

# 基于分型测序技术的粒用高粱遗传多样性和群体结构分析

汪 灿, 周棱波, 高 旭, 张国兵, 程 斌, 曹 宁, 丁延庆, 徐 燕, 邵明波, 张立异  
(贵州省农业科学院旱粮研究所, 贵阳 550006)

**摘要:** 了解粒用高粱的遗传多样性和群体结构, 能有效提高粒用高粱新品种的选育效率。本研究利用基因分型测序技术 (GBS, genotyping by sequencing) 对 120 份粒用高粱材料开展了全基因组基因分型, 共获得了 3456 个多态性的 SNP 标记, 其多态性信息含量指数 (PIC, polymorphism information content) 介于 0.013~0.574 之间, 平均值为 0.381。根据 SNP 标记在 120 份高粱材料中的基因分型数据, 计算了材料间的遗传距离, 其变异范围为 0.084~0.613, 平均遗传距离为 0.365。群体进化树分析和主成分分析都将 120 份高粱材料划分为 3 个类群, 类群 1 主要由包括美国材料 MN-3609 在内的亲缘关系较远的高粱材料组成, 类群 2 主要由中国北方的高粱材料组成, 类群 3 主要由中国南方的高粱材料组成。群体结构分析表明, 当  $K=3$  时,  $\Delta K$  取得最大值, 说明 120 份高粱材料可以划分为 3 个类群, 其划分结果与群体进化树分析和主成分分析基本一致。本研究从基因型多样性水平上阐释了粒用高粱的遗传背景和群体结构, 为中国粒用高粱新品种的选育提供了理论依据。

**关键词:** 粒用高粱; 遗传多样性; 群体结构; 分型测序技术

## Genetic Diversity and Population Structure Analysis of Grain-use Sorghum Based on Genotyping by Sequencing Technology

WANG Can, ZHOU Ling-bo, GAO Xu, ZHANG Guo-bing, CHENG Bin, CAO Ning,  
DING Yan-qing, XU Yan, SHAO Ming-bo, ZHANG Li-yi  
(*Institute of Upland Food Crops, Guizhou Academy of Agricultural Sciences, Guiyang 550006*)

**Abstract:** Understanding the genetic diversity and population structure can effectively improve the breeding efficiency of grain sorghum new cultivar. In this study, genotyping by sequencing (GBS) technology was deployed to unlock the genetic diversity in a collection of 120 grain sorghum materials. Within 3456 SNP markers that were generated, the polymorphism information content (PIC) ranged from 0.013 to 0.574, with the average of 0.381, while SNP-based genetic distance value ranged from 0.084 to 0.613, with the average of 0.365. Both phylogenetic tree analysis and principal component analysis indicated that 120 grain sorghum genotypes were clustered into three groups. Group-1 was mainly composed of sorghum materials with a relatively distant relationship, including the USA material MN-3609. Within Group-2 and group-3, these genotypes were mainly collected from northern and south parts of China, respectively. Population structure analysis showed that the maximum  $\Delta K$  was observed for  $K=3$ , suggested three sub-populations detected in 120 grain sorghum materials. This result was consistent with phylogenetic tree analysis and principal component analysis. Taken together, this resulted deciphered the genetic background and population structure of grain sorghum from the genotype diversity level, which will provide theoretical basis in breeding for grain-use sorghum new cultivar in China.

**Key words:** grain sorghum; genetic diversity; population structure; genotyping by sequencing technology

收稿日期: 2018-09-02 修回日期: 2018-10-23 网络出版日期: 2018-11-30

URL: <http://kns.cnki.net/kcms/detail/11.4996.S.20181128.1010.001.html>

第一作者研究方向为高粱分子育种, E-mail: wangc.1989@163.com

通信作者: 张立异, 研究方向为高粱分子育种, E-mail: 570129883@qq.com

**基金项目:** 国家自然科学基金 (31660400); 贵州省科技计划项目 (黔科合企企 [2018] 4005); 中央引导地方科技发展专项资金项目 (黔科中引地 [2018] 4003); 贵州省农业科学院专项资金项目 (黔农科院院专项 [2014] 034 号)

**Foundation project:** National Natural Science Foundation of China (31660400), Science and Technology Program of Guizhou Province (QKHFQ [2018] 4005), Special Funds for the Central Government Guides Local Science and Technology Development (QKZYD [2018] 4003), Special Funds for Guizhou Academy of Agricultural Sciences (QNKYYZX [2014] 034)

高粱 [*Sorghum bicolor* (L.) Moench] 是我国最早栽培的禾谷类作物之一, 根据用途不同, 一般分为甜高粱、饲草高粱和粒用高粱 3 种类型, 在粒用高粱中又将直链淀粉含量在 0~5% 之间的称为糯高粱<sup>[1-2]</sup>。糯高粱是酿造酱香型白酒的主要原料, 对我国区域经济的发展发挥重要作用<sup>[3-4]</sup>。近年来, 糯高粱主栽品种单一, 品种不断退化的现象已经影响了酒用糯高粱原料的供应, 制约了白酒企业的发展<sup>[5-6]</sup>。遗传多样性是作物改良的基础, 杂种优势群的准确划分是选育优良组合和提高作物新品种选育效率的前提<sup>[5]</sup>。高粱的进化和适应不同环境的能力除了取决于种内遗传变异的大小外, 还依赖于种群的遗传结构是否丰富<sup>[7]</sup>。因此, 开展粒用高粱的遗传多样性和群体结构分析, 对粒用高粱的新品种选育和种质资源保护具有重要意义。

随着分子标记技术的快速发展, 针对高粱的遗传多样性和群体结构, 前人已做了不少研究。如 Ayana 等<sup>[8]</sup>利用 RAPD 标记对厄立特里亚和埃塞俄比亚高粱进行遗传多样性分析, 发现来自不同地区的高粱遗传多样性并不丰富, 在聚类分析中不能按照地域进行类群划分。余传涨等<sup>[9]</sup>利用 SSR 标记将 41 个高粱品种划分为粒用高粱和甜高粱 2 个类群, 但无法清晰地划分 2 类高粱之间的遗传界限。Morris 等<sup>[10]</sup>利用 SNP 标记将全世界 970 份高粱资源根据地理来源和表型类型形成不同的群体结构。Ni 等<sup>[11]</sup>用 25 对 SSR 引物检测了 29 份糯高粱种质资源的遗传多样性, 认为糯高粱品种的亲缘关系与地理来源关系不大。高旭等<sup>[5]</sup>利用 33 个 SSR 标记将 156 份全国粒用高粱根据来源地不同划分为 3 个类群。肖松等<sup>[6]</sup>利用 28 对 SSR 引物将 145 个酱香白酒用糯高粱种质资源划分为 2 个类群, 认为贵州酱香型白酒用糯高粱育种资源具有较高的遗传相似性, 遗传多样性水平较低, 遗传背景较为狭窄。

基因分型测序技术 (GBS, genotyping by sequencing) 是一种常见的简化基因组测序技术, 通过分析获得 SNP 信息并进行基因分型, 可根据研究目的灵活调整所需酶切位点的片段数, 从而控制捕获序列的范围, 能够降低基因组的复杂度、操作简便、节约成本, 特别适合大样本量研究<sup>[12]</sup>。SNP 标记在基因组中具有易于检测、数量多、多态性丰富、通量大且分布广泛等特点, 适于快速、规模化筛查, 易于基因分型, 可以显著提高生物信息学分析的效率<sup>[13]</sup>。

尽管目前国内外利用传统分子标记技术在粒用高粱遗传多样性和群体结构的研究上已取得很大进展, 然而对基于 GBS 技术的粒用高粱遗传多样性和群体结构分析的研究鲜见报道。为此, 本研究以 120 份具有代表性的粒用高粱种质为材料, 利用基于 GBS 技术的 SNP 标记进行粒用高粱的遗传多样性和群体结构分析, 旨在揭示中国粒用高粱的遗传多样性和群体结构特征, 为粒用高粱有价值种质的保存和利用、分子标记辅助育种、育种策略的实施以及关联图谱的构建等提供理论依据和实践参考。

## 1 材料与方法

### 1.1 试验材料

选用 120 份具有代表性的粒用高粱高世代种质为研究材料 (表 1), 其中包括贵州材料 22 份、云南材料 19 份、陕西材料 17 份、辽宁材料 15 份、山东材料 10 份、内蒙古材料 9 份、山西材料 8 份、河北材料 8 份、四川材料 7 份、北京材料 4 份、美国材料 1 份。所用材料由中国农业科学院作物科学研究所和贵州省农业科学院旱粮研究所提供。试验用种均为 2015 年自交套袋收获的种子。

### 1.2 材料种植

2016 年高粱材料种植于贵州省旱粮研究所贵阳试验基地 (26°27'N, 106°35'E)。试验地为黄壤, 含全氮 3.68 g/kg、碱解氮 129.84 mg/kg、全磷 2.03 g/kg、有效磷 50.42 mg/kg、全钾 13.77 g/kg、速效钾 598.62 mg/kg、有机质 40.19 g/kg, pH 值 7.5。采用随机区组设计, 3 次重复, 窝距 70 cm, 行距 70 cm, 小区面积 7.35 m<sup>2</sup> (2.1 m × 3.5 m), 每窝留苗 1 株, 种植 3 行, 每行 5 窝, 于 2016 年 4 月 10 日播种, 人工窝播, 播种前施总养分 ≥ 45% 的高浓度硫酸钾复合肥 (含 N 15%、P<sub>2</sub>O<sub>5</sub> 15%、K<sub>2</sub>O 15%) 300 kg/hm<sup>2</sup> 作为种肥。常规管理。

### 1.3 DNA 提取及基因分型测序

2016 年采集 120 份高粱材料的 4 叶期嫩叶, 参照余传涨等<sup>[9]</sup>的方法提取 DNA, 在 0.8% 的琼脂糖凝胶上 100 V 电泳 40 min 检测 DNA 纯度, 同时用 Qubit® 2.0 荧光测定计 (Invitrogen, Carlsbad, USA) 检测 DNA 浓度。将质检后的样品 DNA 送至华大基因, 首先用限制性内切酶 *EcoR* I 对基因组进行酶切, 酶切后利用 T4 DNA 连接酶 (NEB) 加上带有 barcode 的接头, 对每个样品进行扩增, 然后进行样

表 1 本研究所用的高粱材料

Table 1 Sorghum materials used in this study

编号 Number	名称 Name	来源 Origin	编号 Number	名称 Name	来源 Origin	编号 Number	名称 Name	来源 Origin
GG001	云县高粱	中国云南	GG041	小缨子	中国贵州	GG081	鲁高 369	中国山东
GG002	宾川高粱	中国云南	GG042	红缨子	中国贵州	GG082	鲁高 370	中国山东
GG003	新平高粱	中国云南	GG043	河农 16	中国河北	GG083	鲁高 657	中国山东
GG004	墨江高粱	中国云南	GG044	黔高 7 号	中国贵州	GG084	鲁高 97	中国山东
GG005	宁陕高粱	中国陕西	GG045	MN-3609	美国	GG085	鲁高 80	中国山东
GG006	水 512	中国四川	GG046	京高 426	中国北京	GG086	辽高 43	中国辽宁
GG007	芽糯子	中国山东	GG047	京高 658	中国北京	GG087	阳高高粱	中国山西
GG008	水 421	中国四川	GG048	冀高 2 号	中国河北	GG088	定襄高粱	中国山西
GG009	水 507	中国四川	GG049	冀高 1131	中国河北	GG089	陵川高粱	中国山西
GG010	元谋高粱	中国云南	GG050	冀高 417	中国河北	GG090	大同高粱	中国山西
GG011	禄丰高粱	中国云南	GG051	中 733	中国河北	GG091	广饶高粱	中国山东
GG012	崇明高粱	中国云南	GG052	冀高 367	中国河北	GG092	陕高 206	中国陕西
GG013	水 723	中国四川	GG053	冀高 1135	中国河北	GG093	平利高粱	中国陕西
GG014	水 639	中国四川	GG054	冀高 1134	中国河北	GG094	商南高粱	中国陕西
GG015	水 712	中国四川	GG055	晋高 566	中国山西	GG095	陕高 227	中国陕西
GG016	水 724	中国四川	GG056	晋高 328	中国山西	GG096	安康高粱	中国陕西
GG017	黔高 40	中国贵州	GG057	晋高 285	中国山西	GG097	陕高 189	中国陕西
GG018	辽宁红糯	中国辽宁	GG058	晋高 702	中国山西	GG098	辽高 3008	中国辽宁
GG019	黔高 47	中国贵州	GG059	辽高 32	中国辽宁	GG099	陕高 222	中国陕西
GG020	黔高 1511	中国贵州	GG060	内高 390	中国内蒙古	GG100	石泉高粱	中国陕西
GG021	黔高 1522	中国贵州	GG061	永 123	中国内蒙古	GG101	镇坪高粱	中国陕西
GG022	黔高 1529	中国贵州	GG062	内高 242	中国内蒙古	GG102	陕高 2491	中国陕西
GG023	黔高 1535	中国贵州	GG063	内高 371	中国内蒙古	GG103	陕高 279	中国陕西
GG024	黔高 1537	中国贵州	GG064	内高 524	中国内蒙古	GG104	柞水高粱	中国陕西
GG025	黔高 1538	中国贵州	GG065	内高 433	中国内蒙古	GG105	陕高 273	中国陕西
GG026	黔高 1540	中国贵州	GG066	内高 686	中国内蒙古	GG106	陕高 84-300	中国陕西
GG027	黔高 1545	中国贵州	GG067	内高 249	中国内蒙古	GG107	昭通高粱	中国云南
GG028	黔高 1547	中国贵州	GG068	内高 254	中国内蒙古	GG108	云高 R20	中国云南
GG029	黔高 1549	中国贵州	GG069	辽高 1234	中国辽宁	GG109	辽高 23	中国辽宁
GG030	黔高 1550	中国贵州	GG070	辽高 241	中国辽宁	GG110	砚山高粱	中国云南
GG031	黔高 1599	中国贵州	GG071	辽高 31	中国辽宁	GG111	云高 R32	中国云南
GG032	黔高 1734	中国贵州	GG072	辽高 1102	中国辽宁	GG112	云高 R23	中国云南
GG033	黔高 9511	中国贵州	GG073	辽高 240	中国辽宁	GG113	云高 R21	中国云南
GG034	陕糯 9776	中国陕西	GG074	辽高 1044	中国辽宁	GG114	蒙自高粱	中国云南
GG035	陕糯 9794	中国陕西	GG075	辽高 1030	中国辽宁	GG115	镇雄高粱	中国云南
GG036	京高 565	中国北京	GG076	辽高 1029	中国辽宁	GG116	云高 R09	中国云南
GG037	赤水红梁 2 号	中国贵州	GG077	鲁高 221	中国山东	GG117	宣威高粱	中国云南
GG038	罗平高粱	中国贵州	GG078	辽高 239	中国辽宁	GG118	陆良高粱	中国云南
GG039	京高 459	中国北京	GG079	鲁高 214	中国山东	GG119	华宁高粱	中国云南
GG040	茅红糯 1 号	中国贵州	GG080	鲁高 351	中国山东	GG120	辽高 6 号	中国辽宁

品混合,电泳回收 350~400 bp 区间的 DNA 条带,割取的片段进行纯化,纯化后的产物用于测序,测序反应在 Illumina HiSeq4000 测序平台进行双末端 PE150 的测序。

将测序得到的原始图像数据经 base calling 转化为序列数据,过滤掉含有接头序列的 reads,去除单端测序 read 中 N 的含量超过该条 read 长度比例 10%,且单端测序 read 中含有的低质量 ( $Q \leq 5$ ) 碱基数超过该条 read 长度比例 50% 的序列。然后将有效的高质量测序数据通过 BWA 软件<sup>[14]</sup>比对到 BTx623 参考基因组上 ([https://phytozome.jgi.doe.gov/pz/portal.html#linfo?alias=Org\\_SbicolorRio\\_er](https://phytozome.jgi.doe.gov/pz/portal.html#linfo?alias=Org_SbicolorRio_er))。比对结果经 Samtools 软件<sup>[15]</sup>去除重复并进行 SNP 的检测,利用贝叶斯模型检测群体中的多态性位点。

#### 1.4 数据统计分析

每个 SNP 标记的多态性信息含量指数 (PIC, polymorphism information content) 通过如下公式计算:

$$PIC = 1 - \sum (P_i)^2$$

式中,  $P_i$  为群体中含有第  $i$  个等位基因的比例。利用 Tassel 5.0 软件<sup>[15]</sup>根据个体 SNPs 的检测值计算群体间的遗传距离,2 个个体  $i$  和  $j$  之间的  $p$ -距离通过如下公式计算:

$$D_{ij} = \frac{1}{L} \sum_{l=1}^L d_{ij}$$

式中,  $L$  表示高质量 SNPs 区域长度,  $d_{ij}$  表示 2 个个体为不同基因型值的赋值表。运用 TreeBest (<http://treesoft.sourceforge.net/treebest.shtml>) 软件计算距离矩阵,以此为基础,通过邻近法 (neighbor-joining method) 构建系统进化树,引导值 (bootstrap values) 经过达 1000 次计算获得。利用 Tassel 5.0 软件分别进行基于 SNP 标记、来源和表型的主成分分析。利用 Structure 2.3.4 软件<sup>[16]</sup>的混合模型对 120 份高粱材料的群体结构进行分析,类群数目  $K$  的取值范围为 1~10,将模拟参数迭代 (length of burn-in period) 设为 50000 次,蒙特卡罗迭代 (MCMC, markov chain monte carlo) 设为 100000 次,5 次重复,最后根据后验概率值的变化速率 ( $\Delta K$ ) 来确定关联群体的最佳类群数目。

## 2 结果与分析

### 2.1 测序质量

120 份高粱材料的 GBS 分析共获得了 44.94 Gb 的总测序数据量,过滤后高质量的 clean data 数据

量为 43.67 Gb,平均每个样本 372.65 Mb。测序质量高 ( $Q20 \geq 97.28\%$ ,  $Q30 \geq 92.71\%$ ), GC 分布正常,120 份高粱样本都没有被污染,此次建库测序成功。群体样本与 BTx623 参考基因组的平均比率为 85.52%,序列平均覆盖度为 10.32% (至少有 1 个碱基的覆盖),至少有 4 个碱基覆盖的平均覆盖度为 4.96%。Samtools 软件检测共获得了 14133 个 SNP 位点,经过条件为 dp3、Miss0.2、maf0.05 的过滤后,最后获得了 3456 个高质量的 SNP 标记用于后续分析。这些标记表现出不同水平的遗传多样性,其 PIC 值介于 0.013~0.574 之间,平均值为 0.381。此外,这些 SNP 标记不同程度地全部覆盖在高粱的 10 条染色体上,平均 200 kb 含有 1 个 SNP,能够保障全基因组上的性状位点都能被检测到 (图 1)。

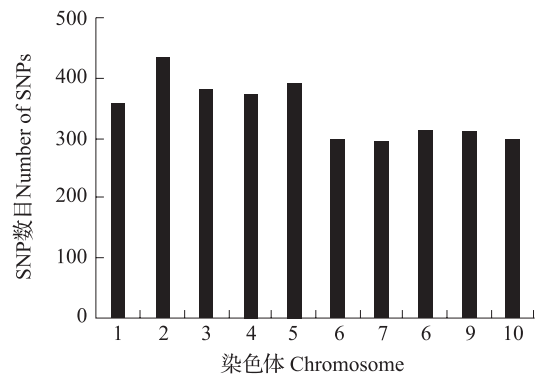


图 1 GBS-SNPs 标记在高粱基因组上的分布  
Fig.1 Distribution of GBS-SNPs on the whole sorghum genome

### 2.2 遗传多样性分析

根据 3456 个 SNP 标记的基因分型数据,计算了 120 份高粱材料间的遗传距离,其变异范围为 0.084~0.613,平均遗传距离为 0.365。其中,我国云南材料云县高粱与蒙自高粱的遗传距离最小,为 0.084,说明二者之间亲缘关系最近。我国陕西材料宁陕高粱与河北材料河农 16 的遗传距离最大,为 0.613,说明二者之间的亲缘关系最远。此外,对于来源贵州的 17 份高粱材料,其变异范围为 0.137~0.538,平均遗传距离为 0.353。

根据群体进化树分析结果 (图 2),结合来源地的不同,在遗传距离 0.1 处将 120 份高粱材料划分为 3 个类群,其中类群 I 主要由 8 份高粱材料组成,包括了美国材料 MN-3609、2 份南方高粱材料 (元谋高粱和黔高 1734) 和 5 份北方高粱材料 (宁陕高粱、辽宁红糯、陕糯 9776、河农 16 和辽高 32)。类群

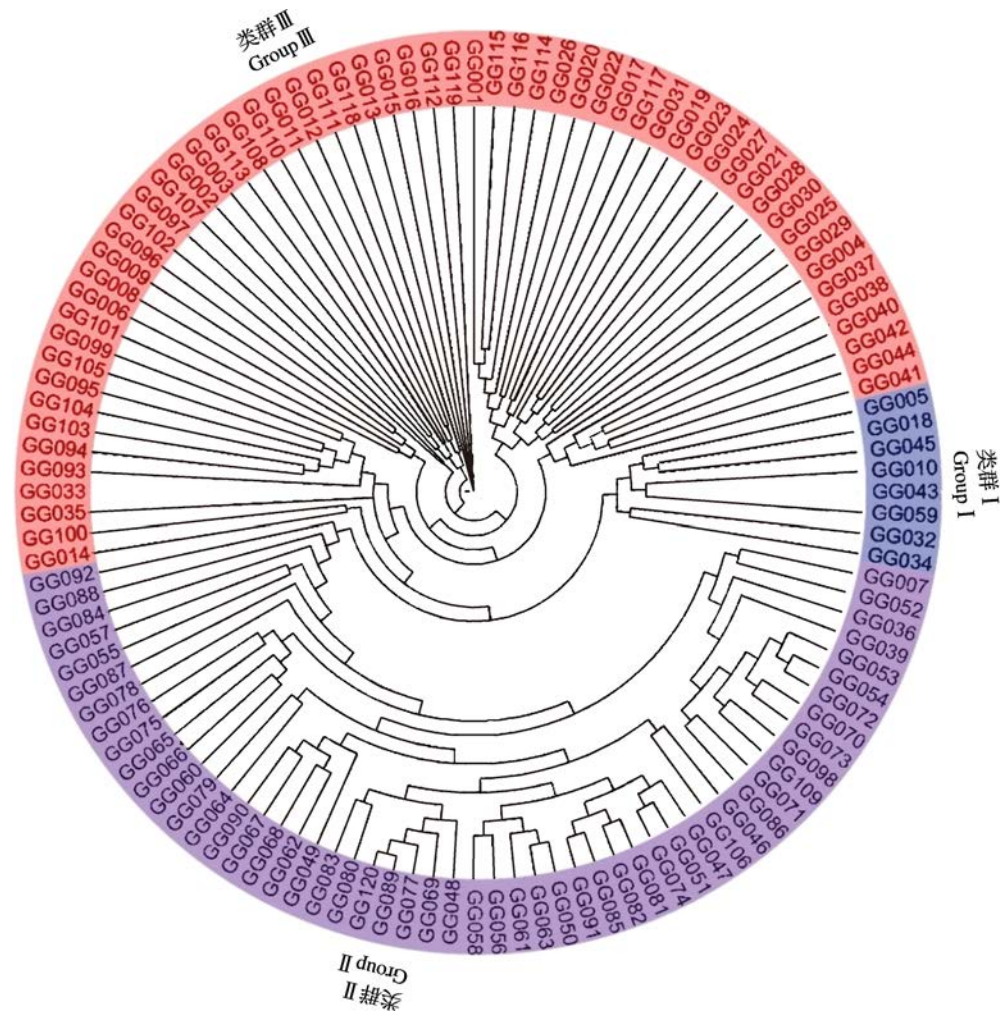


图 2 120 份高粱材料的群体进化树分析

Fig.2 Phylogenetic tree analysis of 120 sorghum materials

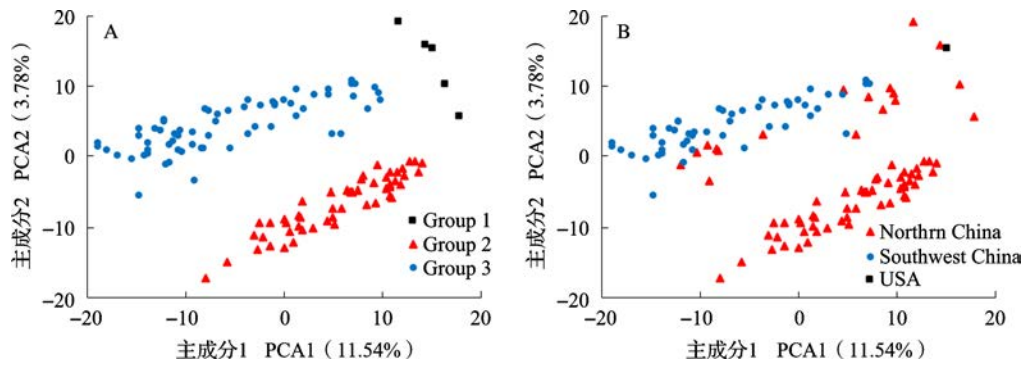
II 包含了 53 份高粱材料, 主要由来源于北方的高粱材料组成。类群 III 包含 59 份高粱材料, 主要由来源于南方的高粱组成。

### 2.3 群体结构分析

基于 SNP 标记的主成分分析结果表明(图 3A), PC1 和 PC2 解释的遗传方差分别为 11.54% 和 3.78%, 120 份高粱群体在 PC1 方向上较为分散, 而在 PC2 方向上聚集得更为紧密。这些高粱材料聚集形成 3 个类群, 类群 1 主要由包括美国材料 MN-3609 在内的 5 份高粱材料组成, 与其他 2 个类群显示较远的亲缘关系, 此类群中的 5 份高粱材料全部包含于群体进化树分析中的类群 I 中, 除了 3 份高粱材料(宁陕高粱、元谋高粱、陕糯 9776)。类群 2 由来源于中国北方的 51 份高粱材料组成, 与群体进化树分析的类群 II 几乎一致, 除了 2 份高粱材料(芽糯子、鲁高 221)。类群 3 主要由来源于中国南方地区的 64 份高粱材料组成, 与群体进化树分析中

类群 III 的组成相比较, 有 6 个高粱材料不一致, 它们分别是水 639、黔高 1535、黔高 1550、黔高 1599、黔高 9511、大同高粱。基于来源地的主成分分析也显示(图 3B), 这些高粱材料主要是根据来源地的不同聚集在一起, 但是也有少量的北方高粱混杂在南方高粱中(类群 3), 美国材料 MN-3609 则和 4 份北方高粱聚集在一起。

群体结构分析显示,  $\Delta K$  随  $K$  值的增大呈双峰曲线变化, 其峰值出现在  $K=3$  和  $K=6$  处, 且最大  $\Delta K$  值出现在  $K=3$  处(图 4A), 结合群体进化树分析和主成分分析结果表明, 120 份高粱材料均被划分为 3 个类群, 所以该关联群体的最佳类群数目为 3, 其  $K=3$  时的  $Q$  值分布见图 4B。通过  $Q$  值分析, 发现其中 64 份高粱材料的  $Q$  值大于或等于 0.6, 推测遗传组分相对单一, 属 3 个类群中的其中一个, 其余 56 份高粱材料的  $Q$  值小于 0.6, 主要集中在另外 2 个类群中, 遗传组分较为复杂。

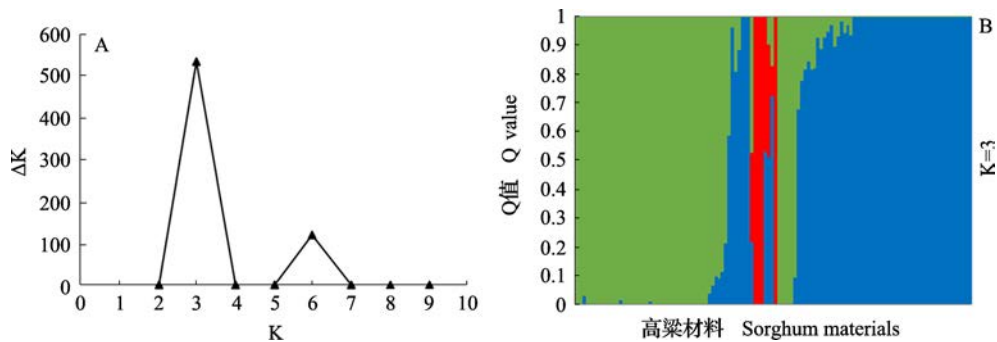


A: 基于 SNP 标记; B: 基于来源

A: based on SNP marker, B: based on origin

图3 120份高粱材料的主成分分析

Fig.3 Principal component analysis of 120 sorghum materials



A: K 取不同值时  $\Delta K$  的对应分布; B: K=3 时的 Q 值分布

A: estimated  $\Delta K$  between successive K from 1 to 9, B: Q-value bar plot in K=3

图4 120份高粱材料的群体结构分析

Fig.4 Population structure analysis of 120 sorghum materials

### 3 讨论

GBS 技术作为第二代深度测序基础上发展起来的简化基因组测序技术,通过采用酶切加标签的方法,使多样本高通量平行测序得以实现,这不仅大大降低了基因测序的成本,也使对大样本全基因组的基因分型成为可能<sup>[17-18]</sup>。GBS 技术可以根据研究目的灵活调整所需捕获酶切位点的片段数,从而控制捕获序列的范围,获得高密度的 SNP 标记,它作为基因分型的重要手段,已经在遗传图谱构建<sup>[19]</sup>、种质鉴定<sup>[20]</sup>、基因多样性研究<sup>[21]</sup>及品种识别<sup>[22]</sup>等领域得到广泛的应用。本研究利用 GBS 技术对 120 份粒用高粱材料进行测序,共获得 44.67 Gb 的测序数据,平均每个样本测序数据为 372.65 Mb,测序质量较高( $Q_{20} \geq 97.28\%$ ,  $Q_{30} \geq 92.71\%$ ),GC 分布正常,且对 120 份高粱样本都没有造成污染,说明此次建库测序成功。此外,该群体样本与高粱基因组的平均比率为 85.52%,说明本试验样本与高粱基因组具有较高的相似性。

因此,利用 GBS 技术对高粱资源进行基因分型有助于高粱遗传多样性和群体结构的分析。在本研究中,利用 GBS 技术对 120 份高粱材料进行基因分型,共获得 3456 个多态性的 SNP 位点,其  $PIC$  值介于 0.013~0.574 之间,平均值为 0.381,说明这些 SNP 标记在 120 份高粱材料中多态性较为丰富。

我国高粱种质资源丰富,但来源复杂,缺乏系统、准确和完善的普查资料,导致很多材料亲缘关系模糊。因此,加强我国高粱种质资源遗传背景的研究具有重要意义。拓宽高粱遗传基础是提高高粱产量和品质的前提<sup>[23]</sup>,为了解我国粒用高粱材料的遗传多样性和群体结构,本研究利用 3456 个 SNP 标记的基因型数据计算了 120 份高粱材料间的遗传距离,其变异范围为 0.084~0.613,平均遗传距离为 0.365,遗传多样性较为丰富,这可能是由于供试材料数量较多,且材料的地理来源较为广泛,能较好地反映我国粒用高粱材料的遗传多样性<sup>[24-25]</sup>。本研究通过群体进化树分析、主成分分析和群体结构

分析调查了 120 份高粱材料的遗传多样性和群体结构,其结果基本一致,但是仍有几个材料在两种方法中被划分到不同的类群中,这可能是由于两种分析方法的计算依据不同,导致一些亲缘关系比较复杂的高粱材料被划分到不同的类群<sup>[26]</sup>。群体进化树分析、主成分分析和群体结构分析都将 120 份高粱材料划分为 3 个类群,类群 1 主要由包括美国材料 MN-3609 在内的亲缘关系较远的高粱材料组成,类群 2 主要由中国北方的高粱材料组成,类群 3 主要由中国南方的高粱材料组成。这在其他的研究中也有相似报道,如高旭等<sup>[5]</sup>基于 SSR 标记根据地理来源将 156 份高粱材料划分为 3 个类群,类群 1 由西南地区的糯高粱材料组成,类群 2 由东北和华北地区的高粱材料组成,类群 3 由剩余地区的高粱材料组成。值得一提的是,本研究将美国高粱 MN-3609 和中国河北的高粱品种河农 16 聚集在类群 1 中,这可能是由于河农 16 是以美国高粱 NK222 和中国优质地方品种象牙白为亲本的杂交高粱品种,而 MN-3609 和 NK222 的亲缘关系较近所致。类群 3 主要由中国南方地区的糯高粱材料组成,这可能是由于近年来贵州、四川、云南等地的一些高粱材料相互引种,遗传差异逐渐缩小所致。

针对高粱的遗传多样性和群体结构,前人利用传统分子标记技术已开展了不少研究,但这些研究在不同程度上还存在一定的局限性。如余传涨等<sup>[9]</sup>利用 SSR 标记技术未能清晰地划分粒用高粱和甜高粱之间的遗传界限;张晗等<sup>[27]</sup>利用 SSR 分子标记技术未能将我国高粱材料按来源或类型进行划分。本研究利用 GBS 技术获得的 SNP 标记一方面说明了中国粒用高粱材料具有丰富的遗传多样性,另一方面根据地理来源清晰地将 120 份高粱材料划分为 3 个不同的类群。因此,利用 GBS 技术对高粱资源进行基因分型更有助于材料的准确划分和科学管理。

#### 参考文献

- [1] Almodares A, Taher R, Chung I M, Fathi M. The effect of nitrogen and potassium fertilizers on growth parameters and carbohydrate contents of sweet sorghum cultivars. *Journal of Environmental Biology*, 2008, 29: 849-852
- [2] Han L P, Steinberger Y, Zhao Y L, Xie G X. Accumulation and partitioning of nitrogen, phosphorus and potassium in different varieties of sweet sorghum. *Field Crops Research*, 2011, 120: 230-240
- [3] 熊先勤, 陈瑞祥, 杨非, 刘正书, 周玉锋. 贵州酒用高粱种质资源考察及鉴定. *山地农业生物学报*, 2003, 22(2): 117-121
- [4] Xiong X Q, Chen R X, Yang F, Liu Z S, Zhou Y F. Inspection and identification of germplasm resources of wine-making sorghum in Guizhou. *Journal of Mountain Agriculture and Biology*, 2003, 22(2): 117-121
- [5] 彭秋. 贵州酿造高粱育种现状、问题及对策. *种子*, 2011, 30(2): 68-71
- [6] Peng Q. Status and problem and countermeasures of brewing sorghum in Guizhou. *Seed*, 2011, 30(2): 68-71
- [7] 高旭, 周棱波, 张国兵, 邵明波, 张立异. 基于 SSR 标记的粒用高粱资源遗传多样性及群体结构. *贵州农业科学*, 2016, 44(9): 13-19
- [8] Gao X, Zhou L B, Zhang G B, Shao M B, Zhang L Y. Genetic diversity and population structure of grain sorghum germplasm resources based on SSR marker. *Guizhou Agricultural Sciences*, 2016, 44(9): 13-19
- [9] 肖松, 周棱波, 张国兵, 邵明波, 乙引, 张立异. 酱香型白酒用糯高粱种质遗传多样性分析. *江苏农业科学*, 2016, 44(4): 45-49
- [10] Xiao S, Zhou L B, Zhang G B, Shao M B, Yi Y, Zhang L Y. Genetic diversity analysis of Maotai-flavor liquor-making waxy sorghum. *Jiangsu Agricultural Sciences*, 2016, 44(4): 45-49
- [11] Wang L M, Jiao S J, Jiang Y X, Yan H D, Su D F, Sun G Q, Yan X F, Sun L F. Genetic structure analysis of sorghum parent lines based on SSR markers. *Cereal Research Communications*, 2013, 41: 359-365
- [12] Ayana A, Bryngelsson T, Bekele E. Genetic variation of Ethiopian and Eritrean sorghum [*Sorghum bicolor* (L.) Moench] germplasm assessed by random amplified polymorphic DNA (RAPD). *Genetic Resources and Crop Evolution*, 2000, 47: 471-482
- [13] 余传涨, 翟国伟, 邹桂花, 陶跃之, 王华. 41 个高粱品种遗传多样性的 SSR 检测. *江苏农业学报*, 2010, 26(2): 248-253
- [14] Yu C Z, Zhai G W, Zou G H, Tao Y Z, Wang H. Assessment of genetic diversity among 41 sorghum varieties using SSR markers. *Jiangsu Journal of Agricultural Sciences*, 2010, 26(2): 248-253
- [15] Morris G P, Ramub P, Deshpandeb S P, Thomas Hash C, Shah T, Upadhyaya H D, Riera-Lizarazu O, Brown P J, Acharya C B, Mitchell S E, Harriman J, Glaubitz J C, Buckler E S, Kresovich S. Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110: 453-458
- [16] Ni X L, Zhao G L, Liu T P, Long W J, Hu J L, Ding G X. Genetic diversity analysis of glutinous sorghum germplasm resources based on SSR markers. *Agricultural Science and Technology*, 2016, 17: 499-504
- [17] Elshire R J, Glaubitz J C, Sun Q, Poland J A, Kawamoto K, Buckler E S, Mitchell S E. A robust, simple genotyping by sequencing (GBS) approach for high diversity species. *PLoS One*, 2011, 6: e19379
- [18] Palaisa K A, Morgante M, Willams M, Rafalski A. Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. *Plant Cell*, 2003, 15: 1795-1806
- [19] Li H, Durbin R. Fast and accurate short read alignment with

- Burrows-Wheeler transform. *Bioinformatics*, 2009, 25: 1754-1760
- [ 15 ] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and SAMtools. *Bioinformatics*, 2009, 25: 2078-2079
- [ 16 ] Bradbury P J, Zhang Z, Kroon D E, Casstevens T M, Ramdoss Y, Buckler E S. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, 2007, 23: 2633-2635
- [ 17 ] Pritchard J K, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*, 2000, 155: 945-959
- [ 18 ] Sonah H, Bastien M, Iquira E, Tardivel A, L egar e G, Boyle B, Normandeau  , Laroche J, Larose S, Jean M, Belzile F. An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. *PLoS One*, 2013, 8: e54603
- [ 19 ] Guajardo V, Sol s S, Sagredo B, Gainza F, Mu oz C, Gasic K, Hinrichsen P. Construction of high density sweet cherry (*Prunus avium* L.) linkage maps using microsatellite markers and SNPs detected by genotyping-by-sequencing (GBS). *PLoS One*, 2015, 10: e0127750
- [ 20 ] Wong M M L, Gujaria-Verma N, Ramsay L, Yuan H Y, Caron C, Diapari M, Vandenberg A, Bett K E. Classification and characterization of species within the genus *lens* using genotyping-by-sequencing (GBS). *PLoS One*, 2015, 10: e0122025
- [ 21 ] Lin M, Cai S B, Wang S, Liu S B, Zhang G R, Bai G H. Genotyping-by-sequencing (GBS) identified SNP tightly linked to QTL for pre-harvest sprouting resistance. *Theoretical and Applied Genetics*, 2015, 128: 1385-1395
- [ 22 ] Lombardi M, Materne M, Cogan N O I, Rodda M, Daetwyler H D, Slater A T, Forster J W, Kaur S. Assessment of genetic variation within a global collection of lentil (*Lens culinaris* Medik.) cultivars and landraces using SNP markers. *BMC Genetics*, 2014, 15: 150
- [ 23 ] 王黎明, 焦少杰, 姜艳喜, 严洪冬, 苏德峰, 孙广全. 利用分子标记分析高粱的遗传多样性及其在种质创新中的应用. *植物遗传资源学报*, 2015, 16(2): 288-293
- Wang L M, Jiao S J, Jiang Y X, Yan H D, Su D F, Sun G Q. Genetic diversity analysis on sorghum using molecular markers and its application in germplasm enhancement. *Journal of Plant Genetic Resources*, 2015, 16(2): 288-293
- [ 24 ] 罗凯, 卢会翔, 吴正丹, 吴雪莉, 尹旺, 唐道彬, 王季春, 张凯. 中国西南地区甘薯主要育种亲本的遗传多样性及群体结构分析. *中国农业科学*, 2016, 49(3): 593-608
- Luo K, Lu H X, Wu Z D, Wu X L, Yin W, Tang D B, Wang J C, Zhang K. Genetic diversity and population structure analysis of main sweet potato breeding parents in Southwest China. *Scientia Agricultura Sinica*, 2016, 49(3): 593-608
- [ 25 ] 张晨, 云岚, 李珍, 王俊, 郭宏宇, 盛誉, 石子英, 徐学宝. 新麦草种质的SSR遗传多样性及群体结构分析. *植物遗传资源学报*, 2019, 20(1): 48-59
- Zhang C, Yun L, Li Z, Wang J, Guo H Y, Sheng Y, Shi Z Y, Xu X B. Genetic diversity and structure analysis in *Psathyrostachys Nevski* population using SSR markers. *Journal of Plant Genetic Resources*, 2019, 20(1): 48-59
- [ 26 ] 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. *中国农业科学*, 2017, 50(9): 1666-1673
- Wang X K, Jiang D, Sun Z Z. Study on phylogeny of 240 mandarin accessions with genotyping-by-sequencing technology. *Scientia Agricultura Sinica*, 2017, 50(9): 1666-1673
- [ 27 ] 张晗, 王健成, 王东建, 姚凤霞, 许金芳, 宋国安, 管延安, 李汝玉. 中国高粱地方品种遗传多样性评价及中、外高粱遗传变异水平比较. *作物学报*, 2011, 37(2): 224-234
- Zhang H, Wang J C, Wang D J, Yao F X, Xu J F, Song G A, Guan Y A, Li R Y. Assessment of genetic diversity in Chinese sorghum landraces using SSR markers as compared with foreign accessions. *Acat Agronomica Sinica*, 2011, 37(2): 224-234