

研究论文

Research Article

斑茅初级核心种质秩数加权平均取样策略

毛钧¹, 刘新龙¹, 苏火生¹, 应雄美¹, 陆鑫¹, 马丽¹, 范源洪¹, 蔡青^{1,2}

1. 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, 661600

2. 云南省农业科学院生物技术与种质资源研究所, 昆明, 650223

✉ 通讯作者: caiqingsri@163.com; ✉ 作者

分子植物育种, 2011年, 第9卷, 第1篇 doi: 10.5376/mpb.cn.2010.09.0001

收稿日期: 2010年10月17日

接受日期: 2010年11月29日

发表日期: 2011年01月06日

这是一篇采用 Creative Commons Attribution License 进行授权的开放取阅论文。只要对本原作有恰当的引用, 版权所有人允许并同意第三方无条件的使用与传播。

引用格式:

毛钧等, 2011, 斑茅初级核心种质秩数加权平均取样策略, 分子植物育种 Vol.9 No.1 (doi: 10.5376/mpb.cn.2011.09.0001)

摘要 本研究以国家甘蔗种质资源圃收集保存的 147 份斑茅野生资源为原始材料, 选取 26 个质量和数量性状, 采用秩数加权分析的方法, 从取样方法、分组原则、组内取样比例 3 个层次探讨构建斑茅初级核心种质的最佳取样策略。取样方法采用聚类取样(C)和随机取样(R); 分组原则包括采集地分组(O)、叶型分组(L)、海拔分组(H)、气候区域分组(Z)、总体聚类分组(C)及不分组随机(N); 组内取样比例包括简单比例(P)、对数比例(L)、平方根比例(S)和多样性比例(G)。9 个总体取样梯度为 10%、15%、20%、25%、30%、35%、40%、45% 和 50%, 每梯度设 3 次重复。通过比较香农多样性指数、变异系数、表型方差、表型频率方差和表型保留比例 5 个参数的秩数加权平均值来检验各取样策略的优劣。结果表明, 斑茅初级核心种质最佳取样策略为叶型分组、对数比例、聚类取样, 最佳总体取样量为 35%。在此基础上, 加入缺失表型材料、优良亲本和抗性材料共计 58 份组成最终的斑茅初级核心种质库, 占总资源数的 39.46%。

关键词 斑茅; 核心种质; 秩数; 加权平均值; 取样策略

Sampling Strategy of Weighted Average Rank for Pre-Core Germplasm in *Erianthus arundianaceus*

Mao Jun¹, Liu Xinlong¹, Su Huosheng¹, Ying Xiongmei¹, Lu Xin¹, Ma Li¹, Fan Yuanhong¹, Cai Qing^{1,2}

1. Sugarcane Research Institute, Yunnan Academy of Agricultural Sciences / Yunnan Key Laboratory of Sugarcane Genetic Improvement, Kaiyuan, 661600, P.R. China

2. Biology & Genetic Germplasm Institute, Yunnan Academy of Agricultural Sciences, Kunming, 650231, P.R. China

✉ Corresponding author, caiqingsri@163.com; ✉ Authors

Abstract In this study, 147 clones of *Erianthus arundianaceus* from the National Nursery of Sugarcane Germplasm Resources (NNSGR) in Kaiyuan China were selected as original materials. Based on 26 quantitative and qualitative traits, the best sampling strategy was discussed by weighted average rank analysis method from 3 levels as follows: 2 sampling methods (random sampling and cluster sampling); 6 grouping principles (original grouping, leaves grouping, height grouping, zone grouping, cluster grouping and not grouping); 4 sampling ratios within group (simple ratios, logarithmic ratios, square root ratios and genetic ratios). At the same time, 9 grades for total sampling percentages (10%, 15%, 20%, 25%, 30%, 35%, 40%, 45% and 50%) were analyzed and each grades included 3 repeats. The sampling strategies and the pre-core collection were evaluated by weighted average rank of five effective parameters (Shannon diversity index, coefficient of variation, variance of phenotypic value, variance of phenotypic frequency and ratio of retaining phenotypic traits). The result indicated that the best sampling strategy for *Erianthus arundianaceus* was set up based on leaves grouping, logarithmic ratio, cluster sampling and 35% total sampling percentage. And some manual selected materials were put in the pre-core collection, which including the rare phenotypic trait clones, the elite parents or resistibility clones. Finally, the pre-core collection of *Erianthus arundianaceus* included 58 clones which accounted for 39.46% of total resources.

Keywords *Erianthus arundianaceus*; Core germplasm; Rank; Weighted average; Sampling strategy

研究背景

斑茅(*Erianthus arundinaceus*)是“甘蔗复合体(*Saccharum complex*)”中重要的野生资源, 具有丛

生性好、萌芽力强、分蘖多、生长势强、宿根性好、适应性广、抗逆性强、抗病虫性强等优异性状, 成为现代甘蔗遗传育种和种质创新中利用较多的种

质资源。国家甘蔗种质资源圃收集保存了采自全国9个省区不同生态类型的斑茅野生材料147份, 如何有效保存和高效利用斑茅野生资源, 是甘蔗种质资源研究的重要任务和研究方向。核心种质(Frankel, 1984; Brown, 1989)的提出和完善为生物种质资源的有效保存和高效利用提供了解决途径。国内外已经在水稻(李志超等, 2000; 李志超等, 2003)、菜豆(Rodino et al., 2003)、大豆(赵丽梅, 2005)和小麦(董玉琛等, 2003; 郝晨阳等, 2008)等各种作物上开展了核心种质构建工作并取得良好效果。甘蔗种质资源方面, 研究人员先后用数量形态学方法构建了甘蔗热带种资源核心库(Balakrishnan et al., 2000)、野生割手密核心库(Tai et al., 2001)和甘蔗杂交品种初级核心库(刘新龙等, 2009)。本研究对斑茅核心种质资源取样策略进行了多层次的分析, 探讨了不同分组原则、不同组内取样比例、不同取样方法、不同总体取样量对核心种质库构建的影响, 并利用秩数加权平均策略构建了斑茅初级核心种质库, 旨在为斑茅野生资源的筛选评价和杂交利用提供科学的理论参考。

1 结果与分析

1.1 加权系数与主效参数确定

巢式方差分析结果见表1, 表中方差贡献率数据均为百分比, 取样策略的方差贡献率为取样方法、分组原则、组内取样比例三者的累计方差贡献率。巢式方差分析结果提示不同的参数在不同的取样层次上受到的影响程度(方差贡献率)是不同的, 在评价各取样策略优劣时应赋予相应的加权。因此, 为放大方差贡献差异和增强可比性, 对同一取样层次不同参数的方差贡献率进行了转换, 使各取样层次的5个参数相对贡献率总和均为100%(横向), 转换后的相对贡献率作为评价该层次取样策略的加权系数(表1括号外数据)。将加权系数大于20%的参数定义为主效参数, 可得以下结果: 取样方法的主效参数为表型方差、多样性指数、表型频率方差和表型保留比例; 分组原则的主效参数为变异系数、表型保留比例和表型频率方差; 组内取样比例的主效参数为变异系数; 取样策略的主效参数为变异系数和表型方差; 总体取样量的主效参数为表型保留比例和多样性指数; 重复取样误差的主效参数为表型方差、表

表 1 各参数在不同取样层次的加权系数

Table 1 Weighting coefficient of parameters for different sampling levels

方差贡献率%	多样性指数	变异系数	表型方差	表型频率方差	表型保留比率
Variance contribution %	I	CV	VPV	VPF	RPR
取样方法	28.74	0.00	28.75	21.72	20.79
Sampling method	(5.81)	(0.00)	(12.09)	(6.92)	(1.10)
分组原则	17.60	22.44	16.96	21.26	21.74
Grouping principle	(6.83)	(33.46)	13.70	13.01	(2.21)
组内取样比例	0.00	93.92	6.08	0.00	0.00
Sampling ratio in group	(0.00)	(15.11)	(0.53)	(0.00)	(0.00)
取样策略(累计方差)	11.41	43.84	23.77	17.99	2.99
Sampling strategy (Cumulative variance)	(12.64)	(48.57)	(26.32)	(19.93)	(3.31)
总体取样量	35.42	4.08	4.27	10.00	46.23
Total sampling ratio	(68.59)	(7.90)	(8.27)	(19.37)	(89.53)
重复取样误差	9.60	22.26	33.45	31.03	3.66
Reapt sampling error	(18.77)	(43.53)	(65.41)	(60.70)	(7.16)

注: I: 多样性指数; CV: 变异系数; VPV: 表型方差; VPF: 表型频率方差; RPR: 表型保留比率; 括号内数值为方差贡献率; 括号外为加权系数(相对方差贡献率)

Note: I: Shannon diversity index; CV: coefficient of variation; VPV: variance of phenotypic value; VPF: variance of phenotypic frequency; RPR: ratio of phenotypic traits retained; In bracket number is variance contribution; out bracket number is weighting coefficient (relative variance contribution)

型频率方差和变异系数。

1.2 秩数加权平均值与取样策略评价

设各参数的原始秩数分别为 I_0 、 CV_0 、 VPV_0 、 VPF_0 和 RPR_0 ，取样方法、分组原则、组内取样比例、取样策略和总体取样量各取样层次的秩数加权平均值分别为 R_m 、 R_g 、 R_i 、 R_s 和 R_t 。根据表 1 中对应的加权系数计算不同取样层次 5 个参数原始秩数的加权平均值，将秩数加权平均值作为该层次取样策略优劣的最终评价参数(计算公式见表 2)。

1.2.1 取样方法评价

取样方法评价结果见表 3，计算公式见表 2 中公式①，表中数据括号内为原始秩数及算术平均值，括号外为加权秩数及加权平均值。取样方法的秩数加权平均值显示聚类取样明显优于随机取样，聚类取样是先根据材料的表型遗传相似度进行归类后再从中选出代表材料，可有效清除遗传冗余，在最大限度上保留原始库的遗传结构。考虑到取样

方法对抽样结果影响显著，故其后分析分为随机和聚类两组进行。

1.2.2 分组原则评价

分组原则评价结果见表 4，参数缩写和数据格式同表 3，计算公式见表 2 中公式②。各分组原则秩数加权平均值大小为：随机取样下，总体聚类分组<采集地分组<叶型分组<不分组<海拔分组<气候区域分组；聚类取样下，叶型分组<海拔分组<总体聚类分组<不分组<采集地分组<气候区域分组。表明在随机取样下，总体聚类分组、采集地分组和叶型分组是较合理的分组原则；在聚类取样下，叶型分组、海拔分组和总体聚类分组是较合理的分组原则。

1.2.3 组内取样比例评价

组内取样比例评价结果见表 5，参数缩写和数据格式同表 3，计算公式见表 2 中公式③。组内取样比例秩数加权平均值大小为：在随机取样下，平方根比例=对数比例=多样性比例<简单比例；在

表 2 不同取样层次的秩数加权平均值计算公式

Table 2 The formula of weighted average rank in different sampling levels

取样层次	秩数加权平均值计算公式
Sampling Levels	The formula of weighted average rank
取样方法	① $R_m = (I_0 \times 28.74\% + CV_0 \times 0\% + VPV_0 \times 28.75\% + VPF_0 \times 21.72\% + RPR_0 \times 20.79\%) / 100\%$
Sampling method	
分组原则	② $R_g = (I_0 \times 17.6\% + CV_0 \times 22.44\% + VPV_0 \times 16.96\% + VPF_0 \times 21.26\% + RPR_0 \times 21.74\%) / 100\%$
Grouping principle	
组内取样比例	③ $R_i = (I_0 \times 0\% + CV_0 \times 93.92\% + VPV_0 \times 6.08\% + VPF_0 \times 0\% + RPR_0 \times 0\%) / 100\%$
Sampling ratio in group	
取样策略	④ $R_s = (I_0 \times 11.41\% + CV_0 \times 43.84\% + VPV_0 \times 23.77\% + VPF_0 \times 17.99\% + RPR_0 \times 2.99\%) / 100\%$
Sampling strategy	
总体取样量	⑤ $R_t = (I_0 \times 35.42\% + CV_0 \times 4.08\% + VPV_0 \times 4.27\% + VPF_0 \times 10\% + RPR_0 \times 46.23\%) / 100\%$
Total sampling ratio	

表 3 取样方法的秩数加权平均值

Table 3 Weighted average rank for sampling method

取样方法	多样性指数	变异系数	表型方差	表型频率方差	表型保留比例	秩数平均值
Sampling method	I	CV	VPV	VPF	RPR	Average rank
随机取样	0.57	0.00	0.58	0.43	0.42	2.00
Random sampling	(2.00)	(2.00)	(2.00)	(2.00)	(2.00)	(2.00)
聚类取样	0.29	0.00	0.29	0.22	0.21	1.00
Cluster sampling	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)

注: I:多样性指数; CV:变异系数; VPV:表型方差; VPF:表型频率方差; RPR:表型保留比率;括号内数值为原始秩数及算术平均值; 括号外为加权秩数及加权平均值

Note: I: Shannon diversity index; CV: coefficient of variation; VPV: variance of phenotypic value; VPF: variance of phenotypic frequency; RPR: ratio of phenotypic traits retained; In bracket number is original rank and arithmetic mean; out bracket number is weighed rank and weighted mean

表 4 分组原则的秩数加权平均值
Table 4 Weighted average rank for grouping principle

取样方法 Sampling method	分组原则 Grouping principle	多样性指数 I	变异系数 CV	表型方差 VPV	表型频率方差 VPF	表型保留比例 RPR	秩数平均值 Average rank	
随机取样 Random sampling	采集地分组 Origin grouping	0.26 (1.50)	0.56 (2.50)	0.42 (2.50)	0.21 (1.00)	0.33 (1.50)	1.79 (1.80)	
	叶型分组 Leaves grouping	0.62 (3.50)	0.22 (1.00)	0.17 (1.00)	0.64 (3.00)	0.33 (1.50)	1.97 (2.00)	
	海拔分组 Height grouping	0.18 (1.00)	0.90 (4.00)	0.59 (3.50)	0.43 (2.00)	0.22 (1.00)	2.31 (2.30)	
	气候区域分组 Zone grouping	0.70 (4.00)	0.67 (3.00)	0.68 (4.00)	0.32 (1.50)	0.43 (2.00)	2.81 (2.90)	
	总体聚类分组 Cluster grouping	0.26 (1.50)	0.45 (2.00)	0.34 (2.00)	0.32 (1.50)	0.33 (1.50)	1.70 (1.70)	
	不分组 Not grouping	0.44 (2.50)	0.67 (3.00)	0.42 (2.50)	0.43 (2.00)	0.33 (1.50)	2.29 (2.30)	
	聚类取样 Cluster sampling	采集地分组 Origin grouping	0.35 (2.00)	0.90 (4.00)	0.51 (3.00)	0.21 (1.00)	0.33 (1.50)	2.30 (2.30)
		叶型分组 Leaves grouping	0.44 (2.50)	0.22 (1.00)	0.34 (2.00)	0.43 (2.00)	0.33 (1.50)	1.75 (1.80)
		海拔分组 Height grouping	0.18 (1.00)	0.90 (4.00)	0.17 (1.00)	0.43 (2.00)	0.22 (1.00)	1.89 (1.80)
		气候区域分组 Zone grouping	0.53 (3.00)	0.45 (2.00)	0.51 (3.00)	0.43 (2.00)	0.43 (2.00)	2.35 (2.40)
总体聚类分组 Cluster grouping		0.35 (2.00)	0.67 (3.00)	0.17 (1.00)	0.43 (2.00)	0.33 (1.50)	1.95 (1.90)	
不分组 Not grouping		0.26 (1.50)	0.79 (3.50)	0.17 (1.00)	0.43 (2.00)	0.33 (1.50)	1.97 (1.90)	

注: 表中参数缩写和数据格式同表 3

Note: Abbreviations and data format as in Table 3

聚类取样下, 对数比例<多样性比例<平方根比例<简单比例。表明在随机取样下, 平方根比例、对数比例和多样性比例都能较好保留群体内的遗传变异, 取样效果明显优于简单比例, 且三种比例间无显著差异; 在聚类取样下, 对数比例取样效果最好, 多样性比例其次, 平方根比例稍差, 但同样明显优于简单比例。

1.2.4 取样策略组合评价

根据表 2 中公式④对取样方法、分组原则、组内取样比例 3 个层次组合形成的 42 种取样策略的优劣进行综合评价, 秩数加权平均值越小表明策略越优。原始库的秩数加权平均值相当于算术平均值, 即 5 个参数的加权系数均为 20%。从表 6 的秩数加权平均值可以看出, 有 2 种取样策略组合明显

优于原始库, 其中叶型分组按对数比例聚类取样(LLC)的策略组合为最佳, 叶型分组按多样性比例聚类取样(LGC)的策略组合其次, 其他的取样策略的加秩数加权平均值均高于原始库。因此, 根据秩数加权平均值的比较分析, 最终确定叶型分组按对数比例聚类取样(LLC)为最佳取样策略。

1.2.5 总体取样量的确定

最佳总体取样量通过比较最优取样策略的 5 个参数在不同总体取样量梯度下的秩数加权平均值来确定, 计算公式见表 2 中公式⑤, 结果见表 7。由表 7 可知, 最佳取样策略 LLC 在不同总体取样量梯度下, 其变异系数、表型方差和表型频率方差均优于原始库, 区别在于表型保留比例和多样性指数两个主效参数。单从多样性指数上看, 总体取样量

表 5 组内取样比例的秩数加权平均值

Table 5 Weighted average rank for sampling ratio within group

取样方法	组内取样比例	多样性指数	变异系数	表型方差	表型频率方差	表型保留比例	秩数平均值
Sampling method	Sampling ratio within group	I	CV	VPV	VPF	RPR	Average rank
随机取样 Random sampling	简单比例	0.00	1.88	0.06	0.00	0.00	1.94
	Simple ratio	(1.00)	(2.00)	(1.00)	(1.00)	(1.00)	(1.20)
	平方根比例	0.00	0.94	0.06	0.00	0.00	1.00
	Square root ratio	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
	对数比例	0.00	0.94	0.06	0.00	0.00	1.00
	Logarithmic ratio	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
	多样性比例	0.00	0.94	0.06	0.00	0.00	1.00
聚类取样 Cluster sampling	Genetic ratio	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
	简单比例	0.00	1.88	0.06	0.00	0.00	1.94
	Simple ratio	(1.00)	(2.00)	(1.00)	(1.00)	(1.00)	(1.20)
	平方根比例	0.00	0.94	0.12	0.00	0.00	1.06
	Square root ratio	(1.50)	(1.00)	(2.00)	(1.00)	(1.00)	(1.30)
	对数比例	0.00	0.94	0.06	0.00	0.00	1.00
	Logarithmic ratio	(1.50)	(1.00)	(1.00)	(1.00)	(1.00)	(1.10)
多样性比例	0.00	0.94	0.09	0.00	0.00	1.03	
Genetic ratio	(2.00)	(1.00)	(1.50)	(1.00)	(1.00)	(1.30)	

注: 表中参数缩写和数据格式同表 3

Note: Abbreviations and data format as in Table 3

在 25% 以上较好。单从表型保留比例上看, 总体取样量 50% 最好, 至少不应低于 30%。综合来看, 总体取样量在 40% 以上较好, 至少不应低于 35%。根据核心种质应以尽可能小的资源份数代表尽可能大的遗传变异的原则, 选择 35% 为最佳总体取样量, 该取样量下多样性指数达 0.6718, 表型保留比例达 94.72%。

通过对最优核心种质库 LLC 35% 丢失表型性状的分析(见表 8), 发现在核心库构建过程中, 有 5 个稀有表型性状极易丢失, 为确保核心库具有最大的表型保留比例, 将分别含有这 5 个稀有表型性状的代表材料各取 1 份纳入核心库, 最终形成的斑茅初级核心库材料总份数为 58 份, 多样性指数提高到 0.683 4, 表型保留比例提高到 100%。

2 讨论

2.1 秩数加权平均值与加权系数

李自超(2003)、张洪亮(2003)、刘新龙(2009)等先后采用计算秩数算术平均值的方法来评价核心种质取样策略的优劣, 构建了水稻、甘蔗等作物的品种初级核心库, 表明秩数分析是评价核心种质取样

策略的有效方法。但秩数算术平均值评价方法的缺陷在于没有考虑到不同取样层次下不同参数产生的影响可能不尽相同。或者说, 这种评价方法是在假设各参数同等重要的前提下进行的, 其评价结果可能掩盖某些取样策略之间细微的差异。针对上述情况, 本研究在前人研究基础上, 引入了加权(权重)的概念, 对不同取样层次的不同参数赋予不同的加权系数, 采用秩数加权平均值作为最终的评价参数。如何计算加权系数才能体现不同参数在不同取样层次的影响大小, 巢式方差分析无疑是一种较好的方法。首先利用巢式方差分析计算各参数的总体方差贡献率; 再对各取样层次的贡献率重新进行横向评估, 得到各参数对于同一取样层次的相对贡献率作为加权系数; 最后根据原始秩数和加权系数计算秩数加权平均值。通过比较原始秩数的算术平均值和加权平均值, 发现加权平均值的优点是可以消除或减弱次要因素影响, 突出主效因子, 对局部细微的差异进行放大, 使评价结果更为客观和科学。在权重相同的情况下, 加权平均值就等于算术平均值, 故加权平均值更具普遍性, 适用范围更广。

表 6 取样策略组合的秩数加权平均值
Table 6 Weighted average rank for different sampling strategy

取样策略 Sampling strategy	多样性指数 I	变异系数 CV	表型方差 VPV	表型频率方差 VPF	表型保留比例 RPR	秩数平均值 Average rank
LLC	0.51 (4.50)	1.10 (2.50)	0.36 (1.50)	0.99 (5.50)	0.07 (2.50)	3.03 (3.30)
LGC	0.74 (6.50)	0.66 (1.50)	0.59 (2.50)	0.99 (5.50)	0.07 (2.50)	3.06 (3.70)
HPC	0.17 (1.50)	2.41 (5.50)	0.36 (1.50)	0.45 (2.50)	0.06 (2.00)	3.45 (2.60)
LGR	0.97 (8.50)	0.44 (1.00)	0.36 (1.50)	1.71 (9.50)	0.07 (2.50)	3.55 (4.60)
LSC	0.40 (3.50)	1.53 (3.50)	0.59 (2.50)	0.99 (5.50)	0.07 (2.50)	3.59 (3.50)
LLR	0.86 (7.50)	0.66 (1.50)	0.36 (1.50)	1.89 (10.50)	0.07 (2.50)	3.83 (4.70)
CGC	0.29 (2.50)	2.41 (5.50)	0.36 (1.50)	0.81 (4.50)	0.07 (2.50)	3.94 (3.30)
CLC	0.29 (2.50)	2.41 (5.50)	0.36 (1.50)	0.99 (5.50)	0.07 (2.50)	4.12 (3.50)
CPC	0.29 (2.50)	2.41 (5.50)	0.36 (1.50)	0.99 (5.50)	0.07 (2.50)	4.12 (3.50)
CSC	0.29 (2.50)	2.85 (6.50)	0.36 (1.50)	0.63 (3.50)	0.07 (2.50)	4.20 (3.30)
ZLC	0.74 (6.50)	1.97 (4.50)	1.07 (4.50)	0.45 (2.50)	0.07 (2.50)	4.31 (4.10)
LSR	0.97 (8.50)	0.88 (2.00)	0.36 (1.50)	2.25 (12.50)	0.07 (2.50)	4.53 (5.40)
CSR	0.63 (5.50)	2.41 (5.50)	0.59 (2.50)	0.99 (5.50)	0.07 (2.50)	4.70 (4.30)
ZPC	0.29 (2.50)	2.85 (6.50)	0.83 (3.50)	0.81 (4.50)	0.07 (2.50)	4.85 (3.90)
LPC	0.29 (2.50)	3.29 (7.50)	0.59 (2.50)	0.63 (3.50)	0.07 (2.50)	4.87 (3.70)
NPC	0.17 (1.50)	3.29 (7.50)	0.36 (1.50)	0.99 (5.50)	0.07 (2.50)	4.88 (3.70)
OLC	0.29 (2.50)	3.29 (7.50)	1.55 (6.50)	0.27 (1.50)	0.07 (2.50)	5.46 (4.10)
HSC	0.17 (1.50)	4.16 (9.50)	0.36 (1.50)	0.99 (5.50)	0.07 (2.50)	5.76 (4.10)
HLC	0.17 (1.50)	4.16 (9.50)	0.24 (1.00)	1.17 (6.50)	0.07 (2.50)	5.82 (4.20)
OSC	0.17 (1.50)	3.29 (7.50)	2.26 (9.50)	0.18 (1.00)	0.07 (2.50)	5.97 (4.40)
ZGC	1.08 (9.50)	1.97 (4.50)	2.26 (9.50)	0.63 (3.50)	0.07 (2.50)	6.02 (5.90)
CPR	0.63 (5.50)	3.29 (7.50)	1.07 (4.50)	0.99 (5.50)	0.07 (2.50)	6.05 (5.10)
OLR	0.51 (4.50)	2.85 (6.50)	2.26 (9.50)	0.63 (3.50)	0.07 (2.50)	6.33 (5.30)
ZSC	0.86 (7.50)	2.41 (5.50)	2.02 (8.50)	0.99 (5.50)	0.07 (2.50)	6.35 (5.90)
OSR	0.40 (3.50)	3.73 (8.50)	1.78 (7.50)	0.63 (3.50)	0.07 (2.50)	6.61 (5.10)
OPC	0.17 (1.50)	5.04 (11.50)	1.31 (5.50)	0.27 (1.50)	0.07 (2.50)	6.86 (4.50)
OGC	0.29 (2.50)	4.16 (9.50)	2.26 (9.50)	0.18 (1.00)	0.07 (2.50)	6.96 (5.00)
CLR	0.74 (6.50)	3.29 (7.50)	2.26 (9.50)	0.81 (4.50)	0.07 (2.50)	7.17 (6.10)
OGR	0.74 (6.50)	3.29 (7.50)	2.73 (11.50)	0.45 (2.50)	0.07 (2.50)	7.29 (6.10)
CGR	0.74 (6.50)	3.73 (8.50)	2.02 (8.50)	1.17 (6.50)	0.07 (2.50)	7.73 (6.50)
HGC	0.17 (1.50)	6.36 (14.50)	0.36 (1.50)	1.17 (6.50)	0.07 (2.50)	8.13 (5.30)
LPR	0.97 (8.50)	3.29 (7.50)	1.78 (7.50)	2.34 (13.00)	0.07 (2.50)	8.45 (7.80)
HPR	0.29 (2.50)	4.60 (10.50)	2.97 (12.50)	0.99 (5.50)	0.07 (2.50)	8.92 (6.70)
ZPR	0.63 (5.50)	5.04 (11.50)	2.26 (9.50)	0.99 (5.50)	0.07 (2.50)	8.99 (6.90)
NPR	0.86 (7.50)	5.04 (11.50)	2.02 (8.50)	1.17 (6.50)	0.07 (2.50)	9.16 (7.30)
ZLR	1.31 (11.50)	4.16 (9.50)	3.57 (15.00)	1.17 (6.50)	0.07 (2.50)	10.29 (9.00)
ZSR	1.20 (10.50)	5.48 (12.50)	3.21 (13.50)	0.99 (5.50)	0.07 (2.50)	10.95 (8.90)
OPR	0.63 (5.50)	6.36 (14.50)	2.97 (12.50)	1.17 (6.50)	0.07 (2.50)	11.20 (8.30)
HGR	0.29 (2.50)	7.01 (16.00)	2.73 (11.50)	1.53 (8.50)	0.07 (2.50)	11.64 (8.20)
HSR	0.29 (2.50)	6.80 (15.50)	3.21 (13.50)	1.35 (7.50)	0.07 (2.50)	11.71 (8.30)
ZGR	1.37 (12.00)	5.92 (13.50)	3.45 (14.50)	0.99 (5.50)	0.09 (3.00)	11.81 (9.70)
HLR	0.40 (3.50)	6.80 (15.50)	2.50 (10.50)	2.07 (11.50)	0.07 (2.50)	11.83 (8.70)
原始库 Original bank	0.20 (1.00)	1.50 (7.50)	1.10 (5.50)	0.30 (1.50)	0.20 (1.00)	3.30 (3.30)

注: 表中参数缩写和数据格式同表 3。取样策略缩写: 第一个字母: O(采集地分组), L(叶型分组), H(海拔分组), Z(气候区域分组), C(总体聚类分组), N(不分组随机); 第二个字母:P(简单比例), S(平方根比例), L(对数比例), G(多样性比例); 第三个字母: C(聚类取样), R(随机取样)

Note: Abbreviations and data format as in Table 3. Abbreviations for sampling strategy: The First letter: O(Original grouping), L(leaves grouping), H(height grouping), Z(zone grouping), C(cluster grouping) and N(not grouping); The second letter: P(simPle ratios), S(square root ratios), L(logarithmic ratios), G(genetic ratios); The third letter: C(cluster sampling), R (random sampling)

表 7 最优取样策略在不同总体取样量下的参数比较

Table 7 Comparison of parameters for different sampling ratio in best sampling strategy

取样策略 Sampling strategy	总体取样量 Sampling ratio	多样性指数 I	变异系数 CV	表型方差 VPV	表型频率方差 VPF	表型保留比例 RPR	秩数平均值 Average rank
LLC	10%	1.42 (4.00)	0.04 (1.00)	0.04 (1.00)	0.10 (1.00)	3.24 (7.00)	4.84 (2.80)
	15%	1.06 (3.00)	0.08 (2.00)	0.06 (1.50)	0.10 (1.00)	2.77 (6.00)	4.08 (2.70)
	20%	1.06 (3.00)	0.04 (1.00)	0.06 (1.50)	0.10 (1.00)	2.77 (6.00)	4.04 (2.50)
	25%	0.71 (2.00)	0.04 (1.00)	0.11 (2.5)	0.10 (1.00)	2.31 (5.00)	3.27 (2.30)
	30%	0.71 (2.00)	0.04 (1.00)	0.11 (2.5)	0.10 (1.00)	2.08 (4.50)	3.04 (2.20)
	35%	0.71 (2.00)	0.04 (1.00)	0.11 (2.5)	0.10 (1.00)	1.62 (3.50)	2.57 (2.00)
	40%	0.53 (1.50)	0.04 (1.00)	0.13 (3.00)	0.10 (1.00)	1.16 (2.50)	1.96 (1.80)
	45%	0.53 (1.50)	0.04 (1.00)	0.11 (2.5)	0.10 (1.00)	1.16 (2.50)	1.93 (1.70)
	50%	0.35 (1.00)	0.04 (1.00)	0.11 (2.5)	0.10 (1.00)	0.92 (2.00)	1.53 (1.50)
原始库 Original bank	100%	0.20 (1.00)	0.60 (3.00)	0.60 (3.00)	0.40 (2.00)	0.20 (1.00)	2.00 (2.00)

注: 表中参数缩写和数据格式同表 3

Note: Abbreviations and data format as in Table 3

表 8 最优核心库 LLC35% 的表型性状丢失情况

Table 8 Phenotypic trait lost in the best pre-core collection LLC35%

表型性状 Phenotypic trait	代码描述 Code discription	总份数 Total Number	丢失表型性状情况及处理 Phenotypic trait lost situation and processing
节间形状 4 Internodes shape 4	圆锥形 Conoidal	2	丢失, 原始库中取样 1 份 Lost, 1 sampled from original pool
节间形状 5 Internodes shape 5	倒圆锥形 Obconoidal	3	部分重复丢失; 未丢失重复里取样 1 份 Lost some replication, 1 sampled from other replication
茎径 10 Salk diameter 10	第 10 组 10 Group	3	部分重复丢失; 未丢失重复里取样 1 份 Lost some replication, 1 sampled from other replication
叶长 9 Lamina length 9	第 9 组 9 Group	3	部分重复丢失; 未丢失重复里取样 1 份 Lost some replication, 1 sampled from other replication
叶宽分组 10 Lamina width	第 10 组 10 Goup	3	部分重复丢失; 未丢失重复里取样 1 份 Lost some replication, 1 sampled from other replication

2.2 取样策略评价与分析

取样方法、分组原则和组内取样比例单独评价的结果表明, 斑茅核心种质较好的取样方法是聚类取样, 较好的分组原则是叶型分组、海拔分组和总体聚类分组, 较好的组内取样比例为对数比例和多样性比例。但方差分析结果表明, 取样方法、分组原则、组内取样比例等因素之间存在互作关系, 因此最好的取样方法、最好的分组原则和最好的组内取样比例不一定能组合出最好的取样策略。鉴于此, 本研究在确定最佳取样策略时, 采用了综合分析的方法, 即先分析 42 种取样策略组合的秩数加权平均值, 寻找实际效果最佳的组合, 再结合方差分析中的主效参数确定最佳组合下的最佳总体取样量, 从

而得到最优的取样结果。

2.3 结论

本文研究结果表明, 在斑茅初级核心种质构建中, 聚类取样优于随机取样。聚类取样中, 叶型分组按对数比例聚类取样(LLC)是分组原则、组内取样比例和取样方法的最佳组合。在最佳总体取样量 35% 的基础上, 将未入选的极值表型、丢失表型和特殊性状(如抗病性)个体加入核心库, 最终形成含 58 份材料的初级核心库, 占总资源的 39.46%, 香农多样性指数 0.6834, 变异系数 37.577 4, 表型方差 0.494 8, 表型频率方差 0.107 9, 表型保留比例 100%, 效率明显优于原始库。

3材料与方法

3.1实验材料与数据标准

以国家甘蔗种质资源圃保存的147份斑茅为原始材料, 根据资源圃数据库系统中每份资源的基本情况(如种质类型、采集地、海拔)、形态特征(如根、茎、叶)和农艺性状(如田间锤度、株高、茎径)等原始数据, 数据规范和标准化整理参照《甘蔗种质资源描述规范和数据标准》(蔡青和范源洪, 2006)和《农作物种质资源鉴定技术规程-甘蔗》(蔡青和范源洪, 2008), 选取性状包括21个质量性状(叶姿、叶色、脱叶性、叶鞘毛群、内叶耳、外叶耳、生长势、茎形、节间形状、曝光前节间颜色、曝光后节间颜色、节间长度、蜡粉带颜色、木栓、生长裂缝、生长带形状、气根、根点排列、芽形、芽位、芽沟)和5个数量性状(株高、茎径、叶长、叶宽、锤度)。对数量性状进行质量标准化处理, 以平均值为基准, ± 0.5 个标准差为间距分为10组。

3.2研究方法

根据前人研究结果(张洪量等, 2003), 本研究选用变异系数(coefficient of variation, CV)、香农多样性指数(Shannon diversity index, I)、表型方差(variance of phenotypic value, VPV)、表型频率方差(variance of phenotypic frequency, VPF)和表型保留比例(ratio of phenotype retained, RPR)5个参数作为斑茅初级核心库的评价指标。从分组原则、组内取样比例、组内取样方法3个层次探讨构建斑茅初级核心种质的最佳取样策略。各参数值进行分组打分, 所得秩数平均值越小的取样策略越好。分组原则有采集地分组(O, 9组)、叶型分组(L, 10组)、海拔分组(H, 10组)、气候区域分组(Z, 5组)、总体聚类分组(C, 3组)及不分组随机(N); 组内取样比例有简单比例(P)、对数比例(L)、平方根比例(S)和多样性比例(G); 组内取样方法分为聚类取样(C)和随机取样(R)。以上3个取样层次组合为42种取样策略。为探讨合适的初级核心库取样量, 设9个总体取样梯度(10%、15%、20%、25%、30%、35%、40%、45%和50%)进行分析比较, 每梯度设3次重复。

斑茅初级核心种质的抽样及5个参数计算在Foxpro系统下编程实现, 聚类分析和Duncan多重检验在SPSS 15.0下进行, Nested巢式方差分析在SAS 8.1系统下进行。取样策略的比较分析采用秩数加权平均的方法, 即先对原始参数进行Nested巢式方差, 分析

各参数在取样方法、分组原则、组内取样比例、总体取样量等水平上各自的总体方差贡献率, 根据总体方差贡献率的大小计算相应的加权系数(相对方差贡献率); 再通过Duncan多重检验方差分析检测不同取样策略下各参数的差异, 秩数值相同表示差异不显著, 秩数值不同表示有显著差异; 最后根据原始秩数和加权系数计算各参数的秩数加权平均值, 秩数加权平均值由小到大排列表示不同取样策略的优劣, 值越小表明该策略越优。

作者贡献

毛钧、刘新龙、苏火生和应雄美是本研究的实验设计和实验研究的执行人; 毛钧、刘新龙及苏火生完成数据分析, 论文初稿的写作; 应雄美、陆鑫及马丽提供实验材料和原始数据; 范源洪和蔡青是项目的构思者及负责人, 指导实验设计, 数据分析, 论文写作与修改。全体作者都阅读并同意最终的文本。

致谢

本研究由云南省应用基础研究计划重点项目(2006C-0013Z)、农业部作物种质资源保护项目(NB2010-2130135-17)、国家人事部留学人员科技活动优秀项目共同资助。作者感谢中国农业大学李志超、张洪亮老师在实验数据分析过程中提供帮助。

参考文献

- Balakrishnan R., Nair N.V., and Sreenivasan T.V., 2000, A method for establishing a core collection of *Saccharum officinarum* germplasm based on quantitative morphological data, *Genet Resour Crop Evol*, 47: 1-9
- Brown A.H.D., 1989, Core collections: A practical approach to genetic resources management. *Genome*, 31: 818-824
- Cai Q., and Fan Y.H., eds., 2006, Descriptors and Data Standard for Sugarcane (*Saccharum officinarum* L.), Agriculture Press, Beijing, China, 2006 (蔡青, 范源洪, 编著, 2006, 甘蔗种质资源描述规范和数据标准, 中国农业出版社, 中国, 北京, pp. 7-31)
- Cai Q., and Fan Y.H., eds., 2008, Technical Code for Evaluating Germ Plasm Resources-Sugarcane (*Saccharum* L.) Agriculture Press, Beijing, China, pp.1-9 (蔡青, 范源洪, 编著, 2008, 农作物种质资源鉴定技术规程-甘蔗, 中国农业出版社, 中国, 北京, pp. 1-9)
- Dong Y.S., Cao Y.S., Zhang X.Y.; Liu S.C., Wang L.F., You G.X.,

- Pang B.S., Li L.H., and Jia J.Z., 2003, Establishment of candidate core collections in Chinese common wheat germplasm, *Journal of Plant Genetic Resources*, 4(1):1-8 (董玉琛, 曹永生, 张学勇, 刘三才, 王兰芬, 游光霞, 庞斌双, 李立会, 贾继增, 2003, 中国普通小麦初选核心种质的产生, *植物遗传资源学报*, 4(1):1-8)
- Frankel O.H., 1984, Genetic perspectives of germplasm conservation, In: Arber W., Llimensee K., Peacock W.J., Starlinger P., eds. *Symposium on Genetic Manipulation: Impact on Man and Society*, Cambridge, UK, Cambridge University Press, pp.161-170
- Hao C.Y., Dong Y.C., Wang L.F., You G.X., Zhang H.N., Ge H.M., Jia J.Z., and Zhang X.Y., 2008, Genetic diversity and construction of core collection in Chinese wheat genetic resources, *Chin Sci Bull*, 53(8): 908-915 (郝晨阳, 董玉琛, 王兰芬, 游光霞, 张洪娜, 盖红梅, 贾继增, 张学勇, 2008, 我国普通小麦核心种质的构建及遗传多样性分析, *科学通报*, 53(8): 908-915)
- Li Z.C., Zhang H.L., Zeng Y.W., Yang Z.Y., Shen S.Q., Sun C.Q., and Wang X.K., 2000, Study on sampling schemes of core collection of local varieties of rice in Yunnan, *Sci. Agric. Sin.*, 33(5): 1-7 (李自超, 张洪亮, 曾亚文, 杨忠义, 申时全, 孙传清, 王象坤, 2000, 云南地方稻种资源核心种质取样方案研究, *中国农业科学*, 33(5): 1-7)
- Li Z.C., Zhang H.L., Cao Y.S., Qiu Z.E., Wei X.H., Tang S.X., Yu P., Wang X.K., 2003, Studies on the sampling strategy for primary core collection of chinese ingenious rice, *Acta Agron. Sin.*, 29(1): 20-24 (李自超, 张洪亮, 曹永生, 裘宗恩, 魏兴华, 汤圣祥, 余萍, 王象坤, 2003, 中国地方稻种资源初级核心种质取样策略研究, *作物学报*, 29(1): 20-24)
- Liu X.L., Cai Q, Ma L., Wu C.W., Lu X., Ying X.M., and Fan Y.H., 2009, Strategy of sampling for pre-core collection of sugarcane hybrid, *Acta Agronomica Sinica*, 35(7): 1209-1216 (刘新龙, 蔡青, 马丽, 吴才文, 陆鑫, 应雄美, 范源洪, 2009, 甘蔗杂交品种初级核心种质取样策略, *作物学报*, 35(7): 1209-1216)
- Rodino A.P., Santalla M., Ron A.M.D., and Singh S.P., 2003, A core collection of common bean from the Iberian peninsula, *Euphytica*, 131:165-175
- Tai P.Y.P., and Miller J.D., 2001, A core collection for *Saccharum spontaneum* L. from the world collection of sugarcane, *Crop Sci.*, 41: 879-885
- Zhang H.L., Li Z.C., Cao Y.S., Qiu Z.E., Yu P., and Wang X.K., 2003, Comparison of parameters for testing the rice core collection in phenotype, *Acta Agron Sin.*, 29(2): 252-257 (张洪亮, 李自超, 曹永生, 裘宗恩, 余萍, 王象坤, 2003, 表型水平上检验水稻核心种质的参数比较, *作物学报*, 29(2): 252-257)
- Zhao L.M., Dong Y.S., Liu B., Hao S., Wang K.J., and Li X.H., 2005, Establishment of a core collection for the Chinese annual wild soybean (*Glycine soja*), *Chin Sci Bull*, 50(10): 989-996 (赵丽梅, 董英山, 刘宝, 郝水, 王克晶, 李向华, 2005, 中国一年生野生大豆核心种质的构建, *科学通报(英文版)*, 50(10): 989-996)



5thPublisher是一个致力于科学与文化传播的中文出版平台

在5thPublisher上发表论文, 任何人都可以免费在线取阅您的论文

- ※同行评审, 论文接受严格的高质量评审
- ※在线发表, 论文一经接受, 即刻在线发表
- ※开放取阅, 任何人都可免费取阅无限使用
- ※快捷搜索, 涵盖谷歌学术搜索与知名数据库
- ※论文版权, 作者拥有版权读者自动授权使用

在线投稿: <http://5th.sophiapublisher.com>