

# 长江春大豆核心种质构建及分析

王丽侠 李英慧 李 伟 朱 莉 关 媛 宁学成  
关荣霞 刘章雄 常汝镇 邱丽娟\*

(中国农业科学院作物科学研究所, 农业部作物种质资源与生物技术重点开放实验室, 北京 100081)

**摘要:** 利用长江春大豆初选核心种质 SSR (simple sequence repeat) 标记和农艺性状表型等基础数据, 对用不同个体取样方法以及不同数据类型建立的核心种质进行评价, 目的是为确定中国大豆 (*Glycine max*) 核心种质的最佳取样策略提供依据。结果表明, 根据 SSR 分子数据聚类, 采用类内随机取样、类内以遗传相似性系数取样以及仅依据遗传相似性系数取样都可用于大豆核心种质构建。但是综合不同评价参数发现, 以类内随机取样最佳, 类内按遗传相似性系数取样次之, 单独以遗传相似性系数取样较差。分析不同 SSR 等位变异保留比例的遗传多样性指数发现, 当保留 90% 和 80% 的 SSR 等位变异时, 核心种质具有更高的遗传多样性。由于与 SSR 分子数据对种质遗传关系评价的不一致性, 农艺性状等基础数据虽然可用来构建核心种质, 但其 SSR 分子水平代表性相对较低。本研究结果还表明, 用不同方法或同一方法不同重复次数取样建立的核心种质具有异质性, 且这种异质性随核心种质取样比例的降低而增大。因此, 虽然可依据不同数据类型确定相应的方法建立核心种质, 但综合表型和分子数据建立的核心种质更具有代表性。

**关键词:** SSR 分子标记, 农艺性状, 聚类分析, 等位变异保留比例, 取样策略

中图分类号: S32      文献标识码: A      文章编号: 1005-0094(2004)06-0578-08

## Establishment of a core collection of Changjiang spring sowing soybean

WANG Li-Xia, LI Ying-Hui, LI Wei, ZHU Li, GUAN Yuan, NING Xue-Cheng, GUAN Rong-Xia, LIU Zhang-Xiong, CHANG Ru-Zhen, QIU Li-Juan

Key Lab of Crop Germplasm Resources and Biotechnology of the Ministry of Agriculture, Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081

**Abstract:** Sampling strategy is very important for core collection establishment. In this study, different methods to construct a core collection of Changjiang spring sowing soybean were performed based on data of SSR (simple sequence repeat) markers and agronomic traits in order to optimize the sampling strategy for core collection. The results showed that, based on cluster analysis of SSR data, each of the three methods, i. e., random sampling within each group, sampling by genetic similarity coefficient within each group and sampling by genetic similarity coefficient, could be used to construct a soybean core collection. When SSR allelic reserving ratios were maintained at 90% and 80%, the core collections had higher genetic diversity indices of SSR alleles than if maintained at 70%. Core collections could also be constructed by agronomic and other basic data if there was no molecular data; however, SSR allelic reserving ratio might be decreased, which suggested that assessment of genetic diversity by SSR data was not always consistent with assessment by agronomic data. We found that core collections were heterogeneous, either when created with different sampling methods or with the same sampling method in different repeats, because randomness always existed in selecting individuals. This indicated that the germplasm to form a core collection was variable, so optimal sampling strategies should be chosen to establish core collections based on different data according to practical or scientific objectives. Integrating phenotypic and genotypic data together would be better for improving the representativeness of core collections.

**Key words:** SSR markers, agronomic traits, cluster analysis, reserving ratio of SSR alleles, sampling strategy

中国是大豆(*Glycine max*)的起源地,国家作物种质资源库收集保存的栽培大豆遗传资源达 23 000 余份,居世界之首(常汝镇等, 1998)。大豆核心种质的建立不仅有利于对大豆资源的深入研究,也有助于优异品种资源的发现和利用。资源总体的遗传多样性评价是建立核心种质的主要依据。由于物种遗传结构和遗传学研究进度不同,目前核心种质中以农艺性状(Grenier *et al.*, 2001; Lasa *et al.*, 2001; Hari *et al.*, 2001)、地理来源(Tai & Miller, 2001)、系谱关系(Martynov *et al.*, 2003)、同工酶(Chandra *et al.*, 2002; Ottosson *et al.*, 2002)等为依据进行遗传多样性评价。然而,上述指标的遗传变异有限,使我们很难对大量资源作深入的分析与评价;且如果仅对表型性状等基础数据作分析,也不能从基因组角度对种质资源进行全面评价。鉴于 SSR(simple sequence repeat)分子标记具操作简单、共显性和反映信息量多的特点,可有效用于大豆种质鉴定和遗传多样性研究(许占友等, 1999; Yong *et al.*, 1999; Narvel *et al.*, 2000; Abe *et al.*, 2003);再加上此项技术操作日趋简化,使高通量分析成为可能。因此,SSR 标记被用于中国栽培大豆核心种质构建研究。

目前核心种质的构建方法多为分区聚类取样(崔艳华等, 2003),但在聚类的基础上如何进行个体选择,还未见报道。本研究利用长江春大豆初选核心种质 SSR 分子标记数据,以聚类分析为基础,进行类内不同个体取样方法的研究,目的是为中国其他生态区域核心种质个体取样提供一定的依据;同时针对大豆的物种特性,对以不同基础数据建立核心种质进行了可行性分析,以期为大豆核心种质构建提供参考。

## 1 材料与方法

### 1.1 材料与农艺性状数据

长江春大豆的 184 份初选核心样本是根据农艺性状以分层聚类法,从长江春大豆总体 2148 份大豆种质资源中取样所得(邱丽娟等, 2003),分别来自四川、湖北、湖南、安徽、江苏、江西和浙江等 7 省份,覆盖 118 个行政县,跨越北纬 25°–32°共 7 个纬度区。选用的 14 个农艺性状包括生育日数、百粒重、株高、粒色、脐色、子叶色、茸毛色、花色、生长习性、结荚习性、粒型、叶形、粗蛋白和粗脂肪含量,这些表

型数据均来自《中国大豆品种资源目录》(王国勋, 1982; 常汝镇和孙建英, 1991; 常汝镇等, 1996)。

### 1.2 SSR 标记分析

用 SDS 法从大豆种子中提取 DNA(关荣霞等, 2003);60 个核心 SSR 位点是谢华等(2003)和王彪等(2003)分别利用 80 份秋大豆和 190 份初选核心种质中的随机样本鉴定和筛选出来的。这些 SSR 位点平均分布在 20 个大豆遗传连锁图上,由 Cregan 等(1999)开发并定位。PCR 扩增条件为:95℃ 变性 5 min,然后以 95℃ 30 s,47℃ 30 s,72℃ 30 s 进行 35 个循环反应,72℃ 延伸 5 min,4℃ 保存。PCR 扩增产物用 6% 的聚丙烯酰胺凝胶电泳分离,以银染法染色。为保证不同胶板间扩增产物统计的准确性,将每板凝胶检测的不同等位变异用 MegaBACE1000 测序仪进行片断分析。依据等位变异的片断大小,将所有样品的 SSR 指纹合并,用于下一步分析。

### 1.3 数据处理

将 SSR 指纹图谱进行数据转换(有带记为 1,无带记为 0),利用统计软件 NTSYSp2.10 中的 UPGMA 法计算品种间遗传相似性系数。差异显著性 *t*-检验用 SAS 软件进行,其余运算均在 Excel 表格中完成。

### 1.4 核心种质构建方法

#### 1.4.1 根据 SSR 分子标记数据聚类构建

方法 1:根据聚类结果,以每类均保留一份种质为原则,随机选取;

方法 2:根据聚类结果,以每类均保留一份种质为原则,类内按种质间遗传相似性系数取样;

方法 3:根据聚类结果,完全按种质间的遗传相似性系数取样,即将遗传相似性系数较大的种质组合分别随机删除一份后,剩余种质再聚类;再从遗传相似性系数较大的成对种质中随机删除一份。以此类推,直到代表性或核心样本量达到要求。

随机法:依据聚类结果,完全随机重复取样。

以上方法均以不同 SSR 等位变异保留比例(100%, 90%, 80%, 70%)进行分析。为减少随机取样误差,每个等位变异保留比例均随机重复取样 2 次,计算核心种质各评价参数时,取其平均值。其中 100% 等位变异保留比例是将种质数压缩到极限,即再删除任何一份种质都会造成等位变异的丢失。

1.4.2 根据农艺性状表现型等基础数据构建

性状群法(ABC法):以大豆品种分类的主要农艺性状种皮色(A)、成熟期(B)和粒大小(C)组成的性状群为依据,每性状群随机取样一份构成核心种质。

地理来源法:以不同纬度来源为依据,每0.1个纬度区随机选取一定数量的种质构成核心种质。

遗传距离法:以14个主要农艺性状计算品种间遗传距离,将遗传距离较小的种质分别随机删除一份后,剩余种质重新计算相互间遗传距离;继续删除遗传距离较小的种质。以此类推,直至核心种质数量达到要求。

随机法:完全随机重复取样。

其中,地理来源法、遗传距离法和随机法的取样量分别参考性状群法,使4种方法的取样量相近。为降低误差,每种方法分别重复取样2次,取其平均值。

1.4.3 核心种质的评价参数

SSR等位变异保留比例(R):

$$R(\%) = \frac{\text{核心种质总等位变异数}}{\text{初选核心种质总等位变异数}} \times 100\% \tag{1}$$

SSR等位变异遗传多样性指数及表型遗传多样性指数(PIC):

$$PIC = 1 - \sum P_i^2 \tag{2}$$

其中 $P_i$ 为某一变异类型的频率,PIC主要反映SSR变异类型的丰富度与不同等位变异分布的均匀度。

平均遗传相似性系数:根据SSR指纹数据以非加权平均法(UPGMA)计算。

平均表型符合度:即所有质量性状符合度( $N_c$ )和数量性状符合度( $N_u$ )的平均值,其中:

$$N_c(\%) = \frac{\text{核心种质中某性状的表型变异数}}{\text{初选核心种质中某性状表型变异数}} \times 100\% \tag{3}$$

$$N_u(\%) = \frac{\text{核心种质某性状极差}}{\text{初选核心种质某性状极差}} \times 100\% \tag{4}$$

性状群和地理来源(县)覆盖率:分别为核心种质种的性状群数或地理来源数占总体性状群数或地理来源数的比例。

特异等位变异比例(U):

$$U(\%) = \frac{\text{特异等位变异数}}{\text{核心种质总等位变异数}} \times 100\% \tag{5}$$

其中特异等位变异是指分析群体中只有一份种质能够检测到的等位变异。

2 结果与分析

2.1 根据SSR分子标记数据建立核心种质

由表1可看出,当SSR等位变异保留比例降至70%时,核心种质压缩比例最低为1.44%,而平均表型符合度也可达74.60%,这远远超出Brown(1989)曾经提出的10%的样品代表70%的遗传多样性的预测。由于不同参数对核心种质评价的侧重点存在差异,本研究中的取样方法在不同评价参数方面也各有优势。比如,特异等位变异比例和平均遗传相似性系数均以方法3的种质间遗传异质性最大,随机法最小;平均遗传多样性指数和表型符合度均以方法1最高,随机法、方法3较低;而平均性状群和地理来源(县)覆盖率以随机法最高,方法3最低。差异显著性 $t$ -检验表明,除方法1平均遗传多样性指数显著大于随机法,平均表型符合度显著大于方法3外,其余评价参数差异均不显著,这可能是由长江春大豆资源群体的遗传背景相对较近导致。不同评价参数的综合分析表明,方法1和方法2优于方法3和随机法,其中前二者差异不显著;由于方法2比方法1取样过程复杂、繁琐,故对于资源量较大群体的核心种质构建,方法1即类内随机取样简单可行。

在不同SSR等位变异保留比例水平上,利用方法1即类内随机取样方法建立核心种质的资源比例与平均表型符合度、性状群覆盖率和地理覆盖率均呈显著正相关( $r=0.99, 0.99, 0.99$ )。这是由所构建核心种质资源的压缩比例过大而造成的:因为长江春大豆初选资源的分布比较分散,184份种质分布于118个县,包含52类品种性状群,仅9个质量性状的表型变异就达43种,而80%SSR等位变异保留比例的核心种质数量为49份,70%SSR等位变异保留比例的核心种质仅为31份。此外,核心种质间的平均遗传相似性系数随取样比例的降低呈显著降低趋势( $r=0.97$ )。随SSR等位变异保留比例的压缩,遗传多样性指数(PIC)也呈下降趋势( $PIC_{90\%} > PIC_{80\%} > PIC_{70\%}$ )。 $t$ -检验结果表明, $PIC_{90\%}$ 与 $PIC_{80\%}$ 无显著差异,但显著高于 $PIC_{70\%}$ ,而 $PIC_{80\%}$ 与 $PIC_{70\%}$ 也无显著性差异。这说明保留90%和80%SSR等位变异,均可在一定程度上降低某些SSR等位变异重复保存的频率,此时核心种质

表 1 在不同 SSR 等位变异保留比例水平上以遗传相似性系数法和随机法建立核心种质的比较  
Table 1 Comparison among core collections constructed by similarity coefficient and random selection at different SSR allelic reserving ratios

评价参数 Evaluation parameter	取样方法 Sampling method	等位变异保留比例 Reserving ratio of SSR alleles				
		100%	90%	80%	70%	平均值 Average
种质数 Number of accessions	方法 1 Method 1	106	77	49	31	65.75
	方法 2 Method 2	103	84	49	32	67.00
	方法 3 Method 3	105	79	45	32	65.25
	随机法 Random	108	100	67	39	78.50
资源比例(%) Ratio of total accessions (%)	方法 1 Method 1	4.93	3.58	2.28	1.44	3.06
	方法 2 Method 2	4.80	3.91	2.28	1.49	3.12
	方法 3 Method 3	4.89	3.68	2.09	1.49	3.04
	随机法 Random	5.03	4.66	3.12	1.82	3.66
特异等位变异比例 Ratio of unique alleles (%)	方法 1 Method 1	27.30	24.56	28.12	34.22	28.55
	方法 2 Method 2	26.76	21.75	28.54	34.79	27.96
	方法 3 Method 3	26.50	21.86	33.94	37.95	30.06
	随机法 Random	26.03	22.35	24.04	30.80	25.81
平均遗传相似性系数 Mean genetic similarity coefficient	方法 1 Method 1	0.211	0.200	0.195	0.193	0.20
	方法 2 Method 2	0.210	0.207	0.203	0.196	0.204
	方法 3 Method 3	0.213	0.205	0.201	0.134	0.188
	随机法 Random	0.216	0.221	0.227	0.227	0.223
平均遗传多样性指数 (PIC) Mean genetic diversity index (PIC)	方法 1 Method 1	0.783	0.789	0.785	0.777	0.784
	方法 2 Method 2	0.782	0.783	0.782	0.777	0.781
	方法 3 Method 3	0.780	0.787	0.777	0.768	0.778
	随机法 Random	0.778	0.772	0.765	0.756	0.768
平均表型符合度(%) Mean correspondence of phenotypes (%)	方法 1 Method 1	99.52	90.95	85.19	82.21	89.47
	方法 2 Method 2	93.92	91.20	87.10	80.70	88.23
	方法 3 Method 3	97.49	85.97	82.29	74.80	85.14
	随机法 Random	94.91	91.01	80.20	74.60	85.18
性状群覆盖率 Coverage of ABC groups (%)	方法 1 Method 1	76.92	59.62	50.00	38.46	56.25
	方法 2 Method 2	78.85	69.23	48.08	36.54	58.18
	方法 3 Method 3	69.23	60.34	53.85	40.39	55.95
	随机法 Random	76.92	63.79	56.74	40.39	59.46
地理来源覆盖率 Coverage of geographic regions (%)	方法 1 Method 1	62.71	48.31	34.75	24.58	42.59
	方法 2 Method 2	63.56	54.66	34.32	23.31	43.96
	方法 3 Method 3	63.56	48.30	31.36	23.73	41.74
	随机法 Random	62.71	62.00	46.62	29.24	50.14

方法 1：在聚类的基础上类内随机选取；方法 2：在聚类的基础上，类内按遗传相似性系数取样；方法 3：仅按照遗传相似性系数取样；随机法：不依据聚类结果的完全随机取样。  
Note: Method 1 means sampling randomly within group based on cluster result; method 2 means sampling according to genetic similarity coefficient based on cluster result; method 3 means sampling according to genetic similarity coefficient merely, and random selection means sampling randomly without any cluster result.

取样比例分别为 3.58% 和 2.28%。因此,本研究将以类内随机取样(方法 1)在 SSR 等位变异保留比例为 90% 和 80% 时构建的 77 份和 49 份核心种质作为长江春大豆核心种质备用,以满足不同研究和利用的需求。

2.2 根据农艺性状等基础数据建立核心种质

由表 2 可看出,根据农艺性状等基础数据以不同方法建立核心种质的资源比例都低于 3%,而平

均表型符合度最低为 80.35%,SSR 等位变异保留比例最低为 73.77%。这说明在初选核心种质的基础上,依据农艺性状等基础数据建立核心种质是可行的。如果将性状群法除外,性状群覆盖率以遗传距离法最高,地理来源法次之,随机法最低。如果将地理来源法除外,地理覆盖率以性状群法最高,遗传距离法次之,随机法最低。平均表型符合度以遗传距离法最高,性状群法次之,随机法最低。但是 SSR

表 2 根据农艺性状等基础数据以不同方法建立核心种质的比较  
Table 2 Comparison of core collections constructed by different methods based on agronomic and other basic data

取样方法 Sampling strategy	性状群法 ABC groups	地理来源法 Geographic origin	遗传距离法 Genetic distance	随机法 Random selection
种质数 Number of accessions	52	54	54	53
资源比例(%) Ratio of total accessions(%)	2.42	2.51	2.51	2.47
SSR 等位变异保留比例 Reserving ratio of SSR alleles(%)	73.77	76.7	75.77	79.16
特异等位变异比例 Ratio of unique alleles(%)	27.69	25.60	28.47	31.16
平均表型符合度 Mean correspondence of phenotypes(%)	85.87	82.07	89.87	80.35
性状群覆盖率 Coverage of ABC groups(%)	100.00	53.85	59.62	50.00
地理覆盖率(县) Coverage of geographic regions(%)	46.0	55.0	44.0	43.5
平均遗传相似性系数 Mean genetic similarity coefficient	0.241	0.231	0.226	0.222
平均遗传多样性指数 Mean genetic diversity index (PIC)	0.745	0.757	0.762	0.764

性状群法指依据大豆种皮色(A)、成熟期(B)和粒大小(C)组成的性状群取样;地理来源法指依据纬度来分区,在不同纬度内分别取样;遗传距离法是指根据主要农艺性状计算的种质间遗传距离取样;随机法则指不依据聚类结果进行的完全随机取样。  
ABC groups means sampling according to seed color(A), maturation period(B) and seed size(C); geographic origin means sampling based on groups divided according to latitude; genetic distance means sampling based on genetic distance according to 14 main agronomic traits; random selection means sampling randomly without any evaluation data.

等位变异保留比例、特异等位变异比例以及平均遗传多样性指数都以随机法最高, *t*-检验表明性状群法的遗传多样性指数极显著低于其他 3 种方法(*P* < 0.01)。

由以上结果看出,根据农艺性状和 SSR 分子标记数据对种质间遗传关系的评价结果并不完全一致。依据不同农艺性状取样和随机取样的结果也说明,总体资源经过初选核心种质压缩以后,种质间已存在较大的遗传差异,再次依据农艺性状压缩的可能性很小。因此,在初选核心种质的基础上,若没有分子评价数据,可采用随机法进行压缩。虽然随机取样在表型符合度方面低于据农艺性状再次压缩的结果,但其 SSR 分子水平的代表性较高。

根据农艺性状表型数据以不同方法建立的核心种质间其共有种质数各不相同(表 3)。相比之下,同种方法不同重复次数取样的共有种质比例较高,其中性状群法两次重复取样的共有种质数最高(26 份),但也仅为核心种质数(52 份,表 2)的一半。这说明即使取样方法相同,由于取样过程中存在随机选取的可能性,使得不同重复次数取样建立的核心种质存在异质性。因此可根据实际需要,选取合适

方法建立核心种质,再根据不同的需求随时补充特异种质资源。由于表型变异分布的不均匀性,即使以农艺性状表型数据建立核心种质,也会丢失一些稀有表型的种质。因此,应通过表型代表性的评价,适当补充与完善特异种质,以提高核心种质的实用性和代表性。

### 3 讨论

Brown(1989)提出核心种质是在样本数不低于 3000 份时,以 10% 的样本代表总体资源 70% 以上的遗传多样性。目前已报道的 30 余个物种的核心种质研究中,资源总量达 10 000 份以上的核心种质比例变化较小,多在 10% 左右;资源量较少时,取样比例变化范围较大,为 6.54% – 29.04%(崔艳华等, 2003)。相比之下,用于核心种质构建的资源总量低于本研究数量的约有 20 余个物种,其中荷兰大麦(*Hordeum vulgare*)的核心种质取样比例最低,为 6.54%。本研究中长江春大豆总资源 2148 份,根据农艺性状初步压缩为 184 份,占 8.58%;而根据 SSR 分子标记,在保证高于 70% 遗传代表性时的取样比例最低可达 1.44%。可见核心种质的取样比例变

表 3 根据农艺性状表型数据以不同方法及不同重复次数取样建立的核心种质之间的共有种质数  
Table 3 Number of common germplasms between collections by different sampling methods

取样方法 Sampling methods	性状群法 1 ABC group 1	性状群法 2 ABC group 2	地理来源法 1 Geographic origin 1	地理来源法 2 Geographic origin 2	随机法 1 Random 1	随机法 2 Random 2
性状群法 2 ABC group 2	26					
地理来源法 1 Geographic origin 1	24	18				
地理来源法 2 Geographic origin 2	15	16	24			
随机法 1 Random 1	15	17	13	14		
随机法 2 Random 2	9	12	14	16	16	
遗传距离法 Genetic distance	20	19	17	13	13	16

性状群法 1 和性状群法 2 是指分别依据大豆种皮色(A)、成熟期(B)和粒大小(C)组成的性状群进行的两次不同取样;地理来源法 1 和地理来源法 2 是依据纬度来分区,在不同纬度内分别进行的两次不同的取样;遗传距离法是指根据主要农艺性状计算的种质间遗传距离取样;随机法 1 和随机法 2 则指不依据聚类结果进行的两次不同的完全随机取样。  
ABC groups 1 and ABC groups 2 mean two different samplings according to seed color(A), maturation period(B) and seed size(C); geographic origin 1 and geographic origin 2 mean two different samplings based on groups divided according to latitude; genetic distance means sampling based on genetic distance according to 14 main agronomic traits; randomly selection 1 and random selection 2 mean two different random samplings without any evaluation data.

化范围很大,在以农艺性状表型数据压缩的基础上,根据 SSR 分子标记可以大幅度提高核心种质的压缩比例和代表性。此外,由于不同物种的遗传结构和基因组遗传重复程度存在差异,利用 DNA 分子标记构建核心种质,取样比例也可能存在差异。因此,相对中国大豆其他生态类型(黄淮春大豆除外)来说,本研究中长江春大豆的地理分布和农艺性状表型变异范围都较小,推测其他生态类型的核心种质取样比例相对要高一些。也就是说,取样比例会因不同大豆生态资源群体的遗传变异程度而有所变化。

本研究发现,利用 SSR 数据聚类取样,类内随机取样和类内按遗传相似性系数取样均优于单独按照遗传相似性系数取样和随机取样。这说明 SSR 数据的聚类结果真实地反映了种质间的遗传关系,按照类别可以有效地依据遗传关系远近将种质分开。由于核心种质取样面临的资源量较大,考虑到取样过程的工作量,类内按遗传相似性系数取样相对比较繁琐,因此,类内随机取样为最佳选择。也就是说,根据所需要的核心资源份数,直接将总体资源聚成相应类别,每类随机选取一份即可组成所需核心种质。

种皮色、成熟期和粒大小是大豆品种分类的主要参考性状,本研究根据大豆品种分类中的 ABC 性

状群建立核心种质,虽然资源量比用 SSR 分子数据压缩有所增大,但也仅为 2.42%。另外地理来源法、遗传距离法(农艺性状表型数据聚类)也可用于核心种质构建,但是 SSR 分子水平代表性有所降低。

可见在构建大豆核心种质时,先利用农艺性状等基础数据聚类对大豆总体资源进行压缩后,再根据 SSR 分子标记以类内随机取样进行分析,不仅可以降低核心种质的取样比例,还具有较大的分子和表型水平的代表性。而这种由农艺性状到 SSR 分子标记分析逐步建立核心种质的方法和最初提出大豆核心种质的构建策略是一致的(邱丽娟等,2003)。

不同方法甚至于同一方法不同重复次数取样构建的核心种质之间具有异质性,这种异质性随着取样比例的降低而变大,即共有种质数减少,说明同一资源群体中能够入选核心种质的种质不是固定的。这是因为任何取样方法中都会存在一定的随机性,例如根据性状群取样,只要某一性状群的种质不唯一,其中的任何种质都有入选的可能;而且,同一性状群的种质数越多,则由所有性状群入选种质的组合数就越多,即核心种质间存在差异的可能性就越大。同样,依据聚类结果取样,也会存在随机性,例如采用类内遗传相似性系数取样,对于遗传相似性

系数最大的种质的删除就有随机性,而这种随机性,将导致下一轮聚类中种质的遗传关系改变,从而导致核心种质的异质性。相对来说,随机取样的不固定性更大,任何种质都有入选的机会。而依据排列组合的理论,核心种质比例越小,每份种质在不同次取样过程中被选中的几率就越低。

关于核心种质的评价体系还缺乏统一认识。李自超等(1999)和张洪亮等(2003)认为对核心种质代表性检测时表型保留比例是不可缺少的检测指标,表型频率方差、遗传多样性指数、表型方差、变异系数等也是评价核心种质的有效参数。作者认为基于核心种质概念出发,只要能包括大部分的表型变异类型即可,而各种表型变异的分布结构不一定要与总体资源结构完全一致,尤其是表型变异分布频率差异比较大的物种。因此,本研究中对不同核心种质的农艺性状代表性评价时,仅仅从表型的符合度方面进行了考虑。

另外,研究中还发现,以不同方法建立核心种质的表型代表性差异主要体现在一些表型变异极不均匀的性状上,如稀有的粒色、子叶色、叶形和粒形等方面。无论以农艺性状还是以 SSR 分子标记数据建立核心种质都难以避免这些特异表型的丢失。因此,构建核心种质时,综合农艺性状和 SSR 分子数据对整体资源进行分析,能较全面地反映基因组信息,并有效提高核心种质的代表性和实用性。

参考文献

Abe, J., Xu, D. H., Suzuki, Y., Kanazawa, A. and Shimamoto, Y. 2003. Soybean germplasm pools in Asia revealed by nuclear SSRs. *Theoretical and Applied Genetics*, **106**: 445 – 453.

Brown, A. H. D. 1989. Core collection: a practical approach to genetic resources management. *Genome*, **31**: 818 – 824.

Chandra, S., Huaman, Z., Hari Krishna, S. and Ortiz, R. 2002. Optimal sampling strategy and core collection size of Andean tetraploid potato based on isozyme data — a simulation study. *Theoretical and Applied Genetics*, **104**: 1325 – 1334.

Chang, R. Z. (常汝镇) and Sun, J. Y. (孙建英). 1991. *Catalogues of Chinese Soybean Germplasm and Resources: Continuation I* (中国大豆品种资源目录 续编一). China Agricultural Press, Beijing. (in Chinese)

Chang, R. Z. (常汝镇), Sun, J. Y. (孙建英) and Qiu, L. J. (邱丽娟). 1998. Progress on Chinese soybean germplasm

resources collection. *Crops* (作物杂志), **3**: 7 – 9. (in Chinese)

Chang, R. Z. (常汝镇), Sun, J. Y. (孙建英), Qiu, L. J. (邱丽娟) and Chen, Y. W. (陈一舞). 1996. *Catalogues of Chinese Soybean Germplasm and Resources: Continuation II* (中国大豆品种资源目录 续编二). China Agricultural Press, Beijing. (in Chinese)

Cregan, P. B., Jarvik, T., Bush, A. L., Shoemaker, R. C., Lark, K. G., Kahler, A. L., Kaya, N., VanToai, T. T., Lohnes, D. G., Chung, J. and Specht, J. E. 1999. An integrated genetic linkage map of the soybean genome. *Crop Science*, **39**: 1464 – 1490.

Cui, Y. H. (崔艳华), Qiu, L. J. (邱丽娟), Chang, R. Z. (常汝镇) and Lü, W. H. (吕文河). 2003. Advance of research on core collection of plant germplasm resources. *Journal of Plant Germplasm Resources* (植物遗传资源学报), **4**: 279 – 284. (in Chinese with English abstract)

Grenier, C., Hamon, P. and Bramel-Cox, J. P. 2001. Core collection of sorghum. II. Comparison of three random sampling strategies. *Crop Science*, **41**: 241 – 246.

Guan, R. X. (关荣霞), Chang, R. Z. (常汝镇) and Qiu, L. J. (邱丽娟). 2003. Rapid isolation of soybean DNA for SSR analysis. *Soybean Science* (大豆科学), **22**: 73 – 74. (in Chinese with English abstract)

Hari, D. U., Paula, J. B., Rodomiro, O. and Sube, S. 2001. Development of a chickpea core subset using geographic distribution and quantitative traits. *Crop Science*, **41**: 206 – 210.

Lasa, J. M., Igartua, E., Ciudad, F. J., Codesal, P., Garcia, E. V., Gracia, M. P., Medina, B., Romagosa, I., Molina-Cano, J. L. and Montoya, J. L. 2001. Morphological and agronomical diversity patterns in the Spanish barley core collection. *Hereditas*, **135**: 217 – 225.

Li, Z. C. (李自超), Zhang, H. L. (张洪亮), Sun, C. Q. (孙传清) and Wang, X. K. (王象坤). 1999. Studies and prospects of core collection in plant germplasm resources. *Journal of China Agricultural University* (中国农业大学学报), **4**(5): 51 – 62. (in Chinese with English abstract)

Martynov, S. P., Dobrotvorskaya, T. V., Dotlacil, L., Stehno, Z., Faberova, I. and Bares, I. 2003. Genealogical approach to the formation of the winter wheat core collection. *Russian Journal of Genetics*, **39**: 917 – 923. (in Russian with English abstract)

Narvel, J. M., Fehr, W. R., Chu, W. C., David, G. and Shoemaker, R. C. 2000. Simple sequence repeat diversity among soybean plant introductions and elite genotypes. *Crop Science*, **40**: 1452 – 1458.

Ottosson, F., Von Bothmer, R. and Diaz, O. 2002. Genetic variation in three species of *Hordeum*, and the selection of

- accessions for the Barley Core Collection. *Hereditas*, **137**: 7 – 15.
- Qiu, L. J. (邱丽娟), Cao, Y. S. (曹永生), Chang, R. Z. (常汝镇), Zhou, X. A. (周新安), Wang, G. X. (王国勋), Sun, J. Y. (孙建英), Xie, H. (谢华), Zhang, B. (张博), Li, X. H. (李向华), Xu, Z. Y. (许占有) and Liu, L. H. (刘立宏). 2003. Establishment of Chinese soybean (*Glycine max*) core collection. I. Sampling strategy. *Scientia Agricultura Sinica*(中国农业科学), **36**: 1442 – 1449. (in Chinese with English abstract)
- Tai, P. Y. P. and Miller, J. D. 2001. A core collection for *Saccharum spontaneum* L. from the world collection of sugarcane. *Crop Science*, **41**: 879 – 885.
- Wang, B. (王彪), Chang, R. Z. (常汝镇), Tao, L. (陶莉), Guan, R. X. (关荣霞), Yan, L. (阎丽), Zhang, M. H. (张明恢), Feng, Z. F. (冯忠孚) and Qiu, L. J. (邱丽娟). 2003. Identification of SSR primer numbers for analyzing genetic diversity of Chinese cultivated soybean. *Molecular Plant Breeding*(分子植物育种), **1**: 82 – 88. (in Chinese with English abstract)
- Wang, G. X. (王国勋). 1982. *Catalogues of Chinese Soybean Germplasm and Resources* (中国大豆品种资源目录). China Agricultural Press, Beijing. (in Chinese)
- Xie, H. (谢华), Chang, R. Z. (常汝镇), Cao, Y. S. (曹永生), Zhang, M. H. (张明恢), Feng, Z. F. (冯忠孚) and Qiu, L. J. (邱丽娟). 2003. Selection of core SSR loci by using Chinese autumn soybean. *Scientia Agricultura Sinica* (中国农业科学), **36**: 360 – 366. (in Chinese with English abstract)
- Xu, Z. Y. (许占友), Qiu, L. J. (邱丽娟), Chang, R. Z. (常汝镇), Li, X. H. (李向华), Zheng, C. M. (郑翠明), Liu, L. H. (刘立宏) and Guo, P. (郭蓓). 1999. Using SSR markers to evaluate genetic diversity of soybean in China. *Scientia Agricultura Sinica*(中国农业科学), **32**: 40 – 48. (in Chinese with English abstract)
- Yong, C. K., Hoon, K. J., Song, H. and Kim, N. 1999. Genetic diversity measured by simple sequence repeat variations among the wild soybean collected along riverside of five major rivers in Korea. *Genes and Genetic Systems*, **74**: 169 – 177.
- Zhang, H. L. (张洪亮), Li, Z. C. (李自超), Cao, Y. S. (曹永生), Qiu, Z. E. (裘宗恩), Yu, P. (余萍) and Wang, X. K. (王象坤). 2003. Comparison of parameters for testing the rice core collection in phenotype. *Acta Agronomica Sinica*(作物学报), **29**: 252 – 257. (in Chinese with English abstract)

(责任编辑: 周玉荣)