云薯 306 (YS306)、紫云一号-1 (ZiYun1-1)、紫云一号-2 (ZiYun1-2) |
| 3 | 华中农业大学 | 10A018-1、10FM72-23、11A002-13、11A010-4、11A017-6、11A036-4、11A043-1、11A043-2、11A052-1、11FM0028、11FM12-11、11FM14-6、11FM22-3、11FM43-6、12FM01-14、12FM01-3、13FF13-3、13FM40-2、14CH08-3、14EM03-7、15A019-3、15A029-2、15A029-5、15A032-12、15A032-31、15A032-40、15A032-44、15A032-45、15A032-47、15A032-48、15A032-49、15A032-5、15A032-57、15A032-62、15A032-65、15A032-8、15A043-1、15A043-2、15A047-1、15A051-1、15A073-1、15A073-3、15A073-4、15A075-1、15CH28-11、15CH28-3、15CH28-8、15CH51-02、15CH51-17、15CH61-2、15CH62-10、15CH62-5、15EM27-1、15EM27-2、15EM28-1、15EM32-1、15FM04-5、15FM18-05、15FM19-1、15FM42-2、15FM45-05、15FM49-2、AC338、Adriodack、华彩 1 号 (HuaCai1) |
| 4 | 甘肃农业科学院 | 陇薯 7 号 (LS7) |
| 5 | 青海省农林科学院 | 青 9 (Qing9) |
| 6 | 华颂种业股份有限公司 | 华颂 66 (HuaSong66)、华颂 88 (HuaSong88) |
| 7 | 内蒙古铃田生物技术公司 | 红美人 (HongMeiRen) |
| 8 | 意大利 | 意 1 (Y1)、意 2 (Y2)、意 3 (Y3)、意 4 (Y4)、意 5 (Y5)、意 6 (Y6) |
| 9 | 秘鲁 | BL12、BL16、BL18、BL22、BL23、BL25、BL26、BL27 |

低温冰箱，分离样品基因组 DNA 后，利用琼脂糖凝胶电泳检测 DNA 样品完整性及有无 RNA 污染，并利用分光光度计对 DNA 样品的质量和浓度进行定量检测。

1.2.2 文库构建、质控及测序 取约 200 ng 基因组 DNA 用限制性内切酶 *Mse I* 和 *Sac I* (纽英伦生物技术有限公司) 按照产品说明书进行完全酶切；酶切产物连接特异性接头，然后对连接产物进行纯化；用高保真聚合酶 KOD-Plus-Neo (东洋纺生物科技有限公司) 按照产品说明书进行 PCR 扩增富集；将所有扩增产物进行低压电泳，切取 300~400 bp 大小的电泳片段，用琼脂糖凝胶回收试剂盒 (QIAGEN, 德国) 进行纯化；用分光光度计对纯化产物进行初步定量，用安捷伦 2100 生物分析仪对文库插入片段大小进行检测；采用实时荧光定量 PCR 手段对文库的有效浓度进行准确定量，有效浓度 > 2 nmol/L 为合格；最后用 Illumina HiSeq 测序平台进行双末端测序。

1.2.3 测序数据过滤 测序得到的原始数据被称为 raw read，含有测序接头序列或低质量的碱基。

使用 Cutadapt (Version 1.13) 软件去除 raw read 中的接头序列，用 Trimmomatic (Version 0.36) 去除测序数据中的低质量碱基，得到 clean read 用于后续分析；测序 read 的长度必须大于 50 bp。

1.2.4 参考基因组比对 使用 BWA (Version 0.7.15-r1140) 软件中的 MEM 算法将测序数据与参考基因组 (http://solanaceae.plantbiology.msu.edu/pgsc_download.shtml) 进行比对，得到 SAM 格式的比对结果，然后使用 Samtools (Version 1.3.1) 将其转换为 BAM 格式，使用 Picard (Version 1.91) 软件中的 SortSam 对 BAM 文件中的序列进行排序，将处理后的 BAM 文件用于覆盖度和覆盖深度统计以及变异位点识别。

1.2.5 变异位点检测及注释 利用 GATK (Version 3.7) 软件包中的 HaplotypeCaller 模块对所有样品生成 gvcf 格式文件，然后用 GenotypeGVCFs 模块对所有样品进行 SNP 和 InDel 变异检测。并按照测序深度 ≥ 5、所有样品基因型缺失率 ≤ 50%、稀有等位基因频率 ≥ 10%、杂合率 ≤ 50%、相对杂合率 ≤ 67% 等标准对变异位点进行筛选，

得到的变异位点为有效变异位点。用 ANNOVAR (Version 2016Feb1) 对变异位点进行注释。依据变异位点在参考基因组上的位置信息以及参考基因组中的基因位置信息, 可得到变异位点在基因组中的区域(基因间区、基因区或 CDS 区等)信息。

1.2.6 马铃薯群体结构分析 通过 Structure 软件对群体结构进行分析, 结果以 ΔK 值变化图和祖先堆叠图展示, 可直观反映个体间的分类关系以及每个样品的“混血”程度。采用 PLINK (Version v1.90p) 软件对群体结构进行主成分分析(Principal Component Analysis, PCA)。PCA 散点图中, 两个样品距离越远, 表明遗传背景差异越大, 而遗传背景相似的个体在 PCA 散点图中聚类在一起。采用 MEGA7 (Version 7.0) 软件中的邻接法(Neighbor-joining methods) 构建系统发生树, 并通过 ggtree (Version 1.7.10) 进行可视化处理。基于分子标记信息, 利用 Cervus 3.0 软件计算多态性信息含量(Polymorphism Information Content, PIC)、观测杂合度(Observed Heterozygosity, H_o)、期望杂合度(Expected Heterozygosity, H_e)等。

1.2.7 SNP 分子标记开发 在综合区分度和检测成本的基础上, 结合 SNP 标记在染色体上的位置信息以及多态性, 选择 120 个在染色体上分布相对均匀的 SNP 标记, 分别在 SNP 位点上下游开发 PCR 引物, PCR 产物长度介于 150~350 bp 之间。

2 结果与分析

2.1 变异位点检测及注释

通过测序最终获得 clean read 7.50×10^8 条、 2.04×10^{11} 个碱基, 平均每个样品 4.05×10^6 条 clean read, 1.10×10^9 个碱基。为了检测变异位点, 将测序得到的 clean read 与马铃薯基因组 (http://solanaceae.plantbiology.msu.edu/pgsc_download.shtml) 进行比对, 并进行群体结构和遗传多样性进行分析。本研究在综合考虑可行性和经济性的基础上对测序覆盖度进行了控制, 样品平均测序覆盖度为 2.62%。利用 GATK 软件包共检测到 39 038 个有效变异位点, 其中 SNP 位点 36 267 个, InDel 位点 2 771 个。SNP 位点在马铃薯 12 条染色体上的平均分布密度为 50.02 个 /Mb SNP 位点, 但在不同染色体上的分布密度不均, 其中 10 号染色体上的 SNP 位点密度最低 (26.22 个

/Mb), 5 号染色体上 SNP 位点密度最高 (83.50 个 /Mb); InDel 变异位点在 12 条染色体上的分布密度介于 1.97~6.05 个 /Mb 之间, 同样 10 号染色体分布密度最低, 5 号染色体分布密度最高 (表 2)。同一条染色体不同区域的变异位点密度也存在较大差异, 如 1 号染色体 10~20 Mb 范围内的变异位点密度最高, 每 100 kb 存在约 90 个变异位点, 而在 70~90 Mb 范围内变异位点密度较低 (图 1)。

表 2 变异位点检测统计结果

Table 2 Detection and statistic results of variation sites

| 染色体编号 Chr. ID | 长度 Length (bp) | SNP | | InDel | |
|------------------|----------------------|--------------|------------------------|--------------|------------------------|
| | | 数量 Number | 密度 Density (/Mb) | 数量 Number | 密度 Density (/Mb) |
| Chr01 | 88663952 | 5724 | 64.56 | 382 | 4.31 |
| Chr02 | 48614681 | 2207 | 45.40 | 188 | 3.87 |
| Chr03 | 62290286 | 2272 | 36.47 | 206 | 3.31 |
| Chr04 | 72208621 | 3009 | 41.67 | 212 | 2.94 |
| Chr05 | 52070158 | 4348 | 83.50 | 315 | 6.05 |
| Chr06 | 59532096 | 2465 | 41.41 | 203 | 3.41 |
| Chr07 | 56760843 | 4010 | 70.65 | 267 | 4.70 |
| Chr08 | 56938457 | 2908 | 51.07 | 235 | 4.13 |
| Chr09 | 61540751 | 2736 | 44.46 | 240 | 3.90 |
| Chr10 | 59756223 | 1567 | 26.22 | 118 | 1.97 |
| Chr11 | 45475667 | 2027 | 44.57 | 183 | 4.02 |
| Chr12 | 61165649 | 2994 | 48.95 | 222 | 3.63 |
| 合计 Total | 725017384 | 36267 | | 2771 | |
| 平均 Mean | | | 50.02 | | 3.82 |

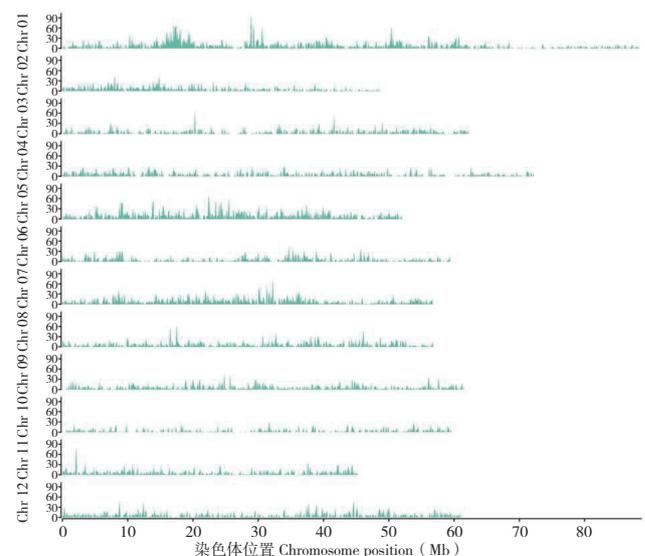


图 1 变异位点在染色体上的分布

Fig. 1 Distribution of variation sites on chromosomes

利用 ANNOVAR 软件结合变异位点在参考基因组和自身基因组上的位置信息对变异位点进行注释, 结果 (表 3) 显示, 共有 26 697 个 SNP 位点位于基因间区, 占有 SNP 位点的 73.65%; 1 392 个 SNP 位点位于基因转录起始位点上游 1 kb 范围内, 1 314 个 SNP 位点位于转录终止位点下游 1 kb 范围内, 116 个 SNP 位点位于一个基因转录起始位点上游 1 kb 范围内并同时在一个基因的转录终止位点下游 1 kb 范围内; 6 729 个 SNP 位点位于基因内部, 占有 SNP 位点的

18.56%, 其中 3 172 个 SNP 位点位于内含子区域, 2 605 个 SNP 位点位于外显子区; 352 个 SNP 位点位于 5' UTR 区, 597 个位点位于 3' UTR 区; 3 个 SNP 位点位于可变转录本上。此外, 共有 1 761 个 (63.57%) InDel 位点位于基因间区; 308 个 InDel 位点位于基因编码区上下游 1 kb 范围内; 25.31% 的 InDel 位点位于基因内部, 其中 406 个位于内含子内, 134 个位于外显子内; 3' UTR 和 5' UTR 区域内的 InDel 位点数分别为 101、56 个; 4 个 InDel 位点位于可变转录本内。

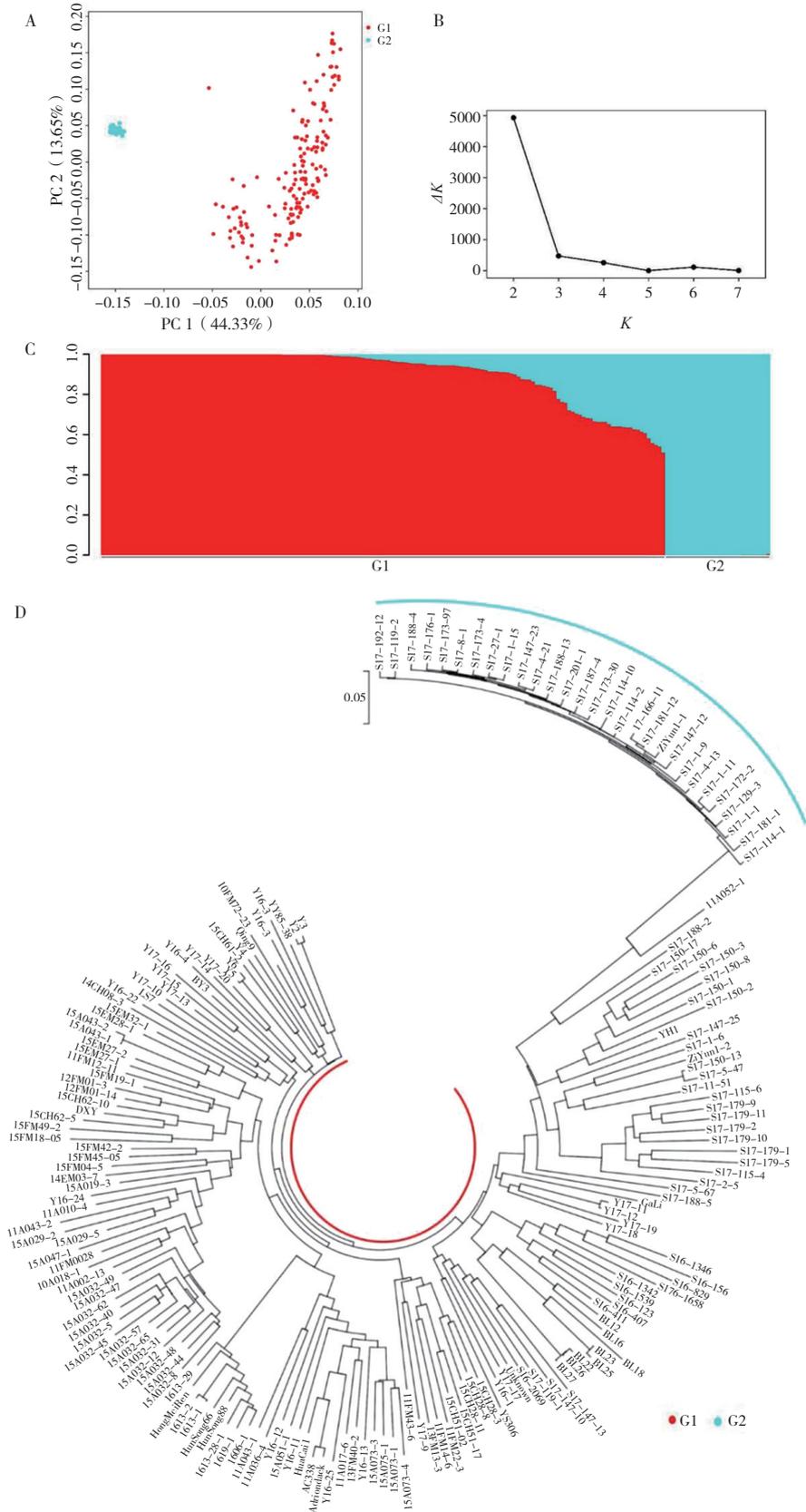
表 3 注释的变异位点分布
Table 3 Distribution of annotated variation sites

| 分类 Category | SNP | | InDel | |
|-------------------------------------|-----------|--------------|-----------|--------------|
| | 数量 Number | 比率 Ratio (%) | 数量 Number | 比率 Ratio (%) |
| 基因间区 Intergenic | 26697 | 73.65 | 1761 | 63.57 |
| 上 / 下游 Upstream/Downstream | 2822 | 7.79 | 308 | 11.12 |
| 上游 Upstream | 1392 | 3.84 | 157 | 5.67 |
| 下游 Downstream | 1314 | 3.63 | 132 | 4.77 |
| 上游和下游 Upstream & Downstream | 116 | 0.32 | 19 | 0.69 |
| 基因区 Genic | 6729 | 18.56 | 701 | 25.31 |
| 内含子 Intronic | 3172 | 8.75 | 406 | 14.66 |
| 外显子 Exonic | 2605 | 7.19 | 134 | 4.84 |
| 同义突变 / 非移码 Synonymous/nonframeshift | 999 | 2.76 | 39 | 1.41 |
| 非同义突变 / 移码 Nonsynonymous/frameshift | 1535 | 4.23 | 90 | 3.25 |
| 新增编码终止 Stopgain | 65 | 0.18 | 5 | 0.18 |
| 通译 Stoploss | 6 | 0.02 | 0 | 0.00 |
| 5' 非编码区 UTR5' | 352 | 0.97 | 56 | 2.02 |
| 3' 非编码区 UTR3' | 597 | 1.65 | 101 | 3.65 |
| 选择性剪切 Alternative splicing | 3 | 0.01 | 4 | 0.14 |

2.3 马铃薯群体结构分析

为了明确本研究所用马铃薯的群体结构, 通过 PCA 法对筛选得到的有效 SNP 标记进行降维处理, 从中提取关键信息将分子标记差异信息反映在二维坐标图上, 坐标轴为对样品差异性解释度最高的特征向量, 两个样品遗传背景差异越小, 则样品分子标记位点信息越相似, 反映在 PCA 图中的距离越近。PCA 散点图显示, 第一特征向量 (PC1) 和第二特征向量 (PC2) 可以将本研究所用群体区分为 G1 和 G2 两个亚群 (图 2A)。利用 Structure 软件对群体结构进行分析, 群体结构分析的 ΔK 值变化趋势图显示, 当 $K=2$ 时对应的 ΔK 值最大, 因而本研究群体的最佳亚群数为

2, 这一结果与 PCA 分析结果一致, 即将本研究群体分为两个亚群较为合适 (图 2B)。G2 亚群包含 29 个品种 (系) 全部来自云南省农业科学院, 分别为 S17-192-12、S17-114-4、S17-188-13、S17-119-2、S17-188-4、S17-114-2、S17-166-11、S17-147-12、S17-114-10、S17-1-1、S17-176-1、S17-173-97、S17-4-13、S17-201-1、S17-173-4、S17-147-23、S17-187-4、S17-1-9、S17-4-21、S17-173-30、S17-1-11、S17-129-3、S17-27-1、S17-8-1、S17-1-15、S17-181-12、S17-172-2、S17-181-1、紫云一号-1; 剩余的 156 个品种 (系) 构成 G1 亚群, 包括来自我国华南、华北、华中、西南、西北等地的品



A: 群体结构 PCA 分析; B: 群体结构分析的 ΔK 值变化; C: 群体结构祖先堆叠图; D: 群体系统发生树
 A: PCA analysis of population structure; B: Change in ΔK value of population structure analysis with Structure;
 C: Stacked diagram of ancestors; D: Phylogenetic tree of population

图 2 马铃薯群体结构分析
 Fig. 2 Analysis of potato population structure

种(系),以及引自秘鲁、意大利等国的品种(系)。祖先堆叠图(图2C)显示,两个亚群间存在明显基因交流,部分样本存在混血。群体系统进化分析结果(图2D)表明,所有G2亚群中的29个品种(系)聚类在一起且其分枝离圆心的距离最长,表明G2亚群中的个体亲缘关系较近,同时G2亚群中的个体在进化上积累了更多变异。

$PIC > 0.5$ 为高度多态性^[9],而本研究群体的平均 $PIC=0.3107$,表明群体遗传多样性中等偏低。同时,群体的期望杂合度 $He=0.3932$,观测杂合度 $Ho=0.1852$,群体自交系数 $Fis=0.553$,也表明群体遗传多样性偏低。

2.4 SNP 分子标记开发

靶向测序(target-seq)技术通过设计多重PCR引物对目标区域DNA片段进行特异性扩增,再通过双端标签区分不同样本,混样建库,进行多样本多位点高通量测序分析,具有信息量大、针对性强、低成本、效率极高、精准度高等特点,被广泛应用于分子标记辅助选择育种(Marker-Assisted Selection, MAS)、数量性状位点(Quantitative Trait Locus, QTL)精细定位、品种指纹分析、品种鉴定、基因编辑位点检测等领域。为了便于后续开展马铃薯亲缘关系鉴定、品种鉴别、群体遗传分析以及分子标记辅助选择育种,我们开发了包含120个马铃薯SNP标记的SNP-Panel,这些SNP标记均匀分布在马铃薯的12条染色体上(图3),并针对这120个SNP标记设计了分离目标DNA区域的特异引物用于后续多样本高通量SNP标记检测。为了验证SNP标记的有效性,选取了60个SNP标记,利用与所选SNP标记对应的引物对46份马铃薯品种(系)进行了靶向高通量测

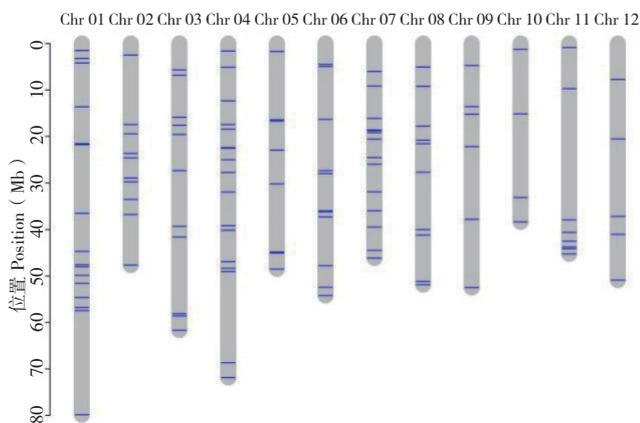


图3 SNP-Panel在马铃薯染色体上的分布

Fig. 3 Distribution of SNP-Panel on potato chromosomes

序,结果表明,60个SNP标记能够准确将46份马铃薯品种(系)区分开。理论上,所选标记越多,越能区分更多的品种(系),因此需要在区分度和检测成本之间寻求平衡。因此在综合区分度和检测成本的基础上,本研究确定了SNP标记的个数为120。

3 讨论

RADseq是在第二代测序技术基础上发展而来的一项基于全基因组酶切位点的简化基因组测序技术,该技术具有检测速度快、高通量、成本低等优点,可以检测出上千万的SNP位点,能够满足群体结构分析、遗传图谱构建和QTL定位等研究需求^[10-11]。依据所使用的限制性内切酶的数量和种类,RADseq技术可细分为Original RADseq、GBS、2b-RADseq、ezRAD、ddRADseq等技术方法。ddRADseq技术由于检测准确、成本低、SNP位点丰富等优点得到广泛应用^[12-13]。本研究采用ddRADseq技术对185个马铃薯品种(系)进行了简化基因组测序分析。基于变异位点在各样品中的测序深度均不小于5、在所有样品中的缺失率不超过50%、低频等位基因频率不低于10%、杂合率不超过50%、相对杂合率不超过67%等标准共获得39 038个高质量变异位点,其中SNP位点36 267个,InDel位点2 771个。高质量变异位点为后续马铃薯群体结构分析、SNP标记Panel开发、分子标记辅助选择育种、QTL位点定位奠定了基础。

物种的遗传多样性越高对生存环境的适应性越强,抵御生物胁迫、非生物胁迫的能力越强,易于拓展其生存分布范围。本研究通过对来自不同地方的185份马铃薯品种(系)的遗传多样性分析发现,多态性信息含量 $PIC=0.3107$,处于中等偏低水平;群体期望杂合度 $He=0.3932$,而观测杂合度 $Ho=0.1852$,即 $Ho < He$ 。这一结果表明这些马铃薯品种(系)的遗传多样性较低,即不同品种(系)间的同质性较高,这与前人研究结果较为一致^[14-18]。为了提升马铃薯品种的遗传多样性,在未来的马铃薯育种工作中应加强对马铃薯种质资源的收集和创新,特别是对马铃薯野生种的利用。马铃薯种质资源丰富,包含107个野生种和4个栽培种。原始栽培种和野生种含有众多调控马铃薯晚疫病、疮痂病、青枯病、病

毒病、黑胫病抗性的等位基因,也含有能够调控马铃薯高淀粉和蛋白质含量、低还原糖含量、耐霜冻、休眠期、早熟等品质和农艺性状的优良等位基因^[8]。加速对马铃薯种质资源的收集、评价、研究、创新和利用,有利于突破当前马铃薯育种的瓶颈^[19-22]。

本研究结果表明,供试马铃薯种(系)间的遗传多样性较低,工作中也发现一些品种间表型差异很小,如果不借助分子标记技术,仅从表型差异很难区分。鉴于此,本研究还开发了包含120个马铃薯 SNP-Panel,这些 SNP 标记均匀覆盖马铃薯 12 条染色体,并针对这些 SNP 标记设计了分离目标 DNA 区域的特异引物,以用于后续进行多样本高通量靶向 SNP 标记检测,为后续开展马铃薯亲缘关系鉴定、品种鉴别、群体遗传分析以及分子标记辅助选择育种奠定基础。

4 结论

本研究对收集的 185 份马铃薯品种(系)种质资源进行简化基因组测序,鉴定了 39 038 个变异位点,其中 SNP 位点 36 267 个, InDel 位点 2 771 个;PCA 分析和 Structure 分析均将供试的 185 份马铃薯品种(系)划分为两个亚类;遗传多样性分析表明,群体多态性信息含量 $PIC=0.3107$,期望杂合度 $He=0.3932$,观测杂合度 $Ho=0.1852$,自交系数 $Fis=0.553$,群体遗传多样性较低;开发了覆盖马铃薯 12 条染色体且包含 120 个 SNP 标记的 SNP-Panel。本研究为后续进行马铃薯遗传多样性研究、分子标记辅助选择育种、分子身份证开发、品种鉴定、遗传图谱构建奠定了一定基础。

参考文献 (References) :

[1] 谢从华,柳俊.中国马铃薯引进与传播之辨析[J].华中农业大学学报,2021,40(4):1-7. DOI:10.13300/j.cnki.hnlkxb.2021.04.001.
XIE C H, LIU J. Review of introduction and spread of potato in China [J]. *Journal of Huazhong Agricultural University*, 2021, 40 (4) : 1-7. DOI:10.13300/j.cnki.hnlkxb.2021.04.001.

[2] SPOONER D M, GHISLAIN M, SIMON R, JANSKY S H, GAVRILENKO Y. Systematics, diversity, genetics, and evolution of wild and cultivated potatoes [J]. *Botanical Review*, 2014, 80 (4) :283-383. DOI: 10.1007/s12229-014-9146-y.

[3] TURRILL W B. Studies on the origin of cultivated plants [J]. *Nature*, 1926, 118: 2967. DOI: 10.1038/118392a0.

[4] 陈亚兰,陈鑫.马铃薯遗传育种技术[M].武汉:武汉大学出版社,2015.
CHEN Y L, CHEN X. Potato genetics and breeding technology [M]. Wuhan: Wuhan University Press, 2015.

[5] WANG J. Genetic diversity of the selected 64 potato germplasm revealed by AFLP markers [J]. *Molecular Plant Breeding*, 2011 (2) : 22-29. DOI: 10.5376/mpb.2011.02.0004.

[6] 谷茂,马慧英,薛世明.中国马铃薯栽培史考略[J].西北农业大学学报,1999:27(1):80-84.
GU M, MA H Y, XUE S M. A study on the cultivation history of potato in China [J]. *Journal of Northwest A & F University*, 1999,27(1): 80-84.

[7] 蔡兴奎,谢从华.中国马铃薯发展历史、育种现状及发展建议[J].长江蔬菜,2016(12):30-33.
CAI X K, XIE C H. Development history, current breeding situation and development suggestions of chinese potato [J]. *Journal of Changjiang Vegetables*, 2016 (12) : 30-33.

[8] 孙慧生.马铃薯育种学[M].北京:中国农业出版社,2003.
SUN H S. Potato breeding [M]. Beijing:China Agriculture Press, 2003.

[9] BOTSTEIN D R, WHITE R L, SKOLNICK M H, DAVIS R W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms [J]. *The American Journal of Human Genetics*, 1980,32(3):314-31.

[10] HENNING F, LEE H J, FRANCHINI P, MEYER A. Genetic mapping of horizontal stripes in Lake Victoria cichlid fishes: benefits and pitfalls of using RAD markers for dense linkage mapping [J]. *Molecular Ecology*, 2014, 23 (21) : 5224-5240. DOI: 10.1111/mec.12860

[11] SUTHERLAND B J G, GOSSELIN T, NORMANDEAU E, LAMOTHE M, ISABEL N, AUDET C, BERNATCHEZ L. Salmonid chromosome evolution as revealed by a novel method for comparing RADseq linkage maps [J]. *Genome Biology and Evolution*, 2016,8(12):3600-3617. DOI: 10.1101/039164.

[12] 杨琼,孙娜,郑敏,罗除成,罗光华,方继朝.水稻二化螟和三化螟基因组 ddRADseq 文库的构建[J].环境昆虫学报,2016,38(6):1114-1120.
YANG Q, SUN N, ZHENG M, LUO C C, LUO G H, FANG J C. Construction of genomic ddRADseq libraries of *Chilo suppressalis* and *Scirpophaga incertulas* (Lepidoptera: Pyralidae) [J]. *Journal of Environmental Entomology*, 2016, 38 (6) : 1114-1120.

[13] KONAR A, CHOUDHURY O, BULLIS R, FIEDLER L, KRUSER J M, STEPHENS M T, GAILING O, SCHLARBAUM S, COGGESHALL M V, STATON M E, CARLSON J E, EMRICH S, ROMERO-SEVERSON J. High-quality genetic mapping with ddRADseq in the non-model tree *Quercus rubra* [J]. *BMC Genomics*, 2017,18(1):417. DOI: 10.1186/s12864-017-3765-8.

[14] 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.
- [15] 宋峥, 王崇, 徐颖华, 雷剑, 柴沙沙, 王连军, 杨新笋. 马铃薯地方种质资源的 SSR 遗传多样性分析 [J]. 分子植物育种, 2021:1-21.
SONG Z, WANG C, XU Y H, LEI J, CHAI S S, WANG L J, YANG X S. SSR analysis on genetic diversity for potato germplasm resources [J]. *Molecular Plant Breeding*, 2021:1-21.
- [16] 高玉坤, 崔江慧, 项晓冬, 薛薇, 殷丛培, 任根增, 李峰, 王维东, 王志刚, 魏世林, 梁红凯, 常金华. 65个马铃薯品种(系)指纹图谱构建和遗传多样性分析 [J]. 农业生物技术学报, 2020, 28 (8):1363-1378.
GAO Y K, CUI J H, XIANG X D, XUE W, YIN C P, REN G Z, LI F, WANG W D, WANG Z G, WEI S L, LIANG H K, CHANG J H. Fingerprint construction and genetic diversity analysis for 65 potato (*Solanum tuberosum*) varieties (lines) [J]. *Chinese Journal of Agricultural Biotechnology*, 2020, 28 (8): 1363-1378.
- [17] 王鹏, 李芳弟, 郭天顺, 窦俊焕, 顾炜清, 罗照霞, 齐小东, 杨晨, 赵中梁, 宋怡, 吕汰. SSR 标记马铃薯育成品种的遗传多样性分析 [J]. 中国马铃薯, 2019,33 (5):257-266.
WANG P, LI F D, GUO T S, DOU J H, JI W Q, LUO Z X, QI X D, YANG C, ZHAO Z L, SONG Y, LU T. Analysis of genetic diversity for potato (*Solanum tuberosum* L.) varieties using SSR markers [J]. *Chinese Potato Journal*, 2019,33 (5):257-266.
- [18] 张海雯, 巩楠, 马斯霜, 甘晓燕, 宋玉霞. 基于 SSR 标记的马铃薯种质遗传多样性及群体结构分析 [J]. 分子植物育种, 2020,18 (12):4144-4152. DOI:10.13271/j.mpb.018.004144.
ZHANG H W, GONG L, MA S S, GAN X Y, SONG Y X. Genetic diversity and population structure analysis of potato (*Solanum tuberosum*) germplasm based on SSR markers [J]. *Molecular Plant Breeding*, 2020,18 (12):4144-4152. DOI:10.13271/j.mpb.018.004144.
- [19] HUANG B Q, RUESS H, LIANG Q Q, COLLEONI C, SPOONER D M. Analyses of 202 plastid genomes elucidate the phylogeny of solanum section petota [J]. *Scientific Reports*, 2019 (9):4454. DOI: 10.1038/s41598-019-40790-5.
- [20] GAVRILENKO T, ANTONOVA O, SHUVALOVA A, KRYLOVA E, ALPATYEVA N, SPOONER D, NOVIKOVA L. Genetic diversity and origin of cultivated potatoes based on plastid microsatellite polymorphism [J]. *Genetic Resources and Crop Evolution*, 2013, 60 (7):1997-2015. DOI: 10.1007/s10722-013-9968-1.
- [21] SPOONER D M, RUESS H, ARBIZU C I, RODRIGUEZ F, SOLIS-LEMUS C. Greatly reduced phylogenetic structure in the cultivated potato clade (*Solanum* section petota pro parte) [J]. *American Journal of Botany*, 2018, 105 (1): 60-70. DOI: 10.1002/ajb2.1008.
- [22] SPOONER D M. Species delimitations in plants: lessons learned from potato taxonomy by a practicing taxonomist [J]. *Journal of Systematics and Evolution*, 2016, 54 (3): 191-203. DOI: 10.1111/jse.12203.

(责任编辑 崔建勋)



李小波, 博士, 研究员, 广东省甘薯马铃薯产业技术体系创新团队马铃薯育种岗位专家, 广东省农业科学院作物研究所马铃薯学科团队学术带头人, 华中农业大学和湖南农业大学校外硕士生导师, 兼任广东省马铃薯产业技术创新联盟常务副理事长, 中国作物学会马铃薯专业委员会委员。

主要从事马铃薯遗传育种和节本增效栽培技术研究与应用。先后主持广东省基础研究重大项目子课题、广东省应用型研发专项、国家重点研发计划子课题、广东省马铃薯良种重大科研联合攻关等各级项目 20 多项; 获广东省农业科学院科技奖二等奖 1 项, 广东省农业技术推广奖一等奖 2 项、二等奖 2 项; 制定地方标准 2 项; 培育马铃薯品种 6 个; 获得植物新品种保护权 1 项。培育的马铃薯品种多次被列为广东省和广州市主导品种, 推广面积达 13 万 hm^2 以上, 社会经济效益显著。主编书籍 2 部, 发表科技论文 40 余篇; 申请和获授权国家发明专利 5 件, 实用新型专利 6 件, 软件著作权 4 件。