

# 利用表型数据构建陆地棉核心种质

代攀虹<sup>1,2</sup>, 孙君灵<sup>2</sup>, 贾银华<sup>2</sup>, 杜雄明<sup>2</sup>, 王 溢<sup>1</sup>

(<sup>1</sup>长江大学农学院, 湖北荆州 434025; <sup>2</sup>中国农业科学院棉花研究所/棉花生物国家重点实验室, 河南安阳 455000)

**摘要:**以 5963 份陆地棉种质资源为材料, 根据品种主要突变性状和品种类型分组成 11 组群, 在分组的基础上利用 21 个表型性状, 用非加权类平均聚类分析法, 构建了 281 份陆地棉核心种质, 占全部种质资源总量的 4.71%。利用不同性状的均值  $t$  测验、方差  $F$  测验、变异系数、多样性指数  $t$  检验、均值、极差、表型方差、变异系数、均值差异百分率、方差差异百分率、极差符合率、变异系数变化率、主成分分析等参数进行核心种质代表性检验和评价。结果表明, 所构建的陆地棉核心种质可以代表全部种质的遗传多样性。

**关键词:** 陆地棉; 核心种质; 聚类分析; 遗传多样性

## Construction of Core Collection of Upland Cotton Based on Phenotypic Data

DAI Pan-hong<sup>1,2</sup>, SUN Jun-ling<sup>2</sup>, JIA Yin-hua<sup>2</sup>, DU Xiong-ming<sup>2</sup>, WANG Mi<sup>1</sup>

(<sup>1</sup>Agricultural College, Yangtze University, Jingzhou Hubei 434025; <sup>2</sup>State Key Laboratory of Cotton Biology/Cotton Research Institute, Chinese Academy of Agricultural Sciences, Anyang Henan 455000)

**Abstract:** A total of 5963 germplasm collection of upland cotton accessions were sampled for constructing a core collection. The whole collection was firstly divided into 11 groups based on the main traits and genotypes. The resultant core collection including 281 accessions was constructed with 21 traits in each subset by unweighted pair-group method with arithmetic means (UPGMA), which accounted for 4.71% of total collection. Furthermore the representativeness between core collection and total collections was assessed by the  $t$  test,  $F$  test of variance, coefficient of variation, the diversity index  $t$  test, average, range, phenotypic variance and variation coefficient, mean difference percentage, variance difference percentage, coincidence rate of range, changeable rate of coefficient of variation and principal component analysis. The results demonstrated that the selected core collection could represent the genetic diversity of initial collection.

**Key words:** upland cotton; core collection; clustering analysis; genetic diversity

棉花是世界上重要的天然纺织纤维作物。陆地棉 (*Gossypium hirsutum* L.) 起源于中美洲墨西哥南部的高原及加勒比地区, 1865 年引进中国, 是世界上种植面积最大的棉种, 也是各个棉产国的主栽类型, 每年陆地棉产量占世界棉花总产量的 95% 以上<sup>[1-2]</sup>。因此, 如何高效地利用陆地棉种质资源材

料为现代育种服务, 提高育种效率和稳步增产显得尤为重要。目前, 中国保存陆地棉资源 7712 份, 其中栽培种有 7362 份<sup>[3]</sup>。如此丰富的遗传资源为陆地棉育种工作提供了大量的材料, 但在育种上能够有效开发利用的却很有限, 同时庞大的资源数量给种质资源的保存、研究、评价和利用带来诸多的困

收稿日期: 2016-01-09 修回日期: 2016-02-24 网络出版日期: 2016-10-12

URL: <http://www.cnki.net/kcms/detail/11.4996.S.20161012.0916.006.html>

基金项目: 农业部保种项目 (2015N WB039); 科技部、财政部国家科技基础条件平台项目 (NICGR2015-012); “十二五”国家支撑计划 (2013BAD01B 03)

第一作者主要从事棉花种质资源研究。E-mail: daipanhong@126.com

通信作者: 王溢, 主要从事棉花栽培及遗传育种研究工作。E-mail: edwmi@126.com

杜雄明, 主要从事棉花种质资源研究工作。E-mail: dxm630723@163.com

难<sup>[4]</sup>。构建核心种质是解决这一矛盾的有效方法,这样既保证了种质材料的遗传多样性,又减少了种质资源的鉴定和分析的数量。核心种质的概念由 O. H. Frankel<sup>[5]</sup> 提出,1989 年 A. H. D. Brown<sup>[6]</sup> 做了进一步完善,其目的就是以最少的资源份数和遗传重复性,最大限度地代表整个种质资源的遗传多样性,李自超等<sup>[7]</sup> 认为核心种质还应包括具有实际应用价值的基因。迄今国内外已完成的核心种质的构建主要基于两种数据类型,一类是依靠农艺、形态、抗性性状数据,如大麦<sup>[8]</sup>、中国小豆<sup>[9]</sup>、野苹果<sup>[10]</sup>、新疆甜瓜<sup>[11]</sup>、灰楸<sup>[12]</sup>、中国芒<sup>[13]</sup>、葡萄<sup>[14]</sup>等作物,另一类是蛋白质、同工酶及分子标记的数据,如水稻<sup>[15]</sup>、大豆<sup>[16]</sup>、桃<sup>[17]</sup>、小麦<sup>[18]</sup>、大麦<sup>[19]</sup>、玉米<sup>[20-21]</sup>、核桃<sup>[22]</sup>等。两种类型的数据都可以用作棉花核心种质的构建,目前关于陆地棉核心种质的构建尚未见报道,为此,尽快完成陆地棉核心种质的构建工作迫在眉睫。O. H. Frankel 等<sup>[23]</sup>、A. H. D. Brown 等<sup>[4]</sup> 提出可以利用地理来源和农艺性状构建核心种质。本研究是根据陆地棉的品种特性、农艺性状等数据资料,初步构建我国陆地棉的核心材料,并采用主成分分析、多样性指数  $t$  检验、初选核心样品和总收集样品性状的平均值、极差、方差、变异系数、均值差异百分率、方差差异百分率、极差符合率、变异系数变化率等对构建的核心种质进行确认并验证其代表性<sup>[10,24-27]</sup>,从而为陆地棉种质资源的有效利用提供参考。

## 1 材料与方法

### 1.1 材料

本研究的 5963 份陆地棉资源数据资料来源于中国农业科学院棉花研究所国家棉花种质资源中期库。其中数据主要包括比强度、长度整齐度、果枝节位、果枝伸展角、花蕊色、花药色、跨长、铃形、铃重、麦克隆值、耐盐性、生育期、叶厚度、叶片大小、叶绒毛、叶色、叶形、衣分、株高、株型、子指等 21 项农艺性状。

### 1.2 方法

**1.2.1 数据标准化** 对陆地棉株型、叶形、花蕊色等 11 个非数值型性状予以赋值(表 1)。

**1.2.2 数据分组** 首先,把整体材料分成两种类型,具有突变性状的遗传材料和普通材料,遗传材料随即被划分成无酚棉、绿色棉、黄色叶、红叶棉、鸡脚棉、光子棉、零式果枝、其他遗传材料等 8 个组群,其余的普通材料根据品种类型划分为地方品种、品系、选育品种 3 个组别。然后,以各组群为单位进行遗传多样性、聚类分析及其核心种质取样。

表 1 陆地棉非数值型性状赋值

Table 1 Coden designed for quantitative traits in upland cotton

性状 Trait	赋值 Coden of quantitative traits
株型 PT	筒型 = 1, 塔型 = 2, 倒塔型 = 3
叶形 LS	阔叶 = 1, 鸡脚 = 3, 超鸡脚 = 4, 皱缩 = 5, 卵圆 = 8
叶片大小 LSI	小 = 1, 中 = 2, 大 = 3
叶厚度 LT	薄 = 1, 中 = 2, 厚 = 3
叶色 LC	浅绿 = 1, 绿 = 2, 深绿 = 3, 黄 = 4, 紫红 = 7, 斑驳 = 8
叶绒毛 LPA	无 = 1, 少 = 2, 中 = 3, 较多 = 4, 多 = 5
花蕊色 SC	无 = 1, 乳白 = 2, 黄 = 3, 粉红 = 4, 红 = 5, 紫 = 6
花药色 AC	乳白 = 1, 黄 = 2, 黄红 = 3, 红 = 5
铃形 BS	圆形 = 1, 卵圆 = 2, 长卵圆 = 3, 圆锥 = 4
果枝伸展角 FSA	大于 45 度 = 1, 等于 45 度 = 2, 小于 45 度 = 3
耐盐性 STT	高抗 = 1, 抗 = 2, 耐 = 3, 不耐 = 4

PT: Plant type; LS: Leaf shape; LSI: Leaf size; LT: Leaf thickness; LC: Leaf color; LPA: Leaf pubescence amount; SC: Stamen color; AC: Anther color; BS: Boll shape; FSA: Fruit spur angle; STT: Salinity to tolerance; the same as below

**1.2.3 多样性指数** 分析多样性指数采用 Shannon-weaver 信息多样性指数( $I$ , index of genetic diversity)

用  $H'$  来表示,其计算公式:如果用  $H' = - \sum_{j=1}^i P_j \ln P_j$  表示任一组中某一性状的多样性程度,其中  $P_j$  表示第  $j$  组中变异类型出现的频率,就可用所有组中多样性指数的平均值  $\overline{H'} = ( \sum_{i=1}^n \sum_{j=0}^l P_{ij} \ln P_{ij} ) / n$ ,表示全部种质该性状的多样性程度,其中,  $n$  是组数,  $l$  是第  $i$  组中性状代码数,  $P_{ij}$  是第  $i$  组中第  $j$  个性状代码数出现的频率<sup>[25,28]</sup>。用各性状相对应的多样性指数的平均值表示某一组或所有种质资源的遗传多样性程度。

**1.2.4 聚类方法** 参照文献[29-30]提出的分组聚类的方法并根据具体情况对方法进行改进,具体方法为:以组为单位,各组内利用上述 21 个表型性状数据进行聚类,根据组内多样性指数大小,每一个小组内根据聚类分析结果随机抽取 1~2 份作为初选核心材料,各组所抽取的全部材料构成备选核心种质。最终确定所用聚类方法为非加权类平均法(UPGMA, unweighted pair-group method with arithmetic means);品种间采用的遗传距离为欧氏距离。

**1.2.5 取样数目** 根据上述聚类分析结果,并结合遗传多样性指数适当调整目标种质数目,遗传多样

性指数高的组适当地增加取样数,而较低的组则减少取样比例<sup>[6,25,31]</sup>;依据取样比例确定各组的取样数<sup>[25]</sup>。统计各组内的所选核心种质数目,并计算该组核心种质取样数目占全部种质资源的比例(5%~10%)。

**1.2.6 核心种质代表性检验** 为了检验核心种质的代表性,本研究选用多样性指数 *t* 检验、均值(Mean)、极差(Range)、表型方差(VPV, variation of phenotypic value)、变异系数(CV, coefficient of variation,)、均值差异百分率(MD, mean difference percentage、方差差异百分率(VD, variance difference percentage)、极差符合率(CR, coincidence rate of range)、变异系数变化率(VR, changeable rate of coefficient of variation)等9个作为核心种质评价指标<sup>[10,24-30]</sup>。通过比较核心种质与全部种质资源的相关参数,以此判断所构建的核心种质是否具有代表性。

根据前人研究结果,以上参数中按照均值差异百分率(MD)小于20%,极差差异百分率(CR)大于80%的原则,认为核心种质能够代表原种质群体的遗传多样性,同时MD越小,VD和VR越大,核心材料越能够代表原种质资源的遗传多样性<sup>[30]</sup>。

**1.3 数据分析** 非加权类平均法(UPGMA)聚类采用MVSP3.2软件进行分析,用Excel计算遗传多样性指数、变异系数等,其余参数均利用SPSS19.0软件进行计算。

2 结果与分析

2.1 各组陆地棉种质资源遗传多样性指数、取样数目及性状变异分析

以品种主要突变性状为依据分组,共计11个组,各组21个性状的多样性指数见表2,其中图1以绿色棉为例说明分组聚类及取样数目的确定,原种质资源与核心种质各组多样性指数比较以及取样比例见表3。由表2可知,11个小组的21个表型性状的多样性指数平均值介于1.16~1.39之间,整体分析红叶棉、品系两个组的遗传多样性指数略高于其他组,这与陆地棉的生长习性紧密相关,同时也说明这两组种质资源较为丰富,遗传多样性较高。图1中以绿色棉聚类结果为例,可以将绿色棉分为两大类,每一类中随机抽取1~2份作为核心材料,最终入选2份作为核心种质,其他各组均按照此方法筛选出备选核心种质。表3中原种质群体各组之间多样性指数的波动不大,而核心种质遗传多样性指数介于0.318~1.325之间,变幅高于原种质群体,原因可能是原种质资源组内种质数目过少,造成取样后遗传多样性指数降低,根据上述核心种质的取样方法,从5963份总种质资源中筛选出了281份材料作为陆地棉的核心种质,中选率为4.71%,对于多样性指数略高的组可适当地减少取样比例,而个别组遗传资源较少的组则要适当地增加取样数目。

表2 陆地棉全部种质资源各组遗传多样性指数

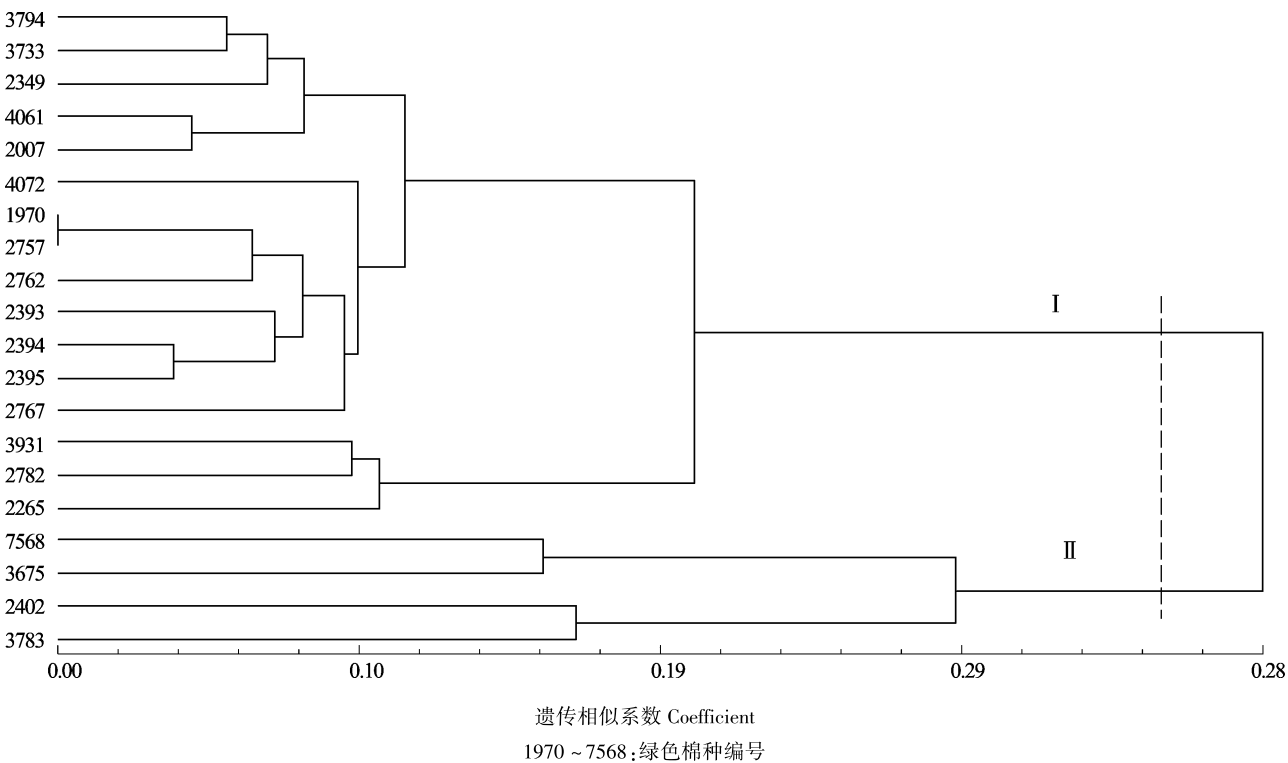
Table 2 Index of genetic diversity in entire germplasm of different groups

性状 Trait	各组多样性指数 Index of genetic diversity of different groups										
	无酚棉	绿色棉	黄色叶	红叶棉	鸡脚棉	光子棉	零式果枝	地方品种	品系	选育品种	遗传材料
	LGC	GC	YLC	RLC	OC	AC	ZTB	L	S	MV	GM
比强度 FS	2.17	2.06	2.01	2.06	2.05	2.02	2.14	2.15	2.18	2.24	2.20
长度整齐度 LU	2.26	2.22	2.24	2.52	2.18	2.00	2.51	2.13	2.33	2.38	2.47
果枝节位 SBN	1.75	1.50	1.24	1.61	1.72	1.81	1.52	1.73	1.76	1.67	1.53
果枝伸展角 FSA	0.26	0.00	0.18	0.63	0.09	0.39	0.97	0.52	0.35	0.40	0.46
花蕊色 SC	1.00	1.18	0.68	1.02	1.07	1.21	0.81	0.66	1.08	1.02	0.96
花药色 AC	0.28	0.69	0.18	0.72	0.31	0.64	0.17	0.00	0.24	0.20	0.29
跨长 SP	2.01	2.28	2.06	2.31	2.17	2.30	2.19	2.26	2.09	2.10	2.25
铃形 BS	0.71	0.69	0.25	0.73	0.68	0.64	0.47	0.70	0.65	0.67	0.53
铃重 BW	1.27	0.94	1.28	1.36	1.40	1.21	1.31	1.33	1.30	1.23	1.27
马克隆值 M	1.02	0.71	1.02	1.24	0.98	1.18	1.20	1.04	1.07	1.01	1.01
耐盐性 SIT	0.66	0.20	0.30	0.35	0.67	0.79	0.54	0.66	0.53	0.36	0.36
生育期 GW	3.27	2.69	3.10	3.02	3.40	2.97	3.33	3.05	3.70	3.67	3.45
叶厚度 LT	0.60	0.61	0.69	0.88	0.70	0.68	0.88	0.72	0.68	0.73	0.69

表 2(续)

性状 Trait	各组多样性指数 Index of genetic diversity of different groups										
	无酚棉	绿色棉	黄色叶	红叶棉	鸡脚棉	光子棉	零式果枝	地方品种	品系	选育品种	遗传材料
	LGC	GC	YLC	RLC	OC	AC	ZTB	L	S	MV	GM
叶片大小 LSI	0.80	0.42	0.82	0.72	1.04	0.87	0.84	0.94	0.80	0.76	0.79
叶绒毛 LPA	1.24	1.05	0.96	1.05	1.04	1.06	0.87	1.07	1.09	1.08	1.12
叶色 LC	0.09	0.56	0.79	0.21	0.15	0.15	0.07	0.68	0.13	0.11	0.08
叶形 LS	0.03	0.20	0.43	0.39	0.36	0.00	0.07	0.60	0.00	0.00	0.00
叶绒毛 LP	2.42	2.04	2.38	2.63	2.60	2.72	2.48	2.26	2.70	2.69	2.68
株高 PH	4.08	2.83	3.26	3.44	3.83	3.06	3.64	3.28	4.35	4.21	3.86
株型 PT	0.40	0.00	0.29	0.47	0.24	0.24	0.17	0.70	0.34	0.20	0.49
子指 SI	1.74	1.58	1.53	1.67	1.75	1.81	1.83	1.82	1.75	1.65	1.71
平均 Average	1.34	1.16	1.22	1.38	1.35	1.32	1.33	1.35	1.39	1.35	1.34

FS: Fiber strength, LU: Length uniformity, SBN: Sympodial brand node, FSA: Fruit spur angle, SL: Span length, BS: Boll shape, BW: Boll weight, M: Micronaire, GP: Growth period, LP: Lint percentage, PH: Plant height, SI: Seed index, LGC: Lack gossypol cottons, GC: Green cotton, YLC: Yellow leaf cotton, RLC: Red leaf cotton, OC: Okra cotton, AC: Absent cotton, ZTB: Zero type brand, L: Landraces, S: Strain, MV: Modern Varieties, GM: Genetic materials, the same as below



1970 ~ 7568: The code of green cotton cultivars

图 1 20 份绿色棉种质的 UPGMA 聚类图

Fig. 1 UPGMA dendrogram of 20 green cotton cultivars

2.2 陆地棉核心种质代表性检验

2.2.1 多样性指数及其 *t* 检验 原种质群体中表型性状的遗传多样性指数变幅为 0.121 ~ 4.353, 核心种质多样性指数的变幅在 0.127 ~ 4.202 之间,由此可见核心种质与原种质群体之间整体的

变异幅度一致,原种质资源群体与核心种质的表型性状中除了果枝伸展角和叶色两个性状的多样性指数 *t* 检验差异极显著外,其他均未达到显著水平,说明所选核心种质具有较好的代表性(表 4)。

表 3 各组材料的数目、中选比例和遗传多样性指数

Table 3 Accessions, sampling percentage and genetic diversity of different groups

组号 Group No.	组名 Groups	份数 Accessions			多样性指数 Genetic diversity ( $H'$ )	
		全部种质 Entire collection	核心种质 Core collection	取样比例(%) Sampling percentage	全部种质 Entire collection	核心种质 Core collection
1	无酚棉 LGC	209	11	5.26	1.271	1.009
2	绿色棉 GC	20	2	10.00	1.139	0.404
3	黄色叶 YLC	44	2	4.55	1.138	0.318
4	红叶棉 RLC	45	3	6.67	1.326	0.643
5	鸡脚棉 OC	107	4	3.74	1.275	0.675
6	光子棉 AC	30	3	10.00	1.244	0.624
7	零式果枝 ZTB	72	4	5.56	1.249	0.571
8	地方品种 L	312	21	6.73	1.280	1.178
9	品系 S	3814	139	3.64	1.316	1.325
10	选育品种 MV	1132	83	7.33	1.269	1.239
11	遗传材料 GM	178	11	6.18	1.271	1.108
共计	Total	5963	281	4.71	1.253	0.827

表 4 全部种质与核心种质 21 个性状遗传多样性指数比较

Table 4 Comparison of diversity indices for 21 traits between the entire and core collection of upland cotton

性状 Trait	全部种质 多样性指数 Entire collection index of genetic diversity	核心种质 多样性指数 Core collection index of genetic diversity	$t$ 值 $t$ value
比强度 FS	2.218	2.259	-0.312
长度整齐度 LU	2.372	2.430	-0.409
果枝节位 SBN	1.743	1.907	-1.515
果枝伸展角 FSA	0.385	0.315	4.684 **
花蕊色 SC	1.072	1.147	-1.278
花药色 AC	0.259	0.252	0.612
跨长 SL	2.169	2.294	-0.935
铃形 BS	0.651	0.674	-0.679
铃重 BW	1.315	1.411	-1.237
麦克隆值 M	1.072	1.152	-1.320
耐盐性 SA	0.777	0.818	-0.877
生育期 GP	3.719	3.605	0.523
叶厚度 LT	0.701	0.735	-0.928
叶片大小 LSI	0.811	0.881	-1.600
叶绒毛 LPA	1.099	1.125	-0.453
叶色 LC	0.197	0.326	-7.015 **
叶形 LS	0.121	0.127	-0.466
衣分 LP	2.756	2.896	-0.814
株高 PH	4.353	4.202	0.589
株型 PT	0.352	0.333	1.220
子指 SI	1.746	1.773	-0.274

$t_{0.05}=1.960$ ;  $t_{0.01}=2.576$ ; \*, \*\*: 分别表示 0.05、0.01 水平有显著差异,下同

$t_{0.05}=1.960$ ,  $t_{0.01}=2.576$ , \* and \*\* represent significant difference at 0.05 and 0.01 level, respectively, the same as below

2.2.2 总种质资源与核心种质性状特征值  $t$  测验、 $F$  测验以及差异检验分析 由表 5 可知,原种质资源群体中表型性状均发生较大的变异,不同性状间的变异程度不同,其中除长度整齐度、跨长和生育期变异程度低于 10% 外,其余表型性状的变异程度均较大,其中果枝伸展角、花蕊色和叶形的变异系数超过 40%,说明该种质资源群体的遗传变异丰富,遗传多样性的分布范围较广,适合进行核心种质的构建。

从 21 个表型性状的评价结果可以看出,核心种质性状的变异系数 ( $CV$ ) 均高于原群体,表明核心种质具有良好的异质性;经  $t$  测验,除了果枝节位、花蕊色和铃形 3 个性状差异有统计学意义外,其他性状均差异不显著;经  $F$  测验,有 18 个性状的方差为齐性,而果枝节位、花蕊色、叶色 3 个性状的方差显著高于总遗传资源,表明核心种质获得更大的变异(表 5)。由表 5 可推算出,核心种质与全部种质的均值差异百分率( $MD$ )和方差差异百分率( $VD$ )均为 85.7%,极差符合率( $CR$ )为 95.8%,变异系数变化率( $VR$ )为 114.1%。按照  $MD < 20\%$ 、 $CR > 80\%$  的原则,同时  $VR$  和  $CR$  越大就越能够代表原种质群体的遗传多样性,由此可知,所选核心种质基本能够代表全部种质资源的遗传多样性。

2.2.3 主成分分析 对核心种质和原种质资源的主成分分析可知(表 6),两者的特征值、贡献率及累计贡献率极为相近。以特征值大于 1 为标准,各选



表 5 陆地棉全部种质与核心种质平均值、极差、方差、变异系数的比较

Table 5 Comparison of mean, range, variance and variation coefficient in entire upland cotton and its core collection

性状 Trait	全部种质 Entire collection				核心种质 Core collection				F 值 F value	t 值 t value
	平均值 Mean	极差 Range	方差 Variation	变异系数 (%) CV	平均值 Mean	极差 Range	方差 Variation	变异系数 (%) CV		
比强度 FS	20.91	7 ~ 31	5.01	10.71	20.89	7 ~ 30.8	6.54	12.24	0.021	0.144
长度整齐度 LU	49.43	28.4 ~ 86.7	7.34	5.48	49.66	37.3 ~ 83.2	12.51	7.12	2.161	-1.159
果枝节位 SBN	6.62	2 ~ 13.4	1.87	20.65	6.84	2 ~ 12.6	2.68	23.93	7.492 *	-2.306 *
果枝伸展角 FPA	1.17	1 ~ 3	0.26	44.12	1.16	1 ~ 3	0.27	45.13	0.110	0.332
花蕊色 SC	1.85	1 ~ 6	0.79	48.07	1.99	1 ~ 6	0.98	49.64	6.424 *	-2.534 *
花药色 AC	1.07	1 ~ 5	0.08	26.33	1.07	1 ~ 5	0.11	31.17	0.015	-0.122
跨长 SL	28.23	10 ~ 36.6	4.76	7.73	28.17	10 ~ 36.6	7.69	9.84	0.170	0.332
铃形 BS	2.01	1 ~ 4	0.43	32.64	2.11	1 ~ 4	0.54	34.66	6.825 *	-2.361 *
铃重 BW	5.23	2 ~ 9	0.72	16.22	5.24	2 ~ 9	1.03	19.39	0.159	-0.399
麦克隆值 M	4.39	1.9 ~ 7.1	0.42	14.74	4.43	2.2 ~ 6.8	0.50	15.97	1.325	-1.065
耐盐性 SA	3.80	1 ~ 4	0.20	11.75	3.76	1 ~ 4	0.29	14.42	1.719	1.311
生育期 GP	133.15	94 ~ 192	119.45	8.21	132.25	94 ~ 163	128.81	8.58	1.843	1.357
叶厚度 LT	2.10	1 ~ 3	0.23	22.93	2.12	1 ~ 3	0.25	23.64	0.455	-0.675
叶片大小 LS	2.12	1 ~ 3	0.30	25.67	2.07	1 ~ 3	0.34	28.36	2.107	1.451
叶绒毛 LPA	2.84	1 ~ 5	1.00	35.29	3.00	1 ~ 5	1.20	36.42	7.285 *	-2.699
叶色 LC	2.07	1 ~ 8	0.36	29.12	2.11	1 ~ 8	0.62	37.29	0.800	-0.705
叶形 LS	1.05	1 ~ 8	0.13	34.61	1.07	1 ~ 8	0.26	47.40	0.464	-0.838
衣分 LP	36.19	13.6 ~ 47.8	15.41	10.85	36.53	18.0 ~ 47.8	27.29	14.30	3.370	-1.189
株高 PH	83.23	29.1 ~ 200.9	374.56	23.25	84.92	29.1 ~ 200.9	452.69	25.06	1.997	-1.413
株型 PT	1.91	1 ~ 3	0.09	15.98	1.93	1 ~ 3	0.08	15.06	1.204	-1.097
子指 SI	11.01	4.2 ~ 16.8	1.88	12.47	11.07	7.1 ~ 16.8	2.19	13.37	0.469	-0.685

表 6 原种质和核心种质主成分分析的特征值和累计贡献率

Table 6 Eigen value and cumulative contribution for the initial collection and the core collection

主成分 Component	原种质 Initial collection			核心种质 Core collection		
	特征值 Eigen value	贡献率(%) Contributive rate	累计贡献率(%) Cumulative contributive rate	特征值 Eigen value	贡献率(%) Contributive rate	累计贡献率(%) Cumulative contributive rate
1	2.436	10.149	10.149	2.224	9.268	9.27
2	2.182	9.091	19.239	2.005	8.354	17.62
3	1.863	7.762	27.002	1.812	7.551	25.17
4	1.633	6.805	33.807	1.763	7.346	32.52
5	1.473	6.137	39.944	1.639	6.829	39.35
6	1.283	5.344	45.289	1.356	5.652	45.00
7	1.107	4.613	49.902	1.327	5.530	50.53
8	1.041	4.337	54.239	1.189	4.955	55.48
9	1.019	4.247	58.486	1.110	4.625	60.11

入 9 个主成分,第一个主成分的特征值分别为 2.436 和 2.224,贡献率分别为 10.149% 和 9.27%,前 9 个主成分的累计贡献率分别是 58.486% 和 60.11%,因此构建核心种质能够有效地排除遗传冗余,提高原种质资源群体的累计贡献率。

3 讨论

3.1 核心种质的构建方法

核心种质的构建方法分为两种,一种是利用表型数据构建核心种质,另一种是利用基因型数据构

建核心种质,而现在的主流是先利用表型数据建立核心种质,然后再利用分子标记等技术对核心种质进行改进和补充。胥婷婷等<sup>[8]</sup>利用3个数量性状构建我国大麦的初级核心种质。徐宁等<sup>[9]</sup>利用15个表型性状构建中国小豆核心种质。刘遵春等<sup>[10]</sup>利用15数量性状构建新疆野苹果的核心种质。随着分子标记的出现及其技术日趋成熟,以其重复性好、速度快、效率高等优点在构建核心种质上得到广泛应用。黎毛毛等<sup>[15]</sup>利用分子标记技术构建了水稻核心种质。邱丽娟等<sup>[16]</sup>同样利用分子标记技术构建了大豆的微核心种质。王红霞等<sup>[22]</sup>基于 AFLP 分子标记技术构建核桃核心种质。本研究采用形态学的方法对构建陆地棉核心种质进行构建,种质资源群体间的遗传差异主要表现在个体间的形态特性、农艺性状的差异以及基因序列的差异。由于形态学上的农艺性状与实际应用相关联,而保存种质资源的目的是利用农艺性状的遗传多样性,因此,在构建陆地棉核心种质过程中利用农艺性状或形态学上的信息必不可少。利用表型数据构建核心种质具有简便、费用低、构建时限短等优点。本研究利用21个农艺性状构建陆地棉核心种质在5%左右的取样比例下,采用随机取样法取样,以欧氏距离进行聚类,以此方法构建的核心种质基本上能够代表陆地棉原种质资源的遗传多样性。但由于表型数据易受外界环境条件以及人为等因素的影响,不能准确反映全部种质的遗传本质,因此在构建核心种质时尽量选择遗传稳定的性状。

### 3.2 分组聚类与取样比例的确定

本研究中采用先对全部种质资源进行分组然后再对各组分别聚类,这样每一类遗传材料都能有比例的筛选出来,同时能够弥补某些性状不稳定带来的变异,通过这一方法就保持了核心种质多样性的遗传变异。

目前构建核心种质没有固定的取样比例,可以根据原种质资源群体的大小进行调整。国内外构建的各类作物的核心收集品占总种质资源的比例一般为10%~30%。H. D. Upadhyaya 等<sup>[32]</sup>用16991份种质群体构建花生核心种质的中选比例约为10%。Y. Zewdie 等<sup>[33]</sup>同样采用10%左右的比例构建高粱的核心种质,徐宁等<sup>[9]</sup>用8.92%的取样比例从4877份小豆筛选出核心种质435份。胡晋等<sup>[24]</sup>用样本量168的群体构建的棉花核心种质的中选率高达30%。而A. H. D. Brown 等<sup>[6]</sup>、O. H. Frankel 等<sup>[34]</sup>提出,当种质资源数目超过3000份时,核心种质的中选比例占整个种质资源的5%~10%,便能够代

表原种质的遗传变异。本研究利用5963份陆地棉种质资源构建陆地棉核心种质时,依据分组聚类结果每组随机取样,并根据多样性指数进行调整,确定最终取样比例为4.71%。结果表明这一取样比例已经能够代表全部种质资源的遗传变异。

### 3.3 核心种质的代表性

J. Hu 等<sup>[30]</sup>、徐海明等<sup>[31]</sup>提出用均值(Mean)、极差(Range)、表型方差(VPV)、变异系数(CV)等对评价核心种质的代表性,对核心种质与全部种质的均值进行 $t$ 测验,对差进行 $F$ 测验,同时可以计算核心种质与全部种质的极差符合率(CR)与变异系数变化率(VR),同时满足显著性0.05时,均值差异小于20%,同时 $CR > 80\%$ ,可认为核心种质能够代表原种质群体的遗传变异。刘遵春等<sup>[10]</sup>用均值差异百分率(MD)、方差差异百分率(VD)、极差符合率(CR)、变异系数变化率(VR)等4个参数作为构建新疆野苹果核心种质的评价指标。本研究利用均值 $t$ 测验、方差 $F$ 测验、变异系数(CV)、多样性指数 $t$ 检验、均值(Mean)、极差(Range)、表型方差(VPV)、变异系数(CV)、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)、主成分分析等参数进行评价,认为所构建核心种质能够较好地代表原种质资源群体的遗传多样性。

主成分分析(PCA, principal component analysis)的空间分布能够近似地反映群体之间的相似程度,目前主成分分析已经应用于核心种质的构建及其代表性评价。刘遵春等<sup>[10]</sup>在构建新疆野苹果核心种质时对15个性状进行主成分分析,取前5个主成分对构建的核心种质进行确认。郑轶琦等<sup>[35]</sup>利用表型性状构建狗牙根核心种质时同样用15个性状取前5个主成分确认核心种质。D. H. Basigalup 等<sup>[36]</sup>在多年生苜蓿群体中,利用主成分分析构建的核心种质优于其他方法。H. D. Upadhyaya 等<sup>[32]</sup>利用标准化的14个数量性状对龙爪樱种质资源进行主成分分析,取前5个主成分分析并结合聚类建立核心种质库。胡标林等<sup>[37]</sup>利用14个表型性状对前9个主成分进行分析,以此来评价水稻核心种质表型性状的遗传多样性。M. Malosetti 等<sup>[38]</sup>对17个表型性状进行主成分分析,取前11个主成分构建玉米核心种质。H. M. Xu 等<sup>[39]</sup>通过主成分分析评价海岛棉核心种质的代表性。J. C. Wang 等<sup>[40]</sup>利用主成分分析进一步验证利用欧氏距离构建棉花核心种质的有效性。本研究利用主成分分析对所构建陆地棉核心种质进行确认,避免核心种质中过多的冗余,与以上

研究结果相一致。

## 4 结论

用表型数据构建的陆地棉核心种质保留原始种质 4.71% 的遗传资源,核心种质与全部种质的均值差异百分率(*MD*)、方差差异百分率(*VD*)、极差符合率(*CR*)、变异系数变化率(*VR*)分别为 85.7%、85.7%、95.8%、114.1%。核心种质的中选比例为原始种质的 4.71%,因此核心种质可以代表全部种质资源的遗传多样性,从而促进了棉花种质资源的高效利用。

## 参考文献

- [1] Cai C, Ye W, Zhang T, et al. Association analysis of fiber quality traits and exploration of elite alleles in upland cotton cultivars/ accessions (*Gossypium hirsutum* L.) [J]. J Integr Plant Biol, 2014, 56(1): 51-62
- [2] 杜雄明, 周忠丽, 贾银华, 等. 中国棉花种质资源的收集与保存[J]. 棉花学报, 2007, 19(5): 346-353
- [3] 杜雄明, 孙君灵, 周忠丽, 等. 棉花资源收集、保存、评价与利用现状及未来[J]. 植物遗传资源学报, 2012, 13(2): 163-168
- [4] Brown A H D. Core collections: a practical approach to genetic resources management [J]. Genome, 1989, 31(2): 818-824
- [5] Frankel O H. Genetic perspectives of germplasm conservation. [M]//Arber W, Limensee K, peacock W J, et al. Symposium on Genetic Manipulation: Impact on Man and Society. Cambridge, UK: Cambridge University Press, 1984: 161-170
- [6] Brown A H D. The case for core collections [M]//Brown A H D, Frankel O H, Marshall D R, et al. The Use of Plant Genetic Resources. Cambridge, UK: Cambridge University Press, 1989: 136-156
- [7] 李自超, 张洪亮, 曾亚文, 等. 云南地方稻种质资源核心种质取样方案研究[J]. 中国农业科学, 2000, 33(5): 1-7
- [8] 胥婷婷, 林峰, 华为, 等. 我国大麦选育品种核心种质的初步构建[J]. 浙江农业学报, 2011, 23(3): 483-488
- [9] 徐宁, 程须珍, 王素华, 等. 以地理来源分组和利用表型数据构建中国小豆核心种质 [J]. 作物学报, 2008, 34(8): 1366-1373
- [10] 刘遵春, 张春雨, 张艳敏, 等. 利用数量性状构建新疆野苹果核心种质的方法 [J]. 中国农业科学, 2010, 43(2): 358-370
- [11] 张永兵, 伊鸿平, 马新力, 等. 新疆甜瓜地方品种资源核心种质构建[J]. 植物遗传资源学报, 2013, 14(1): 52-57
- [12] 李秀兰, 贾继文, 王军辉, 等. 灰楸形态多样性分析及核心种质初步构建[J]. 植物遗传资源学报, 2013, 14(2): 243-248
- [13] 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
- [14] 刘鑫铭, 刘崇怀, 樊秀彩, 等. 葡萄种质资源初级核心群的构建[J]. 植物遗传资源学报, 2012, 13(1): 72-76
- [15] 黎毛毛, 黄永兰, 余丽琴, 等. 利用 SSR 标记构建江西稻种质资源核心种质库的研究[J]. 植物遗传资源学报, 2012, 13(6): 952-957
- [16] 邱丽娟, 李英慧, 关荣霞, 等. 大豆核心种质和微核心种质的构建、验证与研究进展[J]. 作物学报, 2009, 35(4): 571-579
- [17] Cristo-Araújo, M. de, Rodrigues, D. P, Astolfi Filho, S, et al. Peach palm core collection in Brazilian Amazonia [J]. Crop Breed Appl Biotechnol, 2015, 15(1): 18-25
- [18] Hao C Y, Zhang X Y, Wang L F, et al. Genetic diversity and core collection evaluations in common wheat germplasm from the northwest-ern spring wheat region in China [J]. Mol Breed, 2006, 17: 69-77
- [19] Igartua E, Gracia M P, Lasa J M, et al. The Spanish barley core collection [J]. Genet Resour Crop Evol, 1998, 45: 475-481
- [20] Li Y, Shi Y S, Cao Y S, et al. Establishment of a core collection for maize germplasm preserved in Chinese National Genebank using geographic distribution and characterization data [J]. Genet Resour Crop Evol, 2004, 51: 845-852
- [21] Coimbra R R, Miranda G V, Cruz C D, et al. Development of a Brazilian maize core collection [J]. Genet Mol Biol, 2009, 32: 538-545
- [22] 王红霞, 赵书岗, 高仪, 等. 基于 AFLP 分子标记的核桃核心种质的构建 [J]. 中国农业科学, 2013, 46(23): 4985-4995
- [23] Frankel O H, Brown A H D. Plant genetic resources today: A critical appraisal [M]//Holden J H W, Williams J T. Crop Genetic Resources: Conservation and Evaluation. London: George Allen and Unwin, 1984: 249-257
- [24] 胡晋, 徐海明, 朱军, 等. 保留特殊种质材料的核心种质构建方法 [J]. 生物数学学报, 2001, 16(3): 348-352
- [25] 董玉琛, 曹永生, 张学勇, 等. 中国普通小麦核心种质的产生 [J]. 植物遗传资源学报, 2003, 4(1): 1-8
- [26] 王建成, 胡晋, 张彩芳, 等. 建立在基因型值和分子标记信息上的水稻核心种质评价参数 [J]. 中国水稻科学, 2007, 21(1): 51-58
- [27] 韩立德, 徐海明, 胡晋, 等. 核心种质数量性状代表性评价指标的研究 [J]. 生物数学学报, 2006, 21(4): 603-609
- [28] Jain S K, Qualset C O, Bhatt G M, et al. Geographical patterns of phenotypic diversity in a world collection of durum wheats [J]. Crop Sci, 1975, 15(5): 700-704
- [29] 徐海明, 胡晋, 朱军, 等. 构建作物种质资源核心库的一种有效抽样方法 [J]. 作物学报, 2000, 26(2): 157-162
- [30] Hu J, Zhu J, Xu H M, et al. 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(1): 264-268
- [31] 徐海明, 邱英雄, 胡晋, 等. 不同遗传距离聚类法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
- [32] Upadhyaya H D, Ortiz R. A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement [J]. Theor Appl Genet, 2001, 102(8): 1292-1298
- [33] Zewdie Y, Tong N K, Bosland P, et al. Establishing a core collection of Capsicum using a cluster analysis with enlightened selection of accessions [J]. Genet Resour Crop Evol, 2004, 51(2): 147-151
- [34] Frankel O H, Brown A H D. Current plant genetic resources a critical appraisal [M]//Chopra V L. Genetics New Frontiers. New Delhi, India: Oxford and IBH Publishing Co, 1984
- [35] 郑铁琦, 郭琰, 房淑娟, 等. 利用表型数据构建狗牙根核心种质 [J]. 草业学报, 2014, 23(4): 49-60
- [36] Basigalup D H, Barnes D K, Stucker R E, et al. Development of a core collection for perennial Medicago plant introductions [J]. Crop Sci, 1995, 35: 1163-1168
- [37] 胡标林, 万勇, 李霞, 等. 水稻核心种质表型性状遗传多样性分析及综合评价 [J]. 作物学报, 2012, 38(5): 829-839
- [38] Malosetti M, Abadie T. Sampling strategy to develop a core collection of Uruguayan maize landraces based on morphological traits [J]. Genet Resour Crop Evol, 2001, 48(4): 381-390
- [39] Xu H M, Mei Y J, Hu J, et al. Sampling a core collection of Island cotton (*Gossypium barbadense* L.) based on the genotypic values of fiber traits [J]. Genet Resour Crop Evol, 2006, 53(3): 515-521
- [40] Wang J C, Jin H U, Huang X X, et al. Assessment of different genetic distances in constructing cotton core subset by genotypic values [J]. J Zhejiang Univer Sci B, 2008, 9(5): 356-362