

# 基于表型与分子数据的斑茅核心种质构建

毛 钧<sup>1</sup>, 刘新龙<sup>1</sup>, 苏火生<sup>1</sup>, 陆 鑫<sup>1</sup>, 林秀琴<sup>1</sup>, 蔡 青<sup>1,2</sup>, 范源洪<sup>1</sup>

(<sup>1</sup>云南省农业科学院甘蔗研究所/云南省甘蔗遗传改良重点实验室, 开远 661699;

<sup>2</sup>云南省农业科学院生物技术与种质资源研究所, 昆明 650223)

**摘要:**斑茅 (*Saccharum arundinaceum* L.) 是重要的“甘蔗复合群”野生资源, 在我国分布广泛, 但目前尚未在遗传育种研究中得到很好的利用。本研究以国家甘蔗种质资源圃保存的来自全国 9 省区的 147 份斑茅无性系为原始种质, 首先基于表型数据筛选 58 份材料构建了斑茅初级核心种质, 然后基于 SSR 和 AFLP 分子标记数据筛选 16 份材料构建斑茅微核心种质, 分别占原始种质的 39.46% 和 10.88%。通过表型与分子遗传多样性参数的比较分析, 验证了斑茅核心种质具有较好代表性。最后对甘蔗种质资源和遗传育种研究中斑茅核心种质的构建方法和利用策略等问题进行了探讨, 以期“甘蔗复合群”野生类群核心种质的构建及创新利用提供参考, 促进资源研究工作重心由数量保存型向创新利用型转变。

**关键词:**斑茅; 核心种质; 表型; SSR; AFLP

## Constructing Core Collection of *Saccharum arundinaceum* L. Based on Phenotype and Molecular Markers

MAO Jun<sup>1</sup>, LIU Xin-Long<sup>1</sup>, SU Huo-Sheng<sup>1</sup>, LU Xin<sup>1</sup>, LIN Xiu-Qin, CAI Qing<sup>1,2</sup>, FAN Yuan-Hong<sup>1</sup>

(<sup>1</sup>Sugarcane Research Institute, Yunnan Academy of Agricultural Sciences/Yunnan Key Laboratory of Sugarcane Genetic Improvement, Kaiyuan 661699; <sup>2</sup>Biology & Genetic Germplasm Institute, Yunnan Academy of Agricultural Sciences, Kunming 650223)

**Abstract:** As one of the important wild germplasm resources in *Saccharum* complex, *Saccharum arundinaceum* L. widely distributed in China, but has being used inefficiently in sugarcane genetic breeding research. In this study, we gathered 147 clones of *S. arundinaceum* L. from 9 provinces in China and planted in National Germplasm Repository of Sugarcane as original materials. Firstly, 58 clones were selected based on phenotypic data to construct primary core collection and then 16 clones were selected based on SSR and AFLP molecular markers to construct mini core collection, which occupied 39.46% and 10.88% of original materials, respectively. Through comparative analysis of phenotypic and molecular genetic diversity parameters, the core collection of *S. arundinaceum* L. was proved to be a good representative sample. The construction approach and effective utilization of the core collection of *S. arundinaceum* L. in sugarcane germplasm resources and genetic breeding research were discussed. This paper provides a reference for construction and application of core collection from wild populations of *Saccharum* complex, and promotes change of the future research from quantitative conservation to innovative utilization.

**Key words:** *Saccharum arundinaceum* L.; core collection; phenotype; SSR; AFLP

在“甘蔗复合群 (*Saccharum* complex)”的众多成员中, 斑茅 (*Saccharum arundinaceum* L.) 由于具有丛生性好、萌芽力强、分蘖多、生势强、宿根性好、适应性广、抗逆性和抗病虫性强等优异性状, 已成为

收稿日期: 2015-07-08 修回日期: 2015-07-15 网络出版日期: 2016-06-12

URL: <http://www.cnki.net/kcms/detail/11.4996.S.20160612.1120.006.html>

基金项目: 国家科技支撑计划课题 (2013BAD01B00); 云南省应用基础研究计划重点项目 (2006C0013Z); 农业部农作物种质资源保护项目 (2016NWB017); 科技部国家科技基础条件平台 (NICGR2016-044)

第一作者研究方向为甘蔗种质资源鉴定评价与创新利用。E-mail: nj\_raincat@163.com

通信作者: 范源洪, 研究方向为甘蔗种质资源与遗传育种。E-mail: fyhsri@vip.sohu.com

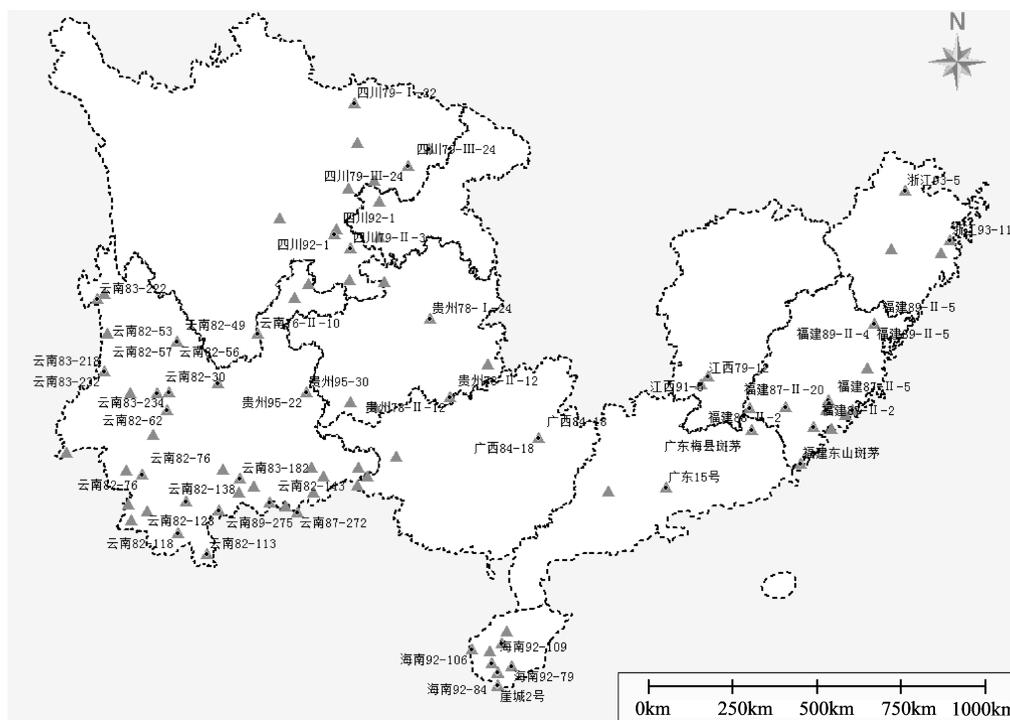
现代甘蔗种质创新和遗传育种研究重点关注的野生资源之一<sup>[1]</sup>。近年来,随着甘蔗野生种质资源的广泛采集,所保育的资源数量急剧增加,若对所有资源开展精细鉴定评价和杂交利用研究,不但费时费力,而且利用效率非常低。核心种质(core collection)概念的提出<sup>[2]</sup>为种质资源的高效保护利用和深入鉴定评价开辟了新的思路。核心种质强调以最少的资源数量最大程度地代表整个资源的遗传多样性<sup>[3]</sup>,因此在核心种质构建过程中取样策略的选择显得尤为关键。国内外关于核心种质构建方法的研究很多,一般来说首先可根据形态表型数据筛选建立初级核心库;然后在初级库的基础上,利用分子标记数据进行DNA水平的遗传冗余压缩,从而得到更具遗传代表性的微核心种质。分子水平构建微核心种质的常用标记有RAPD、AFLP和SSR等。许多一年生作物如水稻<sup>[4-5]</sup>、大豆<sup>[6-7]</sup>和小麦<sup>[8-9]</sup>的核心种质构建工作都取得了良好的效果。关于“甘蔗复合群”遗传育种相关资源的核心种质构建研究,目前已开展了甘蔗热带种<sup>[10]</sup>、杂交品种<sup>[11-12]</sup>、割手密<sup>[13-15]</sup>、

五节芒<sup>[16]</sup>和中国芒<sup>[17]</sup>等类群的核心种质构建与遗传多样性分析。本研究以采自我国斑茅主要分布区的147份斑茅野生资源为原始材料,首先基于表型数据的分析比较构建了斑茅初级核心种质;然后在初级库基础上利用SSR和AFLP标记进行遗传冗余压缩筛选微核心种质,并验证核心种质的代表性;最后对斑茅核心种质在甘蔗遗传育种研究中的利用方式和存在问题进行了探讨。研究旨在为“甘蔗复合群”相关类群核心种质的构建及创新利用提供参考,促进甘蔗野生资源保育和利用工作由数量保存型向质量创新型转变。

## 1 材料与方法

### 1.1 试验材料

147份斑茅材料均来自国家甘蔗种质资源圃,其中59份采自云南,19份采自贵州,18份采自海南,15份采自福建,14份采自四川,10份采自江西,5份采自广西,4份采自浙江,3份采自广东。试验材料采样点分布见图1。



图中数字为斑茅样品采集编号

The numbers on map is the collection series numbers of *E. arundinaceus* samples

图1 斑茅参试样品采样分布图

Fig. 1 Distribution of origins for the *S. arundinaceus* L. samples in this study

### 1.2 基于表型数据的初级核心种质构建方法

形态表型数据标准参考《甘蔗种质资源描述规范和数据标准》<sup>[18]</sup>,并按照文献<sup>[19]</sup>的取样策略对

147份斑茅进行分组、抽样,通过抽样材料子库和原始材料总库的表型遗传多样性比较评价来筛选确定斑茅初级库材料。取样方案及代码见表1,表中括

号内字母表示取样方案缩写代码,代码组合即为该取样策略构建的子库名称,如 LLC50 表示该子库是以叶型分组并按对数比例进行系统聚类取样,然后按 50% 的总体取样比率挑选子库材料。

表 1 表型取样策略及代码

Table 1 Phenotypic sampling strategy and codes

分组原则 Grouping principle	组内取样比例 Sampling ratio in group	取样方法 Sampling method	总体取样比例 Total sampling ratio
采集地分组(O)	简单比例(P)	系统聚类(C) 完全随机(R)	10% ~ 50%
叶型分组(L)	对数比例(L)		
海拔分组(H)	平方根比例(S)		
气候区域分组(Z)	多样性比例(G)		
总体聚类分组(C)			
不分组随机(N)			

表 2 SSR 和 AFLP 引物序列

Table 2 The primer sequences of SSR and AFLP analyses

引物 Primer	正向引物序列 Forward primer sequence	反向引物序列 Reverse primer sequence
Msscir17	AGC ATA GTT TTT GTG GAC	AGC ATA GTT TTT GTG GAC
Msscir21	CGC CAG CCA CAT AAA AGG	CGC CAG CCA CAT AAA AGG
Msscir26	AAA ATC AGA CAA ACA GCA T	AGA AGA AGC AGA TAC AGG T
Msscir33	GCT CAT ATA TCT TCC TGG TC	AGT GGT CTG GTG CTT TGG
Msscir36	CAA CAA TAA CTT AAC TGG TA	CTG TCC TTT TTA TTC TCT TT
Msscir43	ATT CAA CGA TTT TCA CGA G	AAC CTA GCA ATT TAC AAG AG
SMC668	ACG CTT GCG TGC TCC ATT	CCA ATC GTG CCA CTG TAG TAA G
SMC720	CGC ACC GAC GCA CGT CT	GCC AAT GGA ACG GGT CTA
SMC1752	GGC TGA TTT ACA TGA AAC TGT TCT	AAA GCT GGT ATC CCA GCA TAC T
SMC2055	TGA GCA CAT GGT TCT GCA AG	TGA GCA CAT GGT TCT GCA AG
E-ACC/M-CTC	GAC TGC GTA CCA ATT CACC	GAT GAG TCC TGA GTA ACTC
E-ACG/M-CTG	GAC TGC GTA CCA ATT CACG	GAT GAG TCC TGA GTA ACTG
E-AGC/M-CTA	GAC TGC GTA CCA ATT CAGC	GAT GAG TCC TGA GTA ACTA

基于 SSR 和 AFLP 标记数据,采用多态位点优先选择的多次聚类取样策略<sup>[23-24]</sup>构建斑茅微核心种质,设随机取样为对照取样策略,取样总体原则是优先保留稀有基因型并减少有效等位基因的丢失。为了确保所构建的斑茅微核心种质的遗传多样性,相似材料和特殊材料<sup>[25]</sup>的取舍依据以下原则:(1)优先选择 SSR 稀有基因型,若两份材料 SSR 基因型相同,保留 AFLP 位点数较多的材料;(2)若两份材料 SSR 基因

### 1.3 基于分子数据的微核心种质构建方法

参照优化的甘蔗基因组 DNA 提取方法<sup>[20]</sup>提取斑茅初级核心种质叶片 DNA,筛选 10 对 SSR 引物和 3 对 AFLP 引物进行 PCR 扩增,引物序列见表 2。SSR 反应总体系 20  $\mu$ L,包括 10  $\times$  PCR Buffer 2  $\mu$ L、MgCl<sub>2</sub> (25 mmol/L) 2.4  $\mu$ L、DNTPs (10 mmol/L each) 0.4  $\mu$ L、SSR 正反引物(100  $\mu$ mol/L)各 0.05  $\mu$ L,以及 0.5 U Taq DNA 聚合酶和 20 ng 的模版 DNA。SSR 扩增程序:94  $^{\circ}$ C 预变性 3 min;35 个循环(94  $^{\circ}$ C 变性 1 min,54  $^{\circ}$ C 退火 2 min,72  $^{\circ}$ C 延伸 1 min);72  $^{\circ}$ C 延伸 5 min。AFLP 反应体系及扩增程序采用蔡青等<sup>[21]</sup>的方法。SSR 和 AFLP 扩增产物经 95  $^{\circ}$ C 变性后在 5% 的变性聚丙烯酰胺凝胶上电泳分离,用刘新龙等<sup>[22]</sup>的方法进行快速银染,有带记为“1”,无带记为“0”,形成 0-1 矩阵进行数据分析。

型和 AFLP 位点数均相同,保留抗性(抗旱性、抗锈病)较优的材料;(3)若两份材料 SSR 基因型、AFLP 位点数相同,抗性相同或抗性情况不明,保留有效等位基因数较多的材料。

首先采用 SM 匹配系数(SM, Sokal-Michene Matching coefficient)、Jaccard 系数(J, Jaccard index)和 Dice 系数(D, Dice coefficient)3 种相似系数进行 UPGMA 多次聚类,根据每次聚类结果,对相似性最

高的两份材料,采用位点优先选择的方法,剔除 1 份,剩余材料组成一个子库,然后计算子库的 Nei's 基因多样性( $H_e$ )、Shannon 指数( $I$ ),同时使用 T 检验评价子库和原库的遗传多样性是否有显著差异,直到子库的遗传多样性与原库开始有显著差异时,终止多次聚类取样,选择上一个与原库没有显著差异的子库作为该取样策略下的最佳微核心种质。对于不同相似性系数构建的最佳微核心种质,使用 3 次重复随机取样方法,建立随机库作为对照,通过计算微核心种质的 Nei's 基因多样性和 Shannon 指数等参数,与原库进行 T 检验来评价微核心种质的代表性,同时利用分子标记多态条带数据和数量性状形态数据对微核心种质库进行确认<sup>[24,26]</sup>。最后根据“核心种质均值差异百分率小于 20%,同时极差符合率大于 80% 可认为能够代表原种质遗传多样性”的经验标准<sup>[23]</sup>进行最优微核心种质的确定。

#### 1.4 数据分析软件与相关公式

原始数据的 UPGMA 聚类分析使用 NTSYS 2.1 软件,核心种质的遗传多样性分析、T 检验和形态数据确认使用 POPGENE 1.32、SPSS 18.0 和 Excel 2007 软件。

SM( $GS_{Sij}$ )、Jaccard( $GS_{Jij}$ )和 Dice( $GS_{Dij}$ )3 种相似系数计算公式: $GS_{Sij} = (a + d) / (a + b + c + d)$ ;  $GS_{Jij} = a / (a + b + c)$ ;  $GS_{Dij} = 2a / [(a + b) + (a + c)]$ 。式中 a、b、c 和 d 的定义见表 3。

表 3 所有个体对 i 和 j 带型统计

Table 3 Statistical list of band types for all pairs of two individuals i and j

个体编号		个体 j	
Individual number		有带 Band	无带 No band
个体 i	有带 Band	A	b
	无带 No band	c	d

Nei's 基因多样性( $H_e$ )计算公式: $H_e = 1 - \sum_{i=1}^n p_i^2$

Shannon 指数( $I$ )计算公式: $I = - \sum_{i=1}^n p_i \log p_i$

式中  $p_i$  为 SSR 标记第  $i$  个基因位点的基因型频率或 AFLP 标记第  $i$  个基因位点的等位基因频率。

核心种质多态条带百分率(PB)、多态条带保留率(PBR)、变异系数符合率(VR)、极差符合率(CR)、方差差异百分率(VD)和均值差异百分率(MD)的计算公式:

$$PB(\%) = \frac{\text{多态条带数}}{\text{总条带数}} \times 100$$

$$PBR(\%) = \frac{\text{核心库多态条带数}}{\text{原始库多态条带数}} \times 100$$

$$VR(\%) = \frac{\sum_{j=1}^m \text{核心库变异系数}}{\text{原始库变异系数}} \times 100$$

$$CR(\%) = \frac{1}{m} \sum_{j=1}^m \frac{\text{核心库极差}}{\text{原始库极差}} \times 100$$

$$VD(\%) = \frac{\text{核心库与原始库具有显著方差差异的性状个数}}{\text{总性状个数}} \times 100$$

$$MD(\%) = \frac{\text{核心库与原始库具有显著均值差异的性状个数}}{\text{总性状个数}} \times 100$$

## 2 结果与分析

### 2.1 初级核心种质的表型遗传多样性和代表性

斑茅初级核心种质的取样策略评价结果表明,从表型数据上分析,以叶型分组和按对数比例聚类取样(LLC)是构建初级库的最佳取样策略。根据该策略对 147 份斑茅材料进行抽样,在子库取样量达到总库 35% 时,其 Shannon 指数达到 0.6718,表型保留比例达到 94.72%。根据对 LLC35 子库丢失的表型性状(表 4)及入选材料的采集地进行分析,将含有抽样时极易丢失的 5 个稀有表型性状的代表材料各取 1 份纳入子库(采集地不同的材料优先考虑),最终构建的斑茅初级核心种质含 58 份材料,Shannon 指数提高至 0.6834,表型保留比例提高至 100%。

表 4 最优初级库 LLC35 的表型性状丢失情况

Table 4 Phenotypic trait lost in the best primary core collection LLC35

表型性状*	代码描述	总份数	丢失表型性状情况及处理
Phenotypic trait	Code description	Total number	Phenotypic trait lost situation and processing
节间形状 4	节间呈圆锥型	2	丢失,原始库中取样 1 份
节间形状 5	节间呈倒圆锥型	3	部分重复丢失;未丢失重复里取样 1 份
茎径 10	茎径第 10 分组	3	部分重复丢失;未丢失重复里取样 1 份
叶长 9	叶长第 9 分组	3	部分重复丢失;未丢失重复里取样 1 份
叶宽 10	叶宽第 10 分组	3	部分重复丢失;未丢失重复里取样 1 份

\*: 该栏数字为性状对应的表型代码或分组

\*: The numbers in this column represent the phenotypic code of group for traits

## 2.2 初级核心种质的 SSR 和 AFLP 遗传多样性

采用 SSR 和 AFLP 标记对 58 份斑茅初级核心种质材料的遗传多样性进行分析(表 5)。10 对 SSR 引物共扩增出 105 条带,各引物的 Nei's 基因多样性在 0.5166~0.9822 之间,Shannon 指数在 0.3975~

1.7530 之间;3 对 AFLP 引物共扩增出 140 条带,各引物的 Nei's 基因多样性在 0.2166~0.2503 之间,Shannon 指数在 0.3382~0.3848 之间。结果表明,斑茅具有较丰富的遗传多样性,且 SSR 标记检出的遗传多样性高于 AFLP 标记。

表 5 斑茅基于 SSR 和 AFLP 标记的遗传多样性

Table 5 Genetic diversity of *S. arundinaceum* L. based on SSR and AFLP markers

SSR 引物 SSR primers	条带 Bands	Nei's 基因多样性 <i>He</i>	Shannon 指数 <i>I</i>	AFLP 引物 AFLP primers	条带 Bands	Nei's 基因多样性 <i>He</i>	Shannon 指数 <i>I</i>
mssc17	5	0.8799	1.0081	E-ACC/M-CTC	47	0.2452	0.3755
mssc121	9	0.6807	0.6724	E-ACG/M-CTG	46	0.2166	0.3382
mssc126	30	0.9822	1.7530	E-AGC/M-CTA	47	0.2503	0.3848
mssc133	6	0.7640	0.7705				
mssc136	17	0.9810	1.7323				
mssc143	7	0.7955	0.7805				
SMC668	5	0.7794	0.8161				
SMC720	4	0.5166	0.3975				
SMC1752	17	0.9721	1.6201				
SMC2055	5	0.9376	1.2511				
SSR 总计	105	0.8289	1.0802	AFLP 总计	140	0.2374	0.3662

## 2.3 斑茅初级核心种质的聚类分析与微核心种质的遗传多样性检验

使用 NTSYS-2.1 软件对初级核心种质材料进行 UPGMA 多次聚类分析,发现 Jaccard 和 Dice 系数聚类结果相同,SM 系数聚类结果略有不同,且随着剔除样品份数的增多,各子库的遗传多样性呈下降趋势。其中,AFLP 标记有效等位基因数的 Nei's 基因多样性和 Shannon 指数对样品数变化均不敏感,SSR 标记基因型频率的 Nei's 基因多样性对样品数变化也不敏感,但 Shannon 指数则较为敏感。具体表现为,当样品数降低到 15 时,与原始库相比,子库的 AFLP 标记遗传多样性和 SSR 标记的 Nei's 基因多样性变化均未达到 5% 概率水平的显著差异,而 SSR 标记的 Shannon 多样性指数在 5% 概率水平开始出现显著差异。故根据 SSR 标记 Shannon 多样性指数确定斑茅微核心种质的最佳取样量为 16 份样品。

按照 SSR 标记 Shannon 多样性指数出现显著差异前后的取样量 16 份和 15 份进行随机取样构建随机微核心种质(各参数计算按 3 次重复取平均值),

分别编号为 R-16 和 R-15,连同 3 种相似系数构建的 6 个微核心种质(SM-16, SM-15, J-16, J-15, D-16, D-15)一起与初级库进行 T 检验,结果见表 6 和表 7。表中数据显示:在相同取样量下,随机微核心种质的多样性低于聚类微核心种质,且与初级库的遗传多样性差异显著,表明聚类取样优于随机取样;同时在取样量为 16 个样品时,随机取样的 SSR 标记 Shannon 多样性指数与初级库差异显著,而聚类取样无显著差异,因此可认为聚类取样明显优于随机取样。

## 2.4 斑茅微核心种质的代表性

为进一步确认基于 SSR 和 AFLP 标记构建的斑茅微核心种质的代表性,对 2 种分子标记数据的多态条带数以及株高、茎径、叶长、叶宽、锤度等 5 个数量形态性状数据的变异系数、极差、方差、均值进行了统计,并计算各核心种质的多态条带百分率(PB%)、多态条带保留率(PBR%)、变异系数符合率(VR)、极差符合率(CR)、方差差异百分率(VD%)和均值差异百分率(MD%),结果见表 8 和表 9。

表 6 斑茅微核心种质 SSR 遗传多样性 T 检验

Table 6 T-test for SSR genetic diversity in mini core collection of *S. arundinaceum* L.

微核心种质 Mini core collection	Nei's * 基因多样性 $H_e$	平均差 Mean difference	T 值 T value	P 值 Pr >  t	Shannon * 指数 $I$	平均差 Mean difference	T 值 T value	P 值 Pr >  t
SM-16	0.8273	-0.0016	-0.0400	0.9690	0.8910	-0.1892	-2.1347	0.0616
SM-15	0.8240	-0.0049	-0.1300	0.8990	0.8749	-0.2053	-2.4322	<b>0.0378 *</b>
J-16	0.8297	0.0008	0.0220	0.9830	0.8996	-0.1806	-2.1792	0.0572
J-15	0.8213	-0.0076	-0.2170	0.8330	0.8739	-0.2063	-2.5834	<b>0.0295 *</b>
D-16	0.8297	0.0008	0.0220	0.9830	0.8996	-0.1806	-2.1792	0.0572
D-15	0.8213	-0.0076	-0.2170	0.8330	0.8739	-0.2063	-2.5834	<b>0.0295 *</b>
R-16	0.7883	-0.0406	-0.7976	0.4604	0.8395	-0.2407	-2.4880	<b>0.0372 *</b>
R-15	0.7849	-0.0440	-0.8751	0.4262	0.8282	-0.2520	-2.6543	<b>0.0295 *</b>

\* :初级库 SSR 标记的  $H_e = 0.8289, I = 1.0802$ \* :For primary core collection by SSR,  $H_e = 0.8289, I = 1.0802$ 

表 7 斑茅微核心种质 AFLP 遗传多样性 T 检验

Table 7 T-test for AFLP genetic diversity in mini core collection of *S. arundinaceum* L.

微核心种质 Mini core collection	Nei's * 基因多样性 $H_e$	平均差 Mean difference	T 值 T value	P 值 Pr >  t	Shannon * 信息指数 $I$	平均差 Mean difference	T 值 T value	P 值 Pr >  t
SM-16	0.2339	-0.0035	-0.2148	0.8302	0.3516	-0.0146	-0.6260	0.5323
SM-15	0.2316	-0.0058	-0.3495	0.7272	0.3473	-0.0189	-0.8037	0.4230
J-16	0.2311	-0.0063	-0.3946	0.6937	0.3501	-0.0161	-0.7120	0.4777
J-15	0.2300	-0.0074	-0.4581	0.6476	0.3481	-0.0181	-0.7940	0.4286
D-16	0.2311	-0.0063	-0.3946	0.6937	0.3501	-0.0161	-0.7120	0.4777
D-15	0.2300	-0.0074	-0.4581	0.6476	0.3481	-0.0181	-0.7940	0.4286
R-16	0.2160	-0.0214	-1.3415	0.1911	0.3286	-0.0376	-1.6486	0.1081
R-15	0.2152	-0.0222	-1.3851	0.1763	0.3273	-0.0389	-1.7019	0.0961

\* :初级库 AFLP 标记的  $H_e = 0.2374, I = 0.3662$ \* :For primary core collection by AFLP,  $H_e = 0.2374, I = 0.3662$ 

表 8 斑茅微核心种质代表性的分子数据确认

Table 8 Confirmation of molecular data for representativeness of mini core collection of *S. arundinaceum* L.

微核心种质 Mini core collection	SSR 标记多态条带统计 Polymorphic bands of SSR marker			AFLP 标记多态条带统计 Polymorphic bands of SSR marker		
	多态 条带数	多态条带 百分率 (%)	多态条带 保留率 (%)	多态 条带数	多态条带 百分率 (%)	多态条带 保留率 (%)
	PN	PB	PBR	PN	PB	PBR
SM-16	95.00	90.48	94.06	96.00	68.57	86.49
J-16	96.00	91.43	95.05	97.00	69.29	87.39
D-16	96.00	91.43	95.05	97.00	69.29	87.39
R-16	90.00	85.80	89.11	94.00	67.28	84.68
初级核心种质 Primary core collection	101.00	96.19	100.00	111.00	79.29	100.00

表 9 斑茅微核心种质代表性的形态数据确认

Table 9 Confirmation of morphology data for representativeness of mini core collection of *S. arundinaceum* L.

微核心种质 Mini core collection	变异系数 符合率(%) VR	极差符 合率(%) CR	方差差异 百分率(%) VD	均值差异 百分率(%) MD
SM-16	99.45	83.60	0.00	0.00
J-16	100.62	79.15	0.00	0.00
D-16	100.62	79.15	0.00	0.00
R-16	105.17	82.77	0.00	6.67
初级核心种质 Primary core collection	100.00	100.00	0.00	0.00

分子数据确认结果表明:SSR 标记多态条带百分率和保留率均高于 AFLP 标记;聚类取样能够保留初级库 86.49% ~ 95.05% 的多态条带,而随机取样多态条带保留率只有 84.98% ~ 88.78%,且 SM、Jaccard 和 Dice 系数构建的聚类取样微核心种质差别不大。

形态数据确认结果表明:聚类取样微核心种质各数量性状的均值和方差与初级库无显著差异,变异系数与初级库差异也不大(-0.55% ~ +0.62%),在极

表 10 斑茅核心种质构建结果及代表材料

Table 10 The results of core collection of *S. arundinaceum* L. and representative materials

采集地 Sampling sites	原始库 Original materials	初级核心种质 Primary core collection	微核心种质 Mini core collection	代表材料 Representative materials
云南	59	24	4	云南 76- I -13、云南 76- II -10、云南 82-109、云南 82-113、云南 82-118、云南 82-123、云南 82-128、云南 82-138、云南 82-143*、云南 82-30*、云南 82-49*、云南 82-53、云南 82-56、云南 82-57、云南 82-62、云南 82-76*、云南 83-182、云南 83-218、云南 83-222、云南 83-232、云南 83-234、云南 84-241、云南 87-272、云南 89-275
贵州	19	4	1	贵州 95-30、贵州 95-22、贵州 78- II -12*、贵州 78- I -24
海南	18	8	2	海南 92-103、海南 92-79、海南 92-106、海南 92-109*、海南 92-84、海南 92-85、崖城 2 号*、崖城 3 号
福建	15	10	4	福建 87- II -2*、福建 87- II -3、福建 87- II -5*、福建 87- II -10*、福建 87- II -20、福建 88- II -2、福建 88- II -5、福建 89- II -4、福建 89- II -5*、福建东山斑茅
四川	14	5	2	四川 79- III -24*、四川 79- I -28、四川 79- I -22、四川 79- II -3、四川 92-1*
江西	10	2	1	江西 79-12*、江西 91-5
广西	5	1	1	广西 84-18*
浙江	4	2	0	浙江 93-5、浙江 93-11
广东	3	2	1	广东梅县斑茅*、广东 15 号
合计	147	58	16	

\*: 为入选微核心种质的材料

\*: The materials with an asterisk were selected into mini core collections

差符合率上 SM 聚类微核心种质高于 80%, 而 Jaccard 和 Dice 聚类微核心种质低于 80%; 随机微核心种质各数量性状的方差与初级库无显著差异, 均值存在一定差异, 变异系数与初级库存在较大差异(+5.17%), 随机微核心种质的极差符合率为 82.77%, 低于 SM 聚类微核心种质(83.6%) 但高于 Jaccard 和 Dice 聚类微核心种质(79.15%)。

综上所述, 根据“核心种质均值差异百分率小于 20%, 同时极差符合率大于 80% 可认为能够代表原种质遗传多样性<sup>[23]</sup>”的经验标准, 最终确定基于 SSR 标记数据的 SM 相似系数多次聚类方法适用于斑茅微核心种质的构建, 该方法构建的微核心种质含 16 份材料, 经检测其均值差异百分率为 0, 极差符合率为 83.6%, 作为微核心种质库的代表性较好。

## 2.5 斑茅核心种质库最终构建结果

以来自全国 9 省区的 147 份斑茅为原始材料, 首先基于表型数据, 采用叶型分组和按对数比例聚类取样(LLC)的秩数加权平均取样策略, 筛选 58 份材料构建斑茅初级核心种质; 再基于 SSR 和 AFLP 分子标记数据, 采用 SM 相似系数进行 UPGMA 多次聚类, 筛选 16 份材料构建微核心种质。斑茅核心种质最终构建结果及代表材料见表 10。

### 3 讨论

#### 3.1 斑茅核心种质的构建方法指标和取样比例

从核心种质的概念可知,选择有效的取样方法及合理的参数对不同遗传材料间遗传差异进行准确度量是构建核心种质的关键。核心种质构建的基本原则就是剔除相同或相近的个体,将资源数量控制在合理范围内,同时最大可能地保留其遗传多样性,因此确定合理的取样比例也至关重要。目前关于不同取样方法、多样性指标和取样比例对核心种质构建的有效性研究还不多。其中许海明等<sup>[27]</sup>对不同遗传距离聚类 and 抽样方法构建作物核心种质进行了比较研究,认为在核心种质构建中,马氏距离优于欧氏距离,优先取样法略优于偏离度取样法,系统聚类方法最好的是最短距离法(Single linkage),其次是中间距离(Median method)、重心法(Centroid method)和类平均法(UPGMA)。李自超等<sup>[4]</sup>、张洪亮等<sup>[25]</sup>研究认为多样性指数、表型方差、表型频率方差及变异系数为对比不同核心种质取样方法间优劣的有效参数,而表型保留比例为检验核心种质最终有效性和取样比例必不可少的参数。A. H. D. Brown<sup>[3]</sup>根据中性选择学说推导认为,理论上占资源总数5%~10%的核心样品就能够代表整个资源70%以上的变异。国内外已构建的各类作物核心种质取样比例一般在10%~30%之间,具体根据原始群体的大小调整抽样比例,大群体采用较低的取样比例(10%~20%),小群体采用较高的取样比例(20%~30%)<sup>[27]</sup>。当然,由于物种差异、生物遗传变异和进化以及人工选择等诸多因素的影响,不同物种核心种质取样比例的确定不能一概而论,应根据物种遗传特性、遗传结构及建库目的具体分析。

本研究借鉴前人研究经验,基于表型性状和SSR、AFLP分子标记数据,选择Shannon多样性指数、Nei's基因多样性、多态条带百分率、多态条带保留率、变异系数符合率、极差符合率、方差差异百分率和均值差异百分率等参数作为多样性评价指标,利用147份斑茅无性系分别构建了斑茅初级核心种质(58份)和微核心种质(16份)并进行了代表性验证。研究结果表明采用叶型分组按对数比例聚类取样策略(表型数据)和多态位点优先选择的UPGMA多次聚类取样策略(分子数据)比例较适合斑茅核心种质的构建,取样比例以初级核心种质不低于35%,微核心种质不低于10%

为宜。

#### 3.2 斑茅核心种质构建中分子标记和遗传参数的选择

在几种建库常用分子标记中,RAPD标记快速简便,多态性高,但为显性标记,重复性差,且存在共迁移现象,因而不适宜种间或亲缘关系较远材料的遗传分析<sup>[28]</sup>。而SSR(共显性)和AFLP(显性标记,可转化为共显)标记的实验结果稳定可靠,重复性高,多态性丰富,因而广泛用于核心种质构建、遗传图谱分析、资源鉴定等各个方面<sup>[29-32]</sup>。分子标记数据不同于呈现连续变异的数量性状,具有离散分布的特征,因此需要采用与数量性状不同的数据处理方法来计算遗传距离。目前对于分子标记数据主要采用以下3种相似系数来估算样品间的遗传距离,即Dice(或Nei-Li)系数<sup>[33-34]</sup>、Jaccard系数<sup>[35-36]</sup>、SM(或Sokal-Michener)系数<sup>[37-38]</sup>。刘遵春等<sup>[39]</sup>在构建新疆野苹果核心种质时曾对常用遗传距离估算方法进行了比较,认为在50%的取样比例下,上述3种遗传距离构建的核心种质都能够代表原始群体的遗传多样性,其中SM系数估算的遗传距离代表性优于Jaccard和Dice估算的遗传距离。

本研究在基于SSR和AFLP标记数据进行斑茅核心种质构建时也得到了类似的结果,研究发现,随着总体取样量的减小,SSR基因型频率变化较AFLP有效等位基因数变化要敏感,SM、Jaccard和Dice3种相似系数聚类取样结果均优于随机取样,根据T检验、分子标记和数量性状数据确认结果,认为SSR和AFLP是构建核心种质的高效标记,但不同物种效率不同,对于斑茅核心种质的构建,SSR标记优于AFLP标记,同时选用SM系数计算遗传距离有利于提高核心种质的代表性。

#### 3.3 斑茅核心种质的利用策略

上述方法构建的斑茅初级核心种质包含了斑茅大部分的表型遗传变异,因此可作为甘蔗遗传育种的野生血缘亲本备选材料库,开展创新型杂交利用和优良后代筛选评价研究。斑茅微核心库则包含了斑茅大部分的基因遗传变异,因此可作为分子遗传学和基因组学研究的核心材料,用于斑茅分子标记特异引物的筛选开发、基因克隆与测序等分子水平的研究。

#### 参考文献

- [1] 廖兆周,劳方业,周耀辉,等.具有斑茅种质的耐旱甘蔗品系的选育[J].作物学报,2002,28(6):841-846
- [2] Frankel O H. Genetic perspectives of germplasm conservation[C]//

- Arber W, Illemensee K, Peacock W J, et al. Genetic Manipulation: Impact on Man and Society. Cambridge: Cambridge University Press, 1984:161-170
- [3] Brown A H D. Core collections: a practical approach to genetic resources management [J]. Genome, 1989, 31: 818-824
- [4] 李自超, 张洪亮, 曾亚文, 等. 云南地方稻种资源核心种质取样方案研究[J]. 中国农业科学, 2000, 33(5): 1-7
- [5] 李自超, 张洪亮, 曹永生, 等. 中国地方稻种资源初级核心种质取样策略研究[J]. 作物学报, 2003, 29(1): 20-24
- [6] Rodino A P, Santalla M, Ron A M D, et al. A core collection of common bean from the Iberian peninsula [J]. Euphytica, 2003, 131: 165-175
- [7] 赵丽梅, 董英山, 刘宝, 等. 中国一年生野生大豆核心种质的构建[J]. 科学通报: 英文版, 2005, 50(10): 989-996
- [8] 董玉琛, 曹永生, 张学勇, 等. 中国普通小麦初选核心种质的产生[J]. 植物遗传资源学报, 2003, 4(1): 1-8
- [9] 郝晨阳, 董玉琛, 王兰芬, 等. 我国普通小麦核心种质的构建及遗传多样性分析[J]. 科学通报, 2008, 53(8): 908-915
- [10] Balakrishnan R, Nair N V, Sreenivasan T V. A method for establishing a core collection of *Saccharum officinarum* germplasm based on quantitative morphological data [J]. Genet Resour Crop, 2000, 47: 1-9
- [11] 刘新龙, 蔡青, 马丽, 等. 甘蔗杂交品种初级核心种质取样策略[J]. 作物学报, 2009, 35(7): 1209-1216
- [12] 刘新龙, 刘洪博, 马丽, 等. 利用分子标记数据逐步聚类取样构建甘蔗杂交品种核心种质库[J]. 作物学报, 2014, 40(11): 1885-1894
- [13] Tai P Y P, Miller J D. A core collection for *Saccharum spontaneum* L. from the world collection of sugarcane [J]. Crop Sci, 2001, 41: 879-885
- [14] 苏火生, 刘新龙, 毛钧, 等. 割手密初级核心种质取样策略研究[J]. 湖南农业大学学报: 自然科学版, 2011, 37(3): 253-259
- [15] 齐永文, 樊丽娜, 罗青文, 等. 甘蔗细茎野生种核心种质构建[J]. 作物学报, 2013, 39(4): 649-656
- [16] 邓念丹, 刘清波, 蒋建雄, 等. 五节芒核心种质的构建研究概述[J]. 作物研究, 2010, 24(1): 130-134
- [17] 肖亮, 易自力, 蒋建雄, 等. 中国芒 (*Miscanthus sinensis*) 初级核心种质的构建[J]. 植物遗传资源学报, 2014, 15(6): 1196-1201
- [18] 蔡青, 范源洪. 甘蔗种质资源描述规范和数据标准 [M]. 北京: 中国农业出版社, 2006: 7-31
- [19] 毛钧, 刘新龙, 苏火生, 等. 斑茅初级核心种质秩数加权平均取样策略[J]. 分子植物育种, 2011, 9(1): 1001-1009
- [20] 应雄美, 蔡青, 毕艳, 等. 两种不同甘蔗基因组 DNA 提取方法的比较[J]. 中国糖料, 2009(4): 22-23
- [21] 蔡青, 范源洪, Aitken K, 等. 利用 AFLP 进行“甘蔗属复合体”系统演化和亲缘关系研究[J]. 作物学报, 2005, 31(5): 551-559
- [22] 刘新龙, 蔡青, 毕艳, 等. 甘蔗 AFLP 标记和 SSR 标记的 PAGE 胶快速银染检测方法 [J]. 江苏农业学报, 2009, 25(2): 433-435
- [23] Hu J, Zhu J, Xu H M. Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops [J]. Theor Appl Genet, 2000, 101: 264-268
- [24] 胡晋, 徐海明, 朱军. 基因型值多次聚类法构建作物种质资源核心库[J]. 生物数学学报, 2000, 15(1): 103-109
- [25] 胡晋, 徐海明, 朱军. 保留特殊种质材料的核心库构建方法[J]. 生物数学学报, 2001, 16(3): 348-352
- [26] 张洪亮, 李自超, 曹永生, 等. 表型水平上检验水稻核心种质的参数比较[J]. 作物学报, 2003, 29(2): 252-257
- [27] 徐海明, 邱英雄, 胡晋, 等. 不同遗传距离聚类和抽样方法构建作物核心种质的比较[J]. 作物学报, 2004, 30(9): 932-936
- [28] Thormann C E, Ferreira M E, Camargo L E A, et al. Comparison of RFLP and RAPD markers to estimating genetic relationships within and among cruciferous species [J]. Theor Appl Genet, 1994, 88(8): 973-980
- [29] 车克鹏, 许勇, 梁春阳, 等. 西瓜核心种质的 AFLP 指纹图谱和 SCAR 标记[J]. 植物学报, 2003, 45(6): 731-735
- [30] 郝晨阳, 王兰芬, 董玉琛, 等. 我国西北春麦区小麦育成品种遗传多样性的 AFLP 分析[J]. 植物遗传资源学报, 2003, 4(4): 285-291
- [31] 陆鑫, 毛钧, 刘洪博, 等. 甘蔗野生种滇蔗茅种质创新利用研究 I. 甘蔗与滇蔗茅远缘杂交 F1 群体构建与 SSR 分子标记鉴定[J]. 植物遗传资源学报, 2012, 13(2): 321-324
- [32] Zhang H L, Li Z C, Liao D Q, et al. Microsatellite analysis of landrace rice core collection in Yunnan, China [J]. J Agric Biotech, 2003, 11(2): 131-139
- [33] Dice L R. Measures of the amount of ecological association between species [J]. Ecology, 1945, 26: 297-302
- [34] Nei M, Li W H. Mathematical models for studying genetic variation in terms of restriction endonucleases [J]. Proc Natl Acad Sci USA, 1979, 76: 5269-5273
- [35] Jaccard P. Contribution au problème de la l'immigration post-glaciaire de la flore alpine [J]. Bull Soc Vaud Sci Nat, 1900, 37: 547-579
- [36] Jaccard P. Nouvelles recherches sur la distribution florale [J]. Bull Soc Vaud Sci Nat, 1908, 44: 223-270
- [37] Sokal R, Michener C D. A statistical method for evaluating systematic relationships [J]. Univ Kansas Sci Bull, 1958, 38: 1409-1438
- [38] Sokal R. Distance as a measure of taxonomic similarity [J]. Syst Zool, 1961, 10(2): 40-51
- [39] 刘遵春, 刘大亮, 崔美, 等. 整合农艺性状和分子标记数据构建新疆野苹果核心种质 [J]. 园艺学报, 2012, 39(6): 1045-1054