

中国大豆 (*Glycine max*) 核心种质构建

I. 取样方法研究

邱丽娟¹, 曹永生¹, 常汝镇¹, 周新安², 王国勋², 孙建英¹,
谢 华¹, 张 博¹, 李向华¹, 许占有¹, 刘立宏¹

(¹ 中国农业科学院作物品种资源研究所/农业部作物种质资源与生物技术重点开放实验室, 北京 100081;

² 中国农业科学院油料作物研究所, 武汉 500082)

摘要: 以 23 587 份中国栽培大豆为试验材料, 根据农艺性状, 用 20 种方法构建了大豆初级核心种质, 对 3 种分层法、3 种确定取样数法和 2 种个体选择法进行了比较, 明确了栽培大豆核心种质构建的适宜取样方法和取样比例。不同取样方法与总体都进行了品种分类数、各性状符合度、数量性状平均数、各性状多样性指数方差和平均品种距离共 5 个指标的比较分析, 结果表明, 多层次取样方法(品种分类法)对总体的代表性优于二层次或一层次取样法, 按比例和平方根确定取样数方法对总体的代表性优于多样性指数法, 聚类选择的方法对总体的代表性优于随机选择方法。在 20 种方法不同取样比例条件下, 方法 17 的平均品种距离降低幅度大于方法 15。因此, 利用品种分类法进行分层, 用比例法确定取样数目, 根据聚类结果进行个体选择的方法 15 是构建大豆初级核心种质的最佳方法。用最佳方法构建的初级核心种质, 比较不同取样比例的平均品种距离, 确定品种平均距离由缓慢降低到明显降低点(9.0%)为适当的取样比例。

关键词: 中国; 大豆; 核心种质; 取样方法

S565 A

Establishment of Chinese Soybean (*G. max*) Core Collection

I. Sampling Strategy

QIU Li-juan¹, CAO Yong-sheng¹, CHANG Ru-zhen¹, ZHOU Xin-an², WANG Guo-xun²,
SUN Jian-ying¹, XIE Hua¹, ZHANG Bo¹, LI Xiang-hua¹, XU Zhan-you¹, LIU Li-hong¹

(¹ Institute of Crop Germplasm Resources, Chinese Academy of Agricultural Sciences, Key Lab of Ministry of Agriculture, Crop Germplasm and Biotechnology, Beijing 100081; ² Institute of Oil Crops, Chinese Academy of Agricultural Sciences, Wuhan 500082)

Abstract: In this paper, 23 587 cultivated soybeans were analyzed by using agronomic characteristics in order to find out optimal sampling strategy and sample size for developing soybean core collection in China. The 20 sampling strategies were used, 18 of them consisted of 3 stratifications, 3 determinations of sampling number and 2 selections of individuals methods, and the rest 2 were checks. The criteria against whole collection include variety classification, coincidence of 14 characters, mean of 5 quantitative characters, variance of genetic diversity, mean of variety distance. The results indicated that variety classification was better than methods of one or two stratifications, the proportional sample determination was better than methods of square root or genetic diversity, and clustering selection of sample was better than random selection. Therefore, method 15 would be optimal strategy for developing primary Chinese soybean core collection. By using this method, various samples with different proportions of whole collection were compared for mean variety distance, and the sample size of 9% were taken as the best proportion for construction primary core collection because

收稿日期: 2003-03-05

基金项目: 国家重点基础研究发展规划项目“973”“大豆核心种质构建”课题(G1998010203)

作者简介: 邱丽娟(1963-), 女, 研究员, 博士, 博士生导师, 主要从事大豆品种资源研究, Tel: 010-62186650; Fax: 010-62186629; E-mail: qiulijuan@

caas.net.cn

it could keep the variation with the least varieties.

Key words: China; Soybean; Core collection; Sampling methods

大豆起源于中国,据记载已有 5 000 多年的栽培历史,在中国各地都有种植。目前中国国家作物种质资源库已收集保存栽培大豆品种资源 2 万余份,居世界各国大豆品种资源之首。这些资源是大豆品种改良的物质基础,如何深入地对其进行评价和利用是现今的研究焦点。核心种质(core collection)概念的提出和发展^[1,2]为资源研究提供了一条新的思路。到目前为止,国内外已先后构建了小麦^[3]、大麦^[4]、高粱^[5]、花生^[6]、多花菜豆^[7]、甘蔗^[8]等作物的核心种质。Brown^[1,2]认为,根据等位基因频率有两种取样策略,一是在没有遗传材料的任何资料的情况下,采用完全随机取样法;再者是根据研究物种的资料情况,选用恒数、比例或对数取样法的分层取样策略构建核心种质。然而,已构建的核心种质,多采取的是分层取样策略。其共同点是根据物种的基本数据,如来源地、物种分类、生态类型等对整个收集品进行分组,在分组的基础上采用聚类分析和主成分分析取样。究竟哪种取样方法更合理,因不同物种的考虑因素不同,结论不一。

笔者以中国收集保存的大豆种质资源为基础材料,利用 3 种不同分层方法、3 种确定取样数法和 2 种样本选择方法,组合成 18 种不同的构建大豆初选

核心种质的方法,另外,以完全随机法和人工取样法作对照,通过对 20 种不同方法进行比较,确定构建大豆初级核心种质的最佳取样方法,目的是为建立中国大豆核心种质以及大豆资源的深入研究和利用等提供理论依据。

1 材料与方法

1.1 材料

中国国家种质资源库中已保存栽培大豆种质资源 23 587 份。根据国家作物种质资源数据库中的农艺性状数据对这些品种资源进行取样方法研究。

1.2 方法

初级核心种质的组成方法采用了分层-聚类法。聚类分析所用的 15 个性状包括类型、种皮色、茸毛色、花色、结荚习性、叶形、生育习性、粒形、子叶色、脐色、百粒重、生育日数、株高、蛋白质含量和油分含量。首先对质量性状数据进行规范化(表 1),对数值型性状以 0.5 个标准差为间距分为 10 个等级,利用 S-PLUS 开发的大方法聚类分析软件 GEBEI 进行混合聚类分析。样品间距离采用欧氏距离(Euclidean distances),类间距离采用改进的离差平方和法(Ward's method)。

表 1 大豆非数值型性状赋值

Table 1 Coden designed for quantative traits in soybean

性状	Character	赋值	Coden of quantative traits
类型	Planting type	春 = 1, 夏 = 2, 秋 = 3	Spring = 1, Summer = 2, Autumn = 3
种皮色	Seed coat color	黄 = 1, 绿 = 2, 黑 = 3, 褐 = 4, 双 = 5	Yellow = 1, Green = 2, Black = 3, Brown = 4, Bicolor = 5
子叶色	Cotyledon color	黄 = 1, 绿 = 2	Yellow = 1, Green = 2
脐色	Hilum color	黄 = 1, 淡褐 = 2, 褐 = 3, 深褐 = 4, 灰 = 5, 蓝 = 6, 黑 = 7	Yellow = 1, Light brown = 2, Brown = 3, Dark brown = 4, Gray = 5, Blue = 6, Black = 7
粒形	Seed shape	圆 = 1, 扁圆 = 2, 椭圆 = 3, 扁椭圆 = 4, 长椭圆 = 5, 肾形 = 6	Round = 1, Oblate = 2, Ellipse = 3, Flat-ellipse = 4, Long-ellipse = 5, Kindey = 6
生长习性	Growth habit	直立 = 1, 半直 = 3, 半蔓 = 5, 蔓生 = 7	Erect = 1, Sem-erect = 3, Semi-trail = 5, Trail = 7
结荚习性	Podding habit	有限 = 3, 亚有限 = 5, 无限 = 7	Determinate = 3, Semi-determinate = 5, Indeterminate = 7
茸毛色	Pubescence color	灰毛 = 1, 棕毛 = 2	Grey = 1, Brown = 2
花色	Flower color	白花 = 1, 紫花 = 2	White = 1, Purple = 2
叶形	Leaf shape	披针 = 1, 椭圆 = 2, 卵圆 = 3, 圆 = 4	Lanceolate = 1, Ellipse = 2, Ovate = 3, Round = 4

将每个方法都与总体进行不同指标的比较,根据代表总体程度的高低排列秩序。检测的指标包括品种分类数、性状平均数和方差、性状遗传多样性指数和方差、性状符合度、平均品种距离。

有关计算公式如下:

1.2.1 多样性指数取样

$$N_k = \frac{(\bar{H}'_k \times m)}{\sum_{i=1}^m \bar{H}'_i} \quad (1)$$

$$\bar{H}' = - \left(\sum_{i=1}^n \sum_{j=1}^1 P_{ij} \ln P_{ij} \right) / n$$

其中, N_k : 第 k 个类别取样数; N : 总品种数; \bar{H}'_k : 第 k 个类别的平均多样性指数; \bar{H}'_i : 第 i 个类

别的平均多样性指数; m : 类别总数; n : 性状总数; l : 代码总数; P_{ij} : 某一类别中第 i 个性状第 j 个代码出现的频率。

1.2.2 平方根法取样

$$N_j = \frac{\sqrt{n_j}}{\sum_{i=1}^n \sqrt{n_i}} \times N \quad (2)$$

其中, N_j : 第 j 个类别取样数; n_j : 第 j 个类别品种总数; n : 类别数; n_i : 第 i 个类别品种数; N : 总品种数。

1.2.3 显著性测验 多样性指数 t 测验

$$H' = - \sum P_i \ln P_i;$$

$$\text{Var}(H') = [\sum (P_i \ln P_i)^2 - (\sum P_i \ln P_i)^2] / N + (n-1) / (2 \times N \times N) \quad (3)$$

$$t = (H'_1 - H'_2) / [\text{Var}(H'_1) + \text{Var}(H'_2)]^{1/2};$$

$$m = [\text{Var}(H'_1) + \text{Var}(H'_2)]^2 / [\text{Var}^2(H'_1) / N_1 + \text{Var}^2(H'_2) / N_2]$$

其中, H' 、 H'_1 、 H'_2 : 某性状的多样性指数; P_i : 某性状第 i 个代码出现的频率; n : 某性状的代码数; N 、 N_1 、 N_2 : 品种数; m : 自由度; $\text{Var}(H')$ 、 $\text{Var}(H'_1)$ 、 $\text{Var}(H'_2)$: 方差; t : 实测 t 值。

χ^2 测验和 F 测验用 SAS 软件完成。

1.2.4 平均品种距离(D)

$$D = \frac{\sum_{i=1}^n \sum_{j=1}^{n-1} \sqrt{\sum_{k=1}^m (x_{ik} - x_{jk})^2}}{N \times (N-1) / 2} \quad (4)$$

其中, x_{ik} : 第 i 个品种第 k 个性状的值; m : 性状数; N : 品种数。

2 结果与分析

2.1 用 20 种取样方法组成的样本

笔者共用 20 种方法构建了初级核心种质。首先分三个层次: 第一层是根据中国大豆的耕作栽培制度将全国大豆产区分为北方、黄淮和南方 3 个栽培区(简称栽培区); 第二层是在栽培区内根据播种期分为北方春大豆型、黄淮春大豆型、黄淮夏大豆型、长江春大豆型、南方春大豆型、南方夏大豆型和秋大豆型(简称种植型), 或者按行政省进行分区; 第三层是按种皮色(黄、绿、黑、褐、双)、生育期(极早熟、早熟、中早熟、中熟、晚熟、迟熟和极迟熟)、百粒重(大粒: ≥ 18 g; 中粒: 12 ~ 17.9 g; 小粒: ≤ 11.9 g) 3 个性状对品种分群(简称性状群)^[9]。通过不同分层; 将品种分组, 在此基础上, 样本数的确定包括比例法、平方根法和多样性指数法 3 种方法。确定

出取样数后, 分别采用随机或聚类的方法选择个体。这样总共有 18 种取样方法。此外, 设完全随机法(方法 1)和人工选择法(方法 20)为对照。用 20 种方法共建立了 20 个样本(表 2), 每个样本占总体的 15% 左右, 变化范围为 15.8% (方法 17) ~ 16.2% (方法 7)。

2.2 不同取样方法的比较

2.2.1 不同取样方法的品种分类代表性比较 中国栽培大豆的生态区划与影响大豆分布的热量、降水量和光周期这 3 个主要气候条件因素有着密切的关系。按照品种分类系统, 可将中国 23 587 份大豆品种分成 490 群。将 20 种方法构建的初级核心种质的品种分类数与总体品种分类比较(表 3)看出, 方法 14、15、16、17、18、19 的分类均为 490, 对总体分类的代表性为 100%; 而其它方法的品种分类仅占总体的 65.7% ~ 89.8%, 以方法 7 的代表性最低为 65.7%。完全随机取样方法与总体的代表性不是最低, 而人工取样方法对总体的代表性也不是最高, 这两个对照方法(1 和 20)的代表性分别为 75.7% 和 87.6%。对与总体有差别的一层和二层的其它 12 个方法比较, 在相同个体选择法前提下, 比较 3 种确定样本数确定法对总体的代表性, 以平方根法 > 比例法 > 多样性指数法。而在相同确定样本数法条件下, 1 对以随机(方法 6)选择法的代表性比聚类(方法 7)选择略高(4.1%), 其余 5 对对总体的代表性均以聚类方法比随机法略高 2.0% ~ 11.0%。

2.2.2 不同取样方法各性状符合度分析 根据 14 个性状的代码进行符合度分析, 除方法 8 的株高符合度为 90.9% 外, 不同方法的平均符合度差异均来自种皮色的差异, 与总体完全符合的方法为 14、15、11、7、16、17、18 和 19。方法 8 的性状符合度最低, 为 98.1%, 其余样本的符合度为 98.6%。

2.2.3 不同取样方法各性状遗传多样性方差分析

不同方法各性状多样性指数方差都大于总体方差, 差异极显著。根据核心种质的方差与总体差异越大越有代表性的标准, 按照方差, 从大到小进行排序。方法 3 和 11 对总体的代表性最好, 都排在第 1 位; 其次为方法 5, 而方法 12 对总体的代表性最差。这 4 种方法的排序号分别为 1, 1, 2 和 18。在相同分层及确定取样数条件下, 聚类选择好于随机选择。

2.2.4 不同取样方法数量性状平均值的代表性分析 各方法的数量性状平均值与总体性状平均值进行显著性 t 测验, 发现方法 1、8 和 14 的 5 个性状与总体差异均不显著, 表明这 3 个方法对总体的代表

表 2 大豆初级核心种质构建的 20 种不同方法

Table 2 The methods used to construct 20 primary core collections

方法 Method	取样策略 Strategy	分层 Stratification			取样数设定 Determination of sampling number	个体选择 Selection of individuals	样本容量 Total sampling size
		第一层 First grade	第二层 Second grade	第三层 Third grade			
1	随机 Random	-	-	-	比例法 Ratio method	随机 Random	3773
2	分层 Stratification	栽培区 Cultivation region	-	-	比例法 Ratio method	随机 Random	3770
3	分层 Stratification	栽培区 Cultivation region	-	-	比例法 Ratio method	聚类 Cluster	3770
4	分层 Stratification	栽培区 Cultivation region	-	-	平方根法 Square root method	随机 Random	3774
5	分层 Stratification	栽培区 Cultivation region	-	-	平方根法 Square root method	聚类 Cluster	3774
6	分层 Stratification	栽培区 Cultivation region	-	-	多样性指数法 Polymorphic index method	随机 Random	3811
7	分层 Stratification	栽培区 Cultivation region	-	-	多样性指数法 Polymorphic index method	聚类 Cluster	3811
8	分层 Stratification	栽培区 Cultivation region	省 Province	-	比例法 Ratio method	随机 Random	3750
9	分层 Stratification	栽培区 Cultivation region	省 Province	-	比例法 Ratio method	聚类 Cluster	3750
10	分层 Stratification	栽培区 Cultivation region	省 Province	-	平方根法 Square root method	随机 Random	3764
11	分层 Stratification	栽培区 Cultivation region	省 Province	-	平方根法 Square root method	聚类 Cluster	3760
12	分层 Stratification	栽培区 Cultivation region	省 Province	-	多样性指数法 Polymorphic index method	随机 Random	3786
13	分层 Stratification	栽培区 Cultivation region	省 Province	-	多样性指数法 Polymorphic index method	聚类 Cluster	3786
14	分层 Stratification	栽培区 Cultivation region	播种型 Planting type	性状群 Characters	比例法 Ratio method	随机 Random	3769
15	分层 Stratification	栽培区 Cultivation region	播种型 Planting type	性状群 Characters	比例法 Ratio method	聚类 Cluster	3769
16	分层 Stratification	栽培区 Cultivation region	播种型 Planting type	性状群 Characters	平方根法 Square root method	随机 Random	3748
17	分层 Stratification	栽培区 Cultivation region	播种型 Planting type	性状群 Characters	平方根法 Square root method	聚类 Cluster	3718
18	分层 Stratification	栽培区 Cultivation region	播种型 Planting type	性状群 Characters	多样性指数法 Polymorphic index method	随机 Random	3761
19	分层 Stratification	栽培区 Cultivation region	播种型 Planting type	性状群 Characters	多样性指数法 Polymorphic index method	聚类 Cluster	3761
20	分层 Stratification	栽培区 Cultivation region	播种型 Planting type	性状群 Characters	比例法 Ratio method	人工 Artificial	3810

表 3 20 种不同初级核心种质方法的品种分类总数及其占总体分类的百分数¹⁾

Table 3 Numbers of variety classification and percentage of the total collection in 20 methods of primary core collections in soybean

方法 Method	品种分类 Variety classification		排序 Rank	方法 Method	品种分类 Variety classification		排序 Rank
	Total number	Total collection (%)			Total number	Total collection (%)	
1	371	75.7	12	11	428	87.3	4
2	372	75.9	11	12	386	78.7	9
3	416	84.9	5	13	415	84.7	6
4	401	81.4	7	14	490	100	1
5	440	89.8	2	15	490	100	1
6	347	70.8	13	16	490	100	1
7	322	65.7	14	17	490	100	1
8	375	86.5	10	18	490	100	1
9	416	84.5	5	19	490	100	1
10	389	79.4	8	20	429	87.6	3

¹⁾ 根据品种分类占总体百分数的高低从小到大排序

Ranking from small to large based on the classification number from high to low percentage of the total collection

性最好。与总体只有 1 个性状存在显著差异的有方法 6、15、16、17 和 18,这 5 个方法的代表性较好。与总体有 2 个性状存在差异的有方法 4、7 和 12。方法 13 有 3 个性状与总体存在差异;而 5、9、10、20 这 4 个方法与总体有 4 个性状存在差异。在随机选择的条件下,不同分层的 3 种样本数设定法对总体的代表性均为比例法 > 多样性 > 平方根法;而在聚类选择条件下,一次分层和二次分层的 3 种样本数设定法比较,以多样性指数法为最好,三次分层的品种分类以比例法 = 平方根法 > 多样性法。在相同分层和样本数设定条件下,对有差别的随机选择和聚类选择进行比较,均以随机选择法与总体差异小,而聚类选择法对总体的代表性差。

2.2.5 不同方法品种间平均距离的比较 方法内品种之间的平均距离是衡量品种相似性高低的重要指标之一。通过对 20 种方法比较,可以看出(表 4),方法 17 的品种间平均距离最大为 6.17,居 20 种方法之首。品种之间平均距离在 6.00 以上的方法有 10 个,从大到小排列的顺序依次为方法 16 > 6 > 8 > 20,7 > 3 > 9 > 15,5 > 19,其余 7 个方法按品种平均距离从大到小依次为 1 > 14 > 11 > 18 > 8 > 2 > 4,以方法 10 的品种间平均距离最小,为 5.89。平均品种距离越大,表明初选核心种质对总体的代表性越好。因此方法 17 的取样代表性最好,方法 10 的代表性最差。

在相同个体选择条件下,一次分层和二次分层法构成的初级核心种质的品种平均距离从大到小的排列顺序为多样性指数法 > 比例法 > 平方根法,而在三次分层中,方法品种的平均距离从大到小的排列顺序为平方根法 > 比例法 > 多样性指数法。在相同分层及样本数设定条件下,均以聚类选择的平均品种间距离大于随机选择的平均品种距离。

2.2.6 不同取样方法的综合比较 根据分类数、性状符合度、性状多样性指数方差、平均品种平均距离

这 5 个指标对总体的代表性高低从小到大分别排序(表 5)。可以看出,方法 17 的总体代表性最好,其次为方法 15 和 14,再次为方法 3、16 和 19。而方法 8 和 12 的代表性最差,对照方法 1 和 20 的总体代表性居中,分别排在第 11 位和第 7 位。

利用排序结果对不同分层法、取样数确定法和个体选择法进行综合分析。无论是 3 种确定取样数方法还是二种选择方法,对总体的代表性从高到低的趋势为三层法(品种分类) > 一层法(栽培区) > 二层法(栽培区 + 省),但只有用多样性指数确定样本时,三层法(品种分类) > 二层法(栽培区 + 省) > 一层法(栽培区)。比较 3 种取样数确定方法对总体的代表性,总的变化规律是比例法 > 平方根法 > 多样性指数法。与分层和取样数不同,比较 2 种个体选择法时,在 3 种不同分层法和 3 种不同样本数确定法中,均表现为聚类选择对总体的代表性优于随机选择。

根据 5 个指标的综合分析,方法 15 是构建核心方法比较好的方法,但从总的排序结果看,方法 17 排在第 1 位,方法 15 排在第 2 位。将这 2 种方法进行不同取样比较,2 种方法都表现为随着样本数减少,品种间平均距离下降,当取样比例在 7% 左右品种平均距离降至最低;当取样数继续减少时,品种平均距离则呈上升趋势,当取样数与品种分类数相等时,品种平均距离达到最大。然而,当取样比例从 16% 降至 7% 的过程中,方法 17 的品种平均距离减少大于方法 15,这是由于个体数多的类别内有差别品种锐减所造成的。因此,按方法 15 构建核心种质时,能在减少样本同时保持品种间的最大距离。

2.3 不同取样量的比较

用方法 15 构建的核心种质,在取样比例从 10% 降到 7% 时,品种平均距离大幅度降低。为此,在这 2 个取样比例中间又增加了 2 个取样比例 9.20% 和 8.86%,分析表明其品种平均距离分别为 6.00 和

表 4 20 种方法的品种间平均距离及排序¹⁾

Table 4 The mean value of variety distances (MVDD) for 20 methods

方法 Method	品种间 平均距离 MVDD	排序 Rank									
1	5.97	10	6	6.11	3	11	5.95	12	16	6.13	2
2	5.91	15	7	6.05	5	12	5.96	11	17	6.17	1
3	6.03	6	8	5.92	14	13	6.08	4	18	5.93	13
4	5.90	16	9	6.02	7	14	5.96	11	19	6.00	9
5	6.01	8	10	5.89	17	15	6.01	8	20	6.05	5

¹⁾ 根据平均距离的远近从小到大排序

Ranking from large to small based on mean variety distance

表 5 20 种方法不同指标的综合比较¹⁾

Table 5 Comprehensive rank of different criteria for 20 methods in soybean

方法 Method	品种分类 Variety	符合度 Coincidence	数量性状平均数 Mean of quantitative character	品种平均距离 Mean of variety distance	多样性指数方差 Variance of genetic diversity index	总计 Total	排序 Rank
1	12	2	1	10	11	71	11
2	11	2	1	15	13	87	14
3	5	2	8	6	3	50	4
4	7	2	4	16	12	87	14
5	2	2	8	8	1	51	5
6	13	2	3	3	16	80	13
7	14	1	5	5	8	74	12
8	10	3	1	14	15	90	15
9	5	2	7	7	6	55	6
10	8	2	8	17	5	87	14
11	4	1	9	12	2	64	9
12	9	2	4	11	17	91	16
13	6	2	6	4	14	68	10
14	1	1	1	11	10	49	3
15	1	1	3	8	8	45	2
16	1	1	3	2	14	50	4
17	1	1	3	1	4	30	1
18	1	1	2	13	10	60	8
19	1	1	4	9	9	50	4
20	3	2	8	5	7	57	7

¹⁾ 根据累计和低高排出从小到大的位次

Ranking from low to large based on the total number

表 6 按方法 15(最佳方法)不同取样比例的平均品种距离

Table 6 Mean value of variety distance for different samples based on method 15

序号 No.	品种数 Number of variety	占总体百分数 Total collection(%)	平均品种距离 Mean of variety distance
1	3 741	15.86	6.01
2	3 064	12.99	6.01
3	2 397	10.16	6.00
4	2 170	9.20	6.00
5	2 047	8.68	5.98
6	1 944	7.33	5.95
7	1 728	4.67	5.96
8	1 102	3.13	5.97
9	738	0.82	6.02

5.98。比较不同取样比例的样本,随着取样比例从 15.86% 下降到 3.83%,平均品种距离一直呈下降趋势,但降到 0.82% 时,平均品种距离又迅速提高,超过 6.00 而达到最大。因此,笔者确定初级核心种质的取样比例为 9.2%,此时样本容量为 2 170 份。

3 讨论

3.1 大豆分层的依据

中国大豆栽培区划已有近 60 年的历史,期间由于农业耕作制度的变革进行了补充和修订。早在 1943 年,王金陵^[10]根据“中国土地利用调查”等统计

资料,将全国大豆栽培划分为 5 大区域,包括春作大豆区,夏作大豆冬闲区,夏作大豆区,秋作大豆区和大豆两获区,为大豆的栽培区划奠定了基础。解放后,卜慕华和潘铁夫^[11]将中国大豆栽培区域分为北方春作大豆、黄淮流域二作大豆区和南方多作大豆区。研究表明,大豆栽培区划与影响我国大豆分布的热量、降水量和光周期这 3 个主要气候条件因素有着密切的关系。因此,在构建栽培大豆核心种质时首先考虑的分层因素就是大豆栽培区划。这与其它植物构建核心种质所考虑的土壤、纬度等因素是相似的。

栽培大豆是由野生大豆驯化而来的,诸多性状都存在差异,在大豆分类中,王金陵^[12]提出以种皮色、开花期、种粒大小对大豆进行分类。然而,与大豆开花期相比,大豆的生育期生态类型特点在大豆育种、引种及生产中具有更重要意义^[13]。王金陵等^[14]用来自全国的 24 个品种进行光照反应,试验结果表明,同一地区不同播种期的大豆品种,光照阶段存在差异。笔者在卜慕华和潘铁夫^[11]划分的 3 个大豆栽培区基础上,按播种期类型分为北方春大豆,东北春大豆,黄淮春大豆,黄淮夏大豆,南方春大豆,南方秋大豆,南方夏大豆 7 个型,在区和型内按种皮色、生育期、百粒重分群,共分为 490 群。这种分类是品种对光温反应的特点不同所产生的^[9]。笔者在比较不同比例方法时发现,当按品种分类数从每类中选择 1 份品种时,品种之间的平均距离最大,这也证明品种分类能够反映品种之间的差异。

3.2 核心种质取样方法

Yongezawa 等^[15]比较 5 种构建核心种质的取样策略,认为一般情况下按比例法是最佳方法。他们认为,如果在已知多样性的前提下,根据遗传多样性范围按比例进行分层是最好的方法。笔者对 20 种方法的综合比较分析也表明,用比例法确定取样数目是最佳的方法,且随着方法分层的增加,比例法对总体的代表性越好。值得注意的是,在分层少(一层)时,用比例法确定方法数构建核心种质没有平方根法和多样性指数法好,对总体的代表性从高到低的排列顺序平方根法 > 多样性指数法 > 比例法。通过对 3 种分层方法进行比较可以看出,利用栽培生态区(一层)和品种分类(三层)这些与大豆相关特性进行分层法构建核心种质,比用行政区划(二层)法构建核心方法更能代表总体。虽然,在综合排序时,由方法 17 构建的方法列在方法 15 之前,居第一位,但二者同时按不同取样比例构建方法时可以发现,前者平均品种距离随方法数减少而较后者显著下降,这表明按平方发根法确定取样数,对分层后组内样本数尤其是遗传差异大的种质进行锐减,在删除相似品种的同时也删除了差别大的品种。相比之下,按比例法取样则能在缩减种质的同时较好地保持品种间的平均差异。然而,从育种者选择亲本考虑,方法 17 以保持品种分类间的变异为优先,能最大的删除类内的变异,如果从中选择亲本进行育种,可有效地拓宽新品种的遗传基础。另外,排序较高的方法(2, 4, 6, 7, 8, 10, 12)对育种者选用差异大的亲本,也具有指导作用。

尽管已建立不同作物的核心种质,但构建核心种质的方法因资源总体已知数据的多寡、作物的特性差异不同而不同^[16]。本研究利用多个指标构建大豆初选核心种质,正在利用筛选^[17]和验证^[18]的 SSR 核心位点进行分析。

References

- [1] Brown A H D. Core collections: a partial approach to genetic resources management. *Genome*, 1989, 31:818 - 824.
- [2] Brown A H D. The core collection at the crossroads. In: Hodgkin T, Brown AHD, Hitum THL, Morales EAV. *Core Collection of Plant Genetic Resources*. International Plant Genetic Resources Institute (IPGRI), A Wiley - Sayce Publication, 1995:3 - 19.
- [3] Spagnoletti Zueli P L, Qualset C O. Evaluation of five strategies for obtaining a core subset from a large genetic resources collection of durum wheat. *Theoretical and Applied Genetics*, 1993, 87:295 - 304.
- [4] Knupffer H, Van Hintum Th J L. The barley core collection: an international effort. In: Hodgkin T, Brown A H D, Van Hintum Th J L, Morales E A V. (ed) *Core Collections of Plant Genetic Resources*. John Wiley & Sons, West Sussex, England, 1995:171 - 178.
- [5] Greiner C, Hamonb P, Bramel C P G. Colletion of sorgum II. Comparison of three random sampling strategies. *Crop Sciences*, 1994, 41:241 - 246.
- [6] Holbrook C C, Anderson W F, Pittman R N. Selection of a core collection from the U. S. germplasm collection of peanuts. *Crop Sciences*, 1993, 33:859 - 861.
- [7] Tohme J, Jones P, Beebe S, Iwanaga M. The combined use of agroecological and characterization data to establish the CIAT Phaseolus vulgaris core collection. In: Hodgkin T, Brown A H D, Van Hintum Th J L, Morales E A V. (ed) *Core Collections of Plant Genetic Resources*. John Wiley & Sons, West Sussex, England, 1995:95 - 108.
- [8] Tai P Y P, Miller J D. A core collection for *Saccharam spontaneum* L from the world collection of sugarance. *Crop Sciences*, 2001,41:879 - 885.
- [9] 王国勋. 中国栽培大豆分类的原则、系统、性状依据及其标准. 张子金. 中国大豆育种与栽培. 北京:农业出版社:1987, 180 - 214.
Wang G X. Classifying principles, models, factors and criteria of variety classification for Chinese cultivated. Zhang Z J. *Soybean Breeding and Cultivation in China*. Beijing: Agricultural Press, 1987: 180 - 214. (in Chinese)
- [10] 王金陵. 中国大豆栽培区域划分之初步研讨. 农报, 1943, 8(25 - 30):282 - 286.
Wang J L. Preliminary studies on regionlization of cultivated soybean in China. *Agricultural Bulletin*, 1943, 8(25 - 30): 282 - 286. (in Chinese)
- [11] 卜慕华, 潘铁夫. 中国栽培区域探讨. 大豆科学, 1982, 1 (2): 105 - 121.

- Pu M H, Pan T F. A study on the regionalization of soybean production area in China. *Soybean Science*, 1982, 1(2): 105 - 122. (in Chinese)
- [12] 王金陵. 大豆性状之演化. 农报, 1947, 12(5): 6 - 11.
Wang J L. Evolution of soybean cultivars. *Agricultural Bulletin*, 1947, 12(5): 6 - 11. (in Chinese)
- [13] 王国勋. 论我国南北大豆生育期生态类型及在引种工作中的应用. 大豆科学, 1982, 1(1): 33 - 40.
Wang G X. Discussion about the ecological type of duration and its use in introduction of soybean (*Soja max* L.) in China. *Soybean Science*, 1982, 1(1): 33 - 40. (in Chinese)
- [14] 王金陵, 武镛祥, 吴和礼, 孙善澄. 中国南北地区大豆光照生态类型的分析. 农业学报, 1956, 7(2): 169 - 180.
Wang J L, Wu Y X, Wu H L, Sun S C. Analysis on photoperiod ecotypes of cultivated soybean in various latitudes in China. *Agricultural Sciences*, 1956, 7(2): 169 - 180. (in Chinese)
- [15] Yongezawa K, Nomura T, Morishima H. Sampling strategies for use in stratified germplasm collections. In: Hodgkin T. (ed) *Core Collections of Plant Genetic Resources*. John Wiley & Sons, West Sussex, England. 1995: 95 - 108.
- [16] 崔艳华, 邱丽娟, 常汝镇, 吕文河. 植物核心种质研究进展. 植物遗传资源学报, 2003, (2): 84 - 90.
Cui Y H, Qiu L J, Chang R Z, Lu W H. Advances in the core collection of plant germplasm resources. *Plant Germplasm Resource*, 2003, (2): 84 - 90. (in Chinese)
- [17] 谢 华, 邱丽娟, 常汝镇, 曹永生, 张明辉, 冯忠孚. 利用中国秋大豆 [*Glycine max* (L.) Merr] 筛选 SSR 核心位点的研究. 中国农业科学, 2003, 36(4): 360 - 366.
Xie H, Qiu L J, Chang R Z, Cao Y S, Zhang M H, Feng Z F. Selection of core SSR loci by using Chinese autumn soybean [*Glycine max* (L.) Merr]. *Scientia Agricultura Sinica*, 2003, 36(4): 360 - 366. (in Chinese)
- [18] 王 彪, 常汝镇, 陶 莉, 关荣霞, 闫 丽, 张明恢, 冯忠孚, 邱丽娟. 分析中国栽培大豆遗传多样性所需 SSR 引物的数目. 分子植物育种, 2003, 1(1): 82 - 88.
Wang B, Chang R Z, Tao L, Guan Z X, Yan L, Zhang M H, Feng Z F, Qiu L J. Identification of SSR primer numbers for analyzing genetic diversity of Chinese soybean cultivated soybean. *Molecular Plant Breeding*, 2003, 1(1): 82 - 88. (in Chinese)

(责任编辑 孙雷心)