

DOI: 10.3724/SP.J.1006.2010.02099

## 苧麻核心种质构建方法

栾明宝 陈建华\* 许 英 王晓飞 孙志民

中国农业科学院麻类研究所 / 农业部茎纤维生物质与工程微生物重点开放实验室, 湖南长沙 410205

**摘 要:** 选用国家长沙苧麻圃的 790 份种质资源, 在已有的 25 个性状数据的基础上, 采用不同取样方法、不同系统聚类方法、不同遗传距离方法构建苧麻核心种质, 用质量性状的多样性指数均值、表型保留比率均值和数量性状均值差异百分率、方差差异百分率、极差符合率和变异系数变化率等 6 个指标评价不同方法组合(取样方法、聚类方法、遗传距离)构建核心种质的优劣。选出合适的构建方法, 构建苧麻核心种质。结果表明, 不同的取样方法对质量性状和数量性状的遗传多样性影响不同, 就质量性状的最大遗传多样性而言, 选择优先取样+多次聚类随机取样方法比较适宜, 而对数量性状的最大遗传多样性而言, 选择优先取样+多次聚类变异度取样方法则较适宜。用优先取样+多次聚类随机取样方法取样时, 采用最短距离法和重心法构建的核心种质最好, 用优先取样+多次聚类变异度取样时, 采用离差平方和法则是构建苧麻核心种质的最佳聚类方法。苧麻核心种质构建与质量性状的不同遗传距离无关, 但数量性状以欧氏距离最佳。

**关键词:** 苧麻; 核心种质; 取样方法; 聚类方法; 遗传距离

## Method of Establishing Ramie Core Collection

LUAN Ming-Bao, CHEN Jian-Hua\*, XU Ying, WANG Xiao-Fei, and SUN Zhi-Min

Institute of Bast Fiber Crops, Chinese Academy of Agricultural Sciences / Key laboratory of Stem-Fiber Biomass and Engineering Microbiology, Ministry of Agriculture, Changsha 410205, China

**Abstract:** Ramie (*Boehmeria nivea* L. Gaud) is an important bast fiber crop. Its fiber, one of the best bast fibers, is widely used in textile industry and plays an indispensable and irreplaceable role in the national economy. Core collection is a representative subset of entire germplasm collection, consisting of introduced accessions with minimum genetic redundancy and retained most initial collection. To develop the method of establishing ramie core collection, based on 25 agronomic traits data from 790 ramie germplasm of national ramie germplasm nursery, using software provided by professor Xu of Zhejiang university, we sampled a series of subsets at 20% proportion, respectively, by six hierarchical clustering methods (single linkage, complete linkage, median method, centroid method, ward's, unweighted pair-group method with arithmetic means), three sampling methods (stepwise clustering with random sampling strategy and preferred sampling strategy, stepwise clustering with deviation sampling strategy and preferred sampling strategy, random sampling strategy), three genetic distances in disordered qualitative traits (simple matching, jaccard, nei and li), and three genetic distances in ordered qualitative traits and quantitative characteristics (mahalanobis distance, euclidean distance, principal component). The genetic variation among subsets was compared by evaluating the I (index of genetic diversity), RPR (ratio of phenotype retained) of qualitative traits and MD (mean difference percentage), VD (variance difference percentage), CR (coincidence difference percentage), VR (variable rate of coefficient of variation) of quantitative traits. The results showed that different sampling methods have different impacts on qualitative traits and quantitative traits. For the largest genetic diversity of qualitative traits, it is effective to choose stepwise clustering with random sampling strategy and preferred sampling strategy. For the largest genetic diversity of quantitative traits, it is effective to choose stepwise clustering with deviation sampling strategy and preferred sampling strategy. The ramie core collection is the best one constructed by centroid method, single linkage clustering way under stepwise clustering with random sampling strategy and preferred sampling strategy and constructed by ward's clustering way under stepwise clustering with deviation sampling strategy and preferred sampling strategy. Core collection in ramie was not related with genetic distances of qualitative traits, whereas the core collection of ramie con-

本研究由农业部保种项目(NB09-2130135-41), 国家自然科学基金项目(30900913)和中国农业科学院科技经费项目资助。

\* 通讯作者(Corresponding author): 陈建华, E-mail: cjhbt@sina.com

第一作者联系方式: E-mail: luanmingbao2002@126.com, Tel: 0731-88998518

Received(收稿日期): 2010-04-03; Accepted(接受日期): 2010-08-05.

structed by euclidean distance of quantitative traits was the best.

**Keywords:** Ramie; Core collection; Sampling method; Clustering method; Genetic distance

植物遗传资源是新种质选育的基础,种质资源的收集、保存、评价和利用非常重要。然而庞大的种质资源不利于优良种质的保存和利用。由 Frankel 等<sup>[1]</sup>和 Brown<sup>[2]</sup>发展起来的核心种质学说以一定的方法选择整个种质资源的一部分,以最小的资源数量和遗传重复最大程度地代表整个种质资源的多样性,从而很好地解决了这一难题。目前,核心种质的构建不仅在小麦<sup>[3]</sup>、玉米<sup>[4]</sup>、水稻<sup>[5]</sup>等一年生作物上,而且在梅花<sup>[6]</sup>、咖啡<sup>[7]</sup>、桃<sup>[8]</sup>、茶树<sup>[9]</sup>、白桦<sup>[10]</sup>等多年生作物上也获得了成功。

苧麻(*Boehmeria nivea* L. Gaud),又叫“中国草”,起源于中国,为荨麻科苧麻属的多年生宿根性草本植物,是具有中国特色的天然纺织原料<sup>[11]</sup>。中国是世界上种植和使用苧麻历史最悠久的国家,也是苧麻品种变异类型和苧麻属野生种最多的国家。迄今为止,我国是收集、保存苧麻属种质资源最多的国家,仅国家种质长沙苧麻圃就拥有 2 000 多份材料,其中的栽培资源已广泛为育种和生产所利用<sup>[12]</sup>。但至今未见苧麻核心种质构建的报道,构建苧麻核心种质已成为一项刻不容缓的任务。一般采用系统聚类法对原种质资源进行群体分组,在此基础上抽样组成核心种质。不同的取样方法、聚类方法、遗传距离算法都会对核心种质的构建产生影响。由于遗传变异及系统发育等方面的差异,不同作物核心种质构建方法有所不同<sup>[3-10]</sup>。因此,方法的确立是构建苧麻核心种质的前提。本文分析了不同的取样方法、聚类方法、遗传距离算法对构建的苧麻核心种质的影响,旨在筛选出构建苧麻核心种质的最佳方法,同时为其他作物核心种质构建提供借鉴。

## 1 材料与方法

### 1.1 材料与数据

790 份苧麻种质资源种植于中国农业科学院麻类研究所国家种质长沙苧麻圃。采用本课题组多年调查的数据,包括无序质量性状(类别、根型、兜型、叶形、麻骨色、雌蕾色)、有序质量性状(叶面皱纹、叶色、生长整齐度、生长均匀度、绣脚、抗风性、耐旱性、抗花叶病、气候带、分株力)、数量性状(有效株率、工艺成熟天数、株高、茎粗、鲜皮厚度、鲜皮出麻率、原麻产量、单纤维支数、单纤维强力)。

### 1.2 农艺性状调查方法

参照《苧麻种质资源描述规范和数据标准》<sup>[13]</sup>。

### 1.3 核心种质分析方法

采用优先取样+多次聚类随机取样法、优先取样+多次聚类变异度取样方法和完全随机取样法进行取样。采用最短距离法、最长距离法、中间距离法、重心法、不加权类平均法和离差平方和法进行聚类。对无序质量性状采用 simple matching、Jaccard 和 Nei and li 三种遗传距离;对数量性状和有序质量性状采用欧氏遗传距离、马氏遗传距离和主成分遗传距离。构建核心种质参照徐海明等<sup>[14]</sup>的方法。对聚类方法、质量性状和数量性状遗传距离采用正交设计,即正交表之  $L_{18}(6^1 \times 3^3)$ ,空白列用于估计误差,利用 SAS 软件进行方差分析。将上面正交的 18 个组合作为区组,利用 SAS 软件进行取样方法的方差分析。

### 1.4 核心种质遗传参数评价

采用质量性状多样性指数均值<sup>[5]</sup>、质量性状表型比率均值<sup>[5]</sup>评价核心种质质量性状的遗传多样性,采用 Hu 等<sup>[15]</sup>提出的均值差异百分率、方差差异百分率、极差符合率和变异系数变化率等 4 个综合性评价指标评价核心种质数量性状的遗传多样性。

## 2 结果与分析

### 2.1 以不同取样方法构建的核心种质遗传参数的比较

在 20% 的取样水平下,把聚类方法、质量性状和数量性状遗传距离进行正交设计的 18 个组合作为区组,比较 3 种取样方法构建的苧麻核心种质的遗传参数(表 1)。表明质量性状多样性指数均值以优先取样+多次聚类随机取样方法和优先取样+多次聚类变异度取样方法最高,不经过聚类直接采用完全随机取样法最低;质量性状表型保留比率均值以优先取样+多次聚类随机取样方法最高,不经过聚类直接采用完全随机取样法最低。优先取样+多次聚类随机取样方法和优先取样+多次聚类变异度取样方法的数量性状均值差异百分率分别为 4.94% 和 9.26%,差异达到了显著水平,不经过聚类直接采用完全随机取样方法构建的核心种质与原群体间均值无显著差异,即均值差异百分率为 0。优先取样+多次聚类随机取样方法和优先取样+多次聚类变异度

取样方法的数量性状方差差异百分率分别为 28.4% 和 51.24%, 差异极显著, 且全部高于完全随机取样法。对于极差符合率, 不经过聚类直接采用完全随机取样法为 79.75%, 其余 2 种方法为 100%。对于数量性状变异系数变化率, 优先取样+多次聚类变异度取样方法最高, 优先取样+多次聚类随机取样方法次之, 不经过聚类直接采用完全随机取样法最低。在保留原群体遗传变异的前提下, 质量性状多样性指数均值、表型保留比率均值、数量性状方差差异百分率、变异系数变化率和极差符合率越高, 表明核心种质差异大、重复性状材料少, 构建效果

越好。以上分析表明, 优先取样+多次聚类随机取样方法能使质量性状品系间的差异达到最大化, 而优先取样+多次聚类变异度取样方法能使数量性状品系间的差异达到最大化。

2.2 以不同聚类方法构建的核心种质遗传参数的比较

在 20%的取样水平下, 对聚类方法以及质量性状和数量性状遗传距离采用正交设计, 比较两种取样方法下, 6 种聚类方法构建的苕麻核心种质的遗传参数(表 2)。结果表明, 优先取样+多次聚类随机取样方法下, 质量性状多样性指数均值以重心法和最

表 1 以不同取样方法分别构建的核心种质与原种质的遗传参数  
Table 1 Genetic parameters of total collection and core collection with different sampling strategy of quantitative traits

取样方法 Sampling strategy	种质 Collection	多样性指数均值 I	表型保留比率均值 RPR	均值差异百分率 MD	方差差异百分率 VD	极差符合率 CR	变异系数变化率 VR
优先取样+多次聚类随机取样法 A	Core_158	1.12 a	0.99 A	4.94 b	28.40 B	100.00 a	111.91 B
优先取样+多次聚类变异度取样法 B	Core_158	1.12 a	0.95 B	9.26 a	51.24 A	100.00 a	121.22 A
随机取样法 C	Core_158	1.09 b	0.69 C	0.00 c	22.22 C	79.75 b	97.59 C
原种质 Total		1.10	1.00	0	0	100.00	100.00

相同指标标以相同小(大)写字母的数据在 0.05(0.01)水平上差异不显著。

A: stepwise clustering with random sampling strategy and preferred sampling strategy; B: stepwise clustering with deviation sampling strategy and preferred sampling strategy; C: random sampling strategy. Values followed by same small (capital) letters in the same index are not significantly different at 5% (1%) probability levels. I: index of genetic diversity; RPR: ratio of phenotype retained; MD: mean difference percentage; VD: variance difference percentage; CR: coincidence difference percentage; VR: variable rate of coefficient of variation.

表 2 以不同聚类分别构建的核心种质与原种质的遗传参数  
Table 2 Genetic parameters of total collection and core collection using different clustering methods

聚类方法 Clustering method	种质 Collection	多样性指数均值 I	表型保留比率均值 RPR	均值差异百分率 VD	方差差异百分率 CR	极差符合率 CR	变异系数变化率 VR
优先取样+多次聚类随机取样法 A							
最短距离法 Single	Core_158	1.14 Aa	1.00 a	14.82 Aa	40.74 a	100 a	112.51 a
最长距离法 Complete	Core_158	1.12 BCcd	1.00 a	0.00 Bc	40.74 a	100 a	113.63 a
中间距离法 Median	Core_158	1.10 Ce	1.00 a	0.00 Bc	14.81 a	100 a	108.43 b
重心法 Centroid	Core_158	1.13 ABab	1.00 a	7.41 ABb	29.63 a	100 a	113.84 a
不加权类平均法 UPGMA	Core_158	1.11 Cde	1.00 a	0.00 Bc	14.81 a	100 a	111.68 ab
离差平方和法 Ward's	Core_158	1.12 BCbc	0.96 b	7.41 ABb	29.63 a	100 a	111.37 ab
优先取样+多次聚类变异度取样法 B							
最短距离法 Single	Core_158	1.13 ab	1.00 a	7.41 ab	40.71 ab	100 a	116.06 Bc
最长距离法 Complete	Core_158	1.12 ab	0.92 b	14.82 a	59.26 ab	100 a	127.11 Aa
中间距离法 Median	Core_158	1.11 b	0.96 ab	3.70 b	62.96 a	100 a	119.38 ABbc
重心法 Centroid	Core_158	1.12 ab	0.96 ab	11.11 ab	29.63 b	100 a	114.10 Bc
不加权类平均法 UPGMA	Core_158	1.13 a	0.92 b	11.11 ab	62.96 a	100 a	126.64 Aa
离差平方和法 Ward's	Core_158	1.13 ab	0.94 ab	7.41 ab	51.85 ab	100 a	124.03 ABab
原种质 Total		1.10	1.00	0	0	100	100.00

同一取样方法下, 相同指标标以相同小(大)写字母的数据在 0.05(0.01)水平上差异不显著。

A: stepwise clustering with random sampling strategy and preferred sampling strategy. B: stepwise clustering with deviation sampling strategy and preferred sampling strategy. Values followed by same small (capital) letters in the same index are not significantly different at 5% (1%) probability levels under the same sampling strategy. I: index of genetic diversity. RPR: ratio of phenotype retained; MD: mean difference percentage; VD: variance difference percentage; CR: coincidence difference percentage; VR: variable rate of coefficient of variation.

短距离法最高。质量性状表型保留比率均值除离差平方和法外,均为 1。对于数量性状均值差异百分率,最短距离法、重心法和离差平方和法分别为 14.82%、7.41%和 7.41%,其余 3 种方法构建的核心种质与原群体间无显著差异,即均值差异百分率为 0。数量性状方差差异百分率和极差符合率在 6 种聚类方法间差异不显著;数量性状变异系数变化率以中间距离法最低,其他 5 种聚类方法之间差异不显著。以上分析表明,优先取样+多次聚类随机取样方法下,采用最短距离法和重心法构建的核心种质在保持了原群体遗传变异的同时,使品系间的差异达到了最大化,是构建苕麻核心种质的最佳聚类方法。在优先取样+多次聚类变异度取样方法下,质量性状多样性指数均值和数量性状均值差异百分率以中间距离法最低,其余 5 种聚类方法差异不显著;质量性状表型保留比率均值在最短距离法、中间距离法、重心法和离差平方和法之间差异不显著,最短距离法高于最长距离法和不加权类平均法;数量性状方差差异百分率以重心法最低,其余 5 种聚类方法差异不显著;6 种聚类方法的极差符合率均为 100%;数量性状变异系数变化率以最长距离法和不加权类平均法显著高于最短距离法、中间距离法和重心法。以上分析表明,优先取样+多次聚类变异度取样方法下,采用离差平方和法构建的核心种质在保持了原群体遗传变异的同时,使品系间的差异达到了最大化,是构建苕麻核心种质的最佳聚类方法。

2.3 以不同遗传距离方法构建的核心种质遗传参数的比较

2.3.1 以质量性状不同遗传距离构建的核心种质遗传

参数的比较 在 20%的取样水平下,对聚类方法以及质量性状和数量性状遗传距离采用正交设计,比较了在两种取样方法下,无序质量性状以 Simple matching、Jaccard 和 Nei and li 三种遗传距离,有序质量性状以欧氏、马氏和主成分 3 种遗传距离,构建的苕麻核心种质的遗传参数(表 3 和表 4)。结果表明,质量性状利用不同的遗传距离算法,构建的苕麻核心种质除均值差异百分率有差异外,其他遗传参数没有差异。但均值差异百分率都低于 20%。

2.3.2 以数量性状不同遗传距离构建的核心种质遗传参数的比较 在 20%的取样水平下,聚类方法以及质量性状和数量性状遗传距离采用正交设计,比较在两种取样方法下,数量性状利用欧氏、马氏和主成分 3 种不同的遗传距离构建的苕麻核心种质的遗传参数(表 5)表明,优先取样+多次聚类随机取样方法下,欧氏遗传距离法构建的核心种质质量性状多样性指数均值、数量性状方差差异百分率和变异系数变化率均最高,马氏距离和主成分距离构建的次之。表型保留比率均值、均值差异百分率和极差符合率 3 种遗传距离构建的核心种质间差异不显著。优先取样+多次聚类变异度取样方法下,欧氏遗传距离法构建的核心种质的数量性状方差差异百分率和变异系数变化率均最高,马氏距离和主成分距离构建的次之。质量性状多样性指数均值、表型保留比率均值、均值差异百分率和极差符合率在 3 种遗传距离构建的核心种质间差异不显著。以上结果表明,数量性状采用欧氏距离,在保持了原群体遗传变异的同时,使品系间的差异达到了最大化,是构建苕麻核心种质的最佳遗传距离方法。

表 3 以无序质量性状不同遗传距离分别构建的核心种质与原种质的遗传参数  
Table 3 Genetic parameters of total collection and core collection using different genetic distances of disordered qualitative traits

遗传距离 Genetic distance	种质 Collection	多样性 指数均值 I	表型保留 比率均值 RPR	均值差异 百分率 MD	方差差异 百分率 VD	极差符 合率 CR	变异系数 变化率 VR
优先取样+多次聚类随机取样法 A							
Simple matching	Core_158	1.12 a	0.99 a	7.4 a	29.63 a	100 a	112.92 a
Jaccard	Core_158	1.12 a	1.00 a	3.7 a	24.07 a	100 a	111.93 a
Nei and li	Core_158	1.12 a	0.99 a	3.7 a	31.48 a	100 a	110.88 a
优先取样+多次聚类变异度取样法 B							
Simple matching	Core_158	1.12 a	0.94 a	12.96 a	55.56 a	100 a	121.92 a
Jaccard	Core_158	1.12 a	0.95 a	5.56 b	44.44 a	100 a	119.98 a
Nei and li	Core_158	1.13 a	0.96 a	9.26 ab	53.70 a	100 a	121.76 a
原种质 Total		1.10	1.00	0	0	100	100.00

Abbreviations are the same as given in Table 2.

表 4 以有序质量性状不同遗传距离分别构建的核心种质与原种质的遗传参数

Table 4 Genetic parameters of total collection and core collection using different genetic distances of ordered qualitative traits

遗传距离 Genetic distance	种质 Collection	多样性 指数均值 I	表型保留 比率均值 RPR	均值差异 百分率 MD	方差差异 百分率 VD	极差符合率 CR	变异系数 变化率 VR
优先取样+多次聚类随机取样法 A							
Simple matching	Core_158	1.12 a	0.99 a	7.41 a	29.63 a	100 a	112.33 a
Jaccard	Core_158	1.12 a	1.00 a	1.85 b	31.48 a	100 a	111.62 a
Nei and li	Core_158	1.12 a	0.99 a	5.56 ab	24.07 a	100 a	111.78 a
优先取样+多次聚类变异度取样法 B							
Simple matching	Core_158	1.12 a	0.95 a	11.11 a	55.56 a	100 a	122.45 a
Jaccard	Core_158	1.13 a	0.97 a	5.56 a	50.00 a	100 a	120.77 a
Nei and li	Core_158	1.12 a	0.93 a	11.11 a	48.15 a	100 a	120.44 a
原种质 Total		1.10	1.00	0	0	100	100.00

Abbreviations are the same as given in Table 2.

表 5 以数量性状不同遗传距离分别构建的核心种质与原种质的遗传参数

Table 5 Genetic parameters of total collection and core collection using different genetic distances of quantitative traits

遗传距离 Genetic distance	种质 Collection	多样性 指数均值 I	表型保留 比率均值 RPR	均值差异 百分率 MD	方差差异 百分率 VD	极差 符合率 CR	变异系数 变化率 VR
优先取样+多次聚类随机取样法 A							
欧氏 Euclidean	Core_158	1.13 Aa	1.00 a	3.70 a	48.15 Aa	100 a	115.11 Aa
马氏 Mahalanobis	Core_158	1.12 Bb	0.99 a	5.56 a	18.52 Bb	100 a	110.31 Bb
主成分 Principal component	Core_158	1.12 Bb	0.99 a	5.56 a	18.52 Bb	100 a	110.31 Bb
优先取样+多次聚类变异度取样法 B							
欧氏 Euclidean	Core_158	1.12 a	0.93 a	12.96 a	68.52 a	100 a	126.03 a
马氏 Mahalanobis	Core_158	1.13 a	0.96 a	7.41 a	42.59 b	100 a	118.81 b
主成分 Principal component	Core_158	1.13 a	0.96 a	7.41 a	42.59 b	100 a	118.81 b
原种质 Total		1.10	1.00	0	0	100	100.00

Abbreviations are the same as given in Table 2.

3 讨论

在核心种质构建的方法方面，呈连续型变异的数量性状和呈离散型变异的质量性状具有不同的变异模式，因此不能简单地整合。为充分利用数量性状和分子标记的遗传信息，Bar-Hen等<sup>[16]</sup>建议采用二阶段分析策略分析玉米自交系材料间的遗传关系，但该策略不是真正意义上的信息整合，因为每一分析阶段只用到一种类型的数据信息。Islam等<sup>[17]</sup>用非线性的主成分分析整合数量性状和质量性状，但这样的数据转换必定会损失部分遗传信息，而且也受划分类别多少的影响，存在一定的不合理性。徐海明等<sup>[14]</sup>提出了调整欧氏距离度量样品间遗传差异的方法，能有效整合质量性状和数量性状变异的信息。本研究采用该方法可充分利用数据，获得可靠的结果。

多样性指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<sup>[18]</sup>。本研究对质量性状采用了多样性指数均值和表型保留比例两个评价指标，两个指标越高，核心子集越能很好地代表原始群体。Hu等<sup>[15]</sup>提出了均值差异百分率、方差差异百分率、极差符合率和变异系数变化率 4 个遗传多样性评价指标，认为均值差异百分率小于 20%，极差符合率大于等于 80%时可以认为核心种质很好地代表了原群体；均值差异百分率越小，方差差异百分率、极差符合率和变异系数变化率越大，则核心子集越能代表原始群体的遗传多样性。李自超等<sup>[19]</sup>指出平均数或平均数离差不能作为一个核心种质的检验指标。最终核心种质库的性状均值可能会出现一定的变动，但这种变动不应该太大<sup>[18]</sup>。本研究对核心子集数量性状的评价，优先考虑方差差

异百分率、极差符合率和变异系数变化率。对于均值差异百分率,则按照 Hu 等<sup>[15]</sup>的标准,小于 20%。

采用 6 种聚类方法、3 种取样方法以及无序质量性状遗传距离、数量性状和有序质量性状遗传距离各 3 种进行核心种质构建,怎样确定最优组合是一个关键问题。笔者最初采用了随机选定其他因素,仅比较一种因素的策略,逐步确定最优因素,进而确定最优组合。这种确定方法在白桦<sup>[10]</sup>、棉花<sup>[18]</sup>等核心种质构建中得到应用。但是有理由怀疑一开始规定的组合导致最终的结果产生,无法证明不同的组合是否也会得出同样的结果。后来采用参照正交实验的设计方法,设计了 6 种聚类方法、3 种取样方法以及无序质量性状遗传距离、数量性状和有序质量性状遗传距离各 3 种的正交实验,分析结果表明仅有取样方法差异达到了显著水平,其余 4 种因素差异均未达到显著水平,对比第一种方法,结果差异很大,并且可以确定这个正交实验结果存在很大问题。分析原因,发现取样方法中完全随机取样法获得的核心种质结果与其余 4 种因素无关,也就是说,不管采用何种聚类方法、遗传距离,只要采用完全随机取样法得到的核心种质就是一样的。正交设计中完全随机取样法对其他因素效应的屏蔽,影响了对其他因素的分析,导致获得错误的结果。最后笔者把取样方法单独列出,6 种聚类方法以及无序质量性状遗传距离、数量性状和有序质量性状遗传距离各 3 种进行正交设计,分别分析了两种取样方法下聚类方法、质量性状和数量性状的遗传距离对构建核心种质的影响,并比较了不同取样方法对构建核心种质的影响。结果与作者最初采用的第一种方法的结果有差别,但差别不大。差别主要表现在聚类方法的确定。第一种方法认为重心法是最好的聚类方法,而本方法认为在优先取样+多次聚类变异度取样方法下,采用离差平方和法最好;优先取样+多次聚类随机取样方法,利用最短距离法或重心法进行聚类最好。也表明聚类方法和取样方法存在互作效应。

核心种质构建的本质是取样方法的选择。在桃的种质构建中,聚类取样的方法优于完全随机取样<sup>[8]</sup>。白桦的核心种质构建,多次聚类优先取样法比较好,其次为随机取样法,最后为偏离度取样法<sup>[10]</sup>。以总体聚类分组、按对数比例在组内聚类取样为构建甘蔗杂交品种初级核心种质的最佳策略组合<sup>[20]</sup>。本研

究结果表明,优先取样+多次聚类随机取样方法能使质量性状品系间的差异达到最大化,而优先取样+多次聚类变异度取样方法能使数量性状品系间的差异达到最大化,表明不同的取样方法对质量性状和数量性状的遗传多样性影响不同。在苧麻核心种质构建中要根据构建目的进行取样方法的取舍。如果数量性状比较重要,就选择优先取样+多次聚类变异度取样方法。相反,则考虑优先取样+多次聚类随机取样方法。

在核心种质构建过程中,构建策略是关键环节。然而不同的作物由于遗传构成和遗传分化不同,构建策略不应完全相同,应筛选出适合每种作物自身特点的方法策略进行核心种质的构建。以聚类方法构建核心种质比较有效。而不同的聚类方法由于产生不同的分组,会影响构建核心种质的代表性。离差平方和聚类法在构建芝麻<sup>[21]</sup>核心种质时效果最好,类平均法则在水稻<sup>[22]</sup>构建核心种质时效果最优。多年生作物中,最短距离法对白桦<sup>[23]</sup>核心种质构建效果最好。优化最小距离法在杏<sup>[24]</sup>中认为最好。本研究表明,不同的取样方法需要采用不同的聚类方法以构建核心种质。在优先取样+多次聚类随机取样方法下,采用最短距离法和重心法构建的核心种质在保持了原群体遗传变异的同时,使品系间的差异达到了最大化,在质量性状多样性指数均值和变异系数变化率等方面优于其他方法,是构建苧麻核心种质的最佳聚类方法。在优先取样+多次聚类变异度取样方法下,采用离差平方和法则是构建苧麻核心种质的最佳聚类方法,使其构建的核心种质在 6 个评价指标中均表现优异。

遗传距离是度量类群间遗传相似性的综合数量指标,不同遗传距离的计算方法直接影响着群体的分类和核心种质的构建。徐海明<sup>[18]</sup>在棉花上的研究表明,采用马氏距离优于欧氏距离。与以上结果不同,本研究结果表明,质量性状利用不同的遗传距离算法,构建的苧麻核心种质除均值差异百分率有差异外,其他遗传参数没有差异。但均值差异百分率都低于 20%。表明苧麻核心种质构建与质量性状的不同遗传距离无关。而数量性状采用欧氏距离构建的核心种质,数量性状方差差异百分率和变异系数变化率均最高,在保持了原群体遗传变异的同时,使品系间的差异达到了最大化,是构建苧麻核心种质的最佳遗传距离方法。

## 4 结论

为保持质量性状的最大遗传多样性, 应选择优先取样+多次聚类随机取样方法, 利用最短距离法或重心法进行聚类, 任意选择一种质量性状的遗传距离, 而对数量性状采用欧氏距离。相反, 为保持数量性状的最大遗传多样性, 应选择优先取样+多次聚类变异度取样方法, 利用离差平方和法的聚类方法, 任意选择一种质量性状的遗传距离, 而对数量性状采用欧氏距离。

致谢: 核心种质构建得到浙江大学徐海明老师的帮助; 正交分析得到西北农林科技大学李得孝老师、华东师范大学张应山老师的帮助, 在此表示感谢。

## References

- [1] Frankel O H, Brown A H D. Current Plant Genetic Resources—a Critical Appraisal. In: Genetics: New Frontiers. Vol. IV. New Delhi, India: Oxford and IBH Publishing, 1984. pp 1–11
- [2] Brown A H D. Core collection: a practical approach to genetic resources management. *Genome*, 1989, 31: 818–824
- [3] Hao C Y, Zhang X Y, Wang L F, Dong Y S, Shang X W, Jia J Z. Genetic diversity and core collection evaluations in common wheat germplasm from the northwestern spring wheat region in china. *Mol Breed*, 2006, 17: 69–77
- [4] Yao Q-L(姚启伦), Fang P(方平), Yang K-C(杨克诚), Pan G-T(潘光堂). Methods of constructing a core collection of maize landraces in southwest China based on SSR data. *J Hunan Agric Univ (Nat Sci)*(湖南农业大学学报·自然科学版), 2009, 35(3): 225–228 (in Chinese with English abstract)
- [5] 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), Wang X-K(王象坤). Comparison of parameters for testing the rice core collection in phenotype. *Acta Agron Sin (作物学报)*, 2003, 29(2): 252–257 (in Chinese with English abstract)
- [6] Ming J(明军), Zhang Q-X(张启翔), Lan Y-P(兰彦平). Core collection of *Prunus mume* Sieb. et Zucc. *J Beijing For Univ (北京林业大学学报)*, 2005, 27(2): 65–69 (in Chinese with English abstract)
- [7] Prakash N S, Combes M C, Dussert S, Naveen S, Lashermes P. Analysis of genetic diversity in Indian robusta coffee gene pool (*Coffea canephora*) in comparison with a representative core collection using SSRs and AFLPs. *Genet Resour Crop Evol*, 2005, 52: 333–343
- [8] Li Y-X(李银霞), An L-J(安丽君), Jiang Q(姜全). Establishment and evaluation of the core collection of peach (*Prunus persica* L. Bat sch.) cultivars. *J China Agric Univ (中国农业大学学报)*, 2007, 12(5): 22–28 (in Chinese with English abstract)
- [9] Wang X-C(王新超), Liu Z(刘振), Yao M-Z(姚明哲). Sampling strategy to establish a primary core collection of Chinese tea germplasms. *J Tea Sci (茶叶科学)*, 2009, 29(2): 159–167 (in Chinese with English abstract)
- [10] Wei Z-G(魏志刚), Gao Y-C(高玉池), Yang C-P(杨传平). Sampling method for constructing germplasm core collections of *Betula platyphylla*. *J Northeast For Univ (东北林业大学学报)*, 2009, 37(7): 1–4 (in Chinese with English abstract)
- [11] Lu H-R(卢浩然). Bast-Fiber Crops Cultivation Science in China (中国麻类作物栽培学). Beijing: Agricultural Press, 1992. pp 1–10 (in Chinese)
- [12] Xiong H-P(熊和平). Bast-fiber Crops Breeding (麻类作物育种学). Beijing: China Agricultural Science and Technology Press, 2008. pp 46–50 (in Chinese)
- [13] Jie Y-C(揭雨成). Descriptors and Data Standard for Ramie (苕麻种质资源描述规范和数据标准). Beijing: China Agriculture Press, 2007. pp 9–28 (in Chinese)
- [14] Xu H-M(徐海明), Li X-L(李晓玲), Li J-Q(李金泉), Liu G-F(刘桂富), Hu J(胡晋). Study on the strategies of integrating qualitative and quantitative traits for constructing core collection. *J Zhejiang Univ (Agric & Life Sci)* (浙江大学学报·农业与生命科学版), 2005, 31(4): 362–367 (in Chinese with English abstract)
- [15] Hu J, Zhu J, Xu H M. Methods of constructing core collection by stepwise cluster with three sampling strategies based on genotypic values of crops. *Theor Appl Genet*, 2000, 101: 264–268
- [16] Bar-Hen A, Charcosset A, Bourgoin M, Guiard J. Relationship between genetic markers and morphological traits in maize inbred lines collection. *Euphytica*, 1995, 84: 145–154
- [17] Islam F M A, Basford K E, Redden R J, Gonzalez A V, Kroonenberg P M, Beebe S. Genetic variability in cultivated common bean beyond the two major genes. *Genet Resour Crop Evol*, 2002, 49: 271–283
- [18] Xu H-M(徐海明). Study on Methods of Constructing Core Collection of Germplasm and Their Applications in Core Construction (种质资源核心库的构建方法的研究及其应用). Zhejiang: Zhejiang University, 2005. pp 26–71 (in Chinese)
- [19] Li Z-C(李自超), Zhang H-L(张洪亮), Zeng Y-W(曾亚文), Yang Z-Y(杨忠义). Study on sampling schemes of core collection of

- local varieties of rice in Yunnan, China. *Sci Agric Sin* (中国农业科学), 33(5): 1–7 (in Chinese with English abstract)
- [20] Liu X-L(刘新龙), Cai Q(蔡青), Ma L(马丽), Wu C-W(吴才文). Strategy of sampling for pre-core collection of sugarcane hybrid. *Acta Agron Sin* (作物学报), 2009, 35(7): 1209–1216 (in Chinese with English abstract)
- [21] Zhang X, Zhao Y, Cheng Y. Establishment of sesame germplasm core collection in China. *Genet Resour Crop Evol*, 2000, 47: 273–279
- [22] Li C T, Shi C H, Wu J Q. Methods of developing core collections based on the predicted genotypic value of rice (*Oryza sativa* L.). *Theor Appl Gene*, 2004, 108: 1172–1176
- [23] Wei Z-G(魏志刚), Gao Y-C(高玉池), Liu G-J(刘关君). Study on the clustering method of constructing core collection in *Betula platyphylla* germplasm. *J Plant Genet Resour* (植物遗传资源学报), 2009, 10(3): 405–410 (in Chinese with English abstract)
- [24] Zhang Q-P(章秋平), Liu W-S(刘威生) Liu N(刘宁). Establishment and evaluation of primary core collection of apricot (*Armeniaca vulgaris*) germplasm. *J Fruit Sci* (果树学报), 2009, 26(6): 819–825 (in Chinese with English abstract)