

李 飞, 王清连, 李成奇. 陆地棉品种(系)资源的主成分分析和聚类分析[J]. 江苏农业学报, 2015, 31(6): 1211-1217.  
doi: 10.3969/j.issn.1000-4440.2015.06.004

## 陆地棉品种(系)资源的主成分分析和聚类分析

李 飞, 王清连, 李成奇

(现代生物育种河南省协同创新中心, 河南科技学院棉花研究所, 河南 新乡 453003)

**摘要:** 为了客观评价棉花品种(系)资源, 为棉花早熟、高产育种的亲本选配提供理论依据, 选取近年来育成的来自黄河流域、长江流域、西北内陆和北部特早熟 4 大棉区以及国外引进的 172 份陆地棉骨干品种(系)为材料, 对棉花 19 个农艺性状进行了主成分分析和聚类分析。主成分分析结果表明, 19 个性状可以简化为彼此互不相关的 5 个主成分, 这 5 个主成分累计贡献率达 80.253%, 第 2 主成分得分值高的材料生育期短、产量高。聚类分析结果显示, 在遗传距离 5.62 处, 172 份材料划分为 10 个类群, 第 V 类群材料生育期短、植株矮、产量适中, 第 VII 类群材料生育期适中、产量高。

**关键词:** 陆地棉; 品种(系)资源; 农艺性状; 主成分分析; 聚类分析

**中图分类号:** S562.024 **文献标识码:** A **文章编号:** 1000-4440(2015)06-1211-07

## Principal component and clustering analyses of upland cotton variety (line) resources

LI Fei, WANG Qing-lian, LI Cheng-qi

(Collaborative Innovation Center of Modern Biological Breeding, Henan Province/Cotton Research Institute, Henan Institute of Science and Technology, Xinxiang 453003, China)

**Abstract:** To objectively evaluate upland cotton varieties (lines) and provide a reference for parental materials selection with early maturity and high yield, 172 backbone upland cotton varieties (lines) bred in recent years from four cotton areas, Yellow River, Yangtze River, Northwestern inland and Northern extremely early maturation in China and introduced from abroad, were selected as experimental materials, and 19 agronomic traits were tested for principal component analysis and clustering analysis. 19 characters were integrated into 5 principal components with the cumulative contribution of 80.253%. The high-scored materials of the second component were characterized by short growth period and high yield. 172 materials were classified into 10 clusters at the genetic distance of 5.62; materials of cluster V had the characters of short growth period, low plant height and moderate yield, and materials of cluster VII was type of middle growth period and high yield.

**Key words:** upland cotton; variety (line) resource; agronomic trait; principal component; clustering

收稿日期: 2015-08-22

**基金项目:** 国家自然科学基金项目(31371677); 河南省科学技术研究重点项目(14A210006); 河南省高校青年骨干教师资助计划项目(2013GGJS-136)

**作者简介:** 李 飞(1990-), 男, 河南郑州人, 硕士研究生, 主要从事棉花分子育种研究。(Tel) 15670566929; (E-mail) 15670566929@163.com

**通讯作者:** 李成奇, (E-mail) lichq2010@126.com

品种(系)资源是棉花育种的重要基础材料, 合理利用品种(系)资源, 选择适宜的亲本进行杂交是育种工作的关键。目前, 中国种植的陆地棉品种均是由美国的岱字棉、斯字棉、金字棉等种质资源衍生而来, 遗传基础较狭窄<sup>[1-3]</sup>, 限制了亲本材料的有效利用。因此, 对品种(系)资源进行客观评价具有现实意义。主成分分析是一种利用降维的思想, 在损失较少信息的前提下, 把多个指标转化为几个相互

无关的综合指标的多元统计方法<sup>[4-8]</sup>。聚类分析是一种用于生物资源分类和亲缘关系研究的多元统计方法<sup>[9-10]</sup>。近年来,同时利用主成分分析和聚类分析研究小麦<sup>[11-12]</sup>、水稻<sup>[13]</sup>、大豆<sup>[14-15]</sup>、黍稷<sup>[16]</sup>等作物物质资源的报道较多,在棉花上也有一些报道。Brown<sup>[17]</sup>对美国区域试验棉花品种的农艺和纤维品质性状进行了主成分和聚类分析,结果表明,来自密西西比河三角洲、中部和德克萨斯高平原地区品种的遗传基础高于来自东部、墨西哥和圣华金河地区的品种。孙长发等<sup>[18]</sup>调查了河南省春棉区域试验棉花品种的 17 个数量性状,用主成分分析法提取了累计贡献率达 90.10% 的前 4 个主成分,聚类分析法将所有材料分为 5 类。许乃银等<sup>[19]</sup>利用长江流域棉花区域试验品种的纤维品质数据进行了主成分和聚类分析,从 8 个纤维品质性状中提取了 4 个主成

分,并将 20 个试点划分为 4 个纤维品质相似亚区。上述研究均是以参加区域试验的少数品种为材料评价种质资源,对棉花育种实践的指导意义不大。本研究选用中国近年来育成或引进的 172 份陆地棉骨干品种(系)为材料,对棉花 19 个农艺性状进行主成分分析和聚类分析,为深入评价陆地棉品种间(系)的遗传多样性和合理选配亲本组合提供理论依据。

## 1 材料与方法

### 1.1 试验材料

供试的 172 份陆地棉品种(系)中,64 份来自黄河流域棉区,25 份来自长江流域棉区,55 份来自西北内陆棉区,18 份来自北部特早熟棉区,10 份为国外引进材料(表 1)。所有材料均经过多代自交。

表 1 供试材料的来源、数目及名称

Table 1 Source, number and name of experimental materials

来源	数目	名称
黄河流域	64	百棉 1 号、百棉 2 号、百棉 5 号、汾无 195、国欣棉 3 号、邯鄹 802、邯鄹 885、冀棉 1 号、冀棉 7 号、冀棉 12、冀棉 958、晋棉 13 号、晋棉 29 号、晋棉 36 号、晋棉 45 号、鲁棉 1 号、鲁棉 4 号、鲁棉 6 号、鲁棉 10 号、鲁棉研 21 号、鲁棉研 28 号、鲁棉研 29、陕 1155、陕 2365、陕棉 4 号、石远 321、系 9、鑫秋 1 号、豫棉 1 号、豫棉 5 号、豫棉 7 号、豫棉 9 号、豫棉 12 号、豫棉 21 号、中 1707、中棉所 8 号、中棉所 10 号、中棉所 12、中棉所 13、中棉所 14、中棉所 15、中棉所 17、中棉所 18、中棉所 19、中棉所 20、中棉所 22、中棉所 23、中棉所 24、中棉所 25、中棉所 26、中棉所 27、中棉所 30、中棉所 31、中棉所 33、中棉所 34、中棉所 35、中棉所 36、中棉所 37、中棉所 40、中棉所 42、中棉所 50、中棉所 58、中棉所 64、中植棉 2 号
长江流域	25	川棉 56、岱红岱、洞庭 1 号、赣棉 8 号、江苏棉 1 号、荆 8891、钱江 9 号、黔农 465、泗棉 2、泗棉 3、泗棉 4、苏棉 1 号、苏棉 6 号、苏棉 9 号、苏棉 10 号、苏棉 12 号、苏棉 16、湘棉 3、湘棉 10 号、徐州 142、盐棉 48、鄂荆 1 号、鄂棉 3 号、鄂棉 14、鄂沙 28
西北内陆	55	新陆早 1 号、新陆早 2 号、新陆早 3 号、新陆早 4 号、新陆早 5 号、新陆早 6 号、新陆早 7 号、新陆早 8 号、新陆早 9 号、新陆早 10 号、新陆早 11 号、新陆早 12 号、新陆早 13 号、新陆早 15 号、新陆早 16 号、新陆早 17 号、新陆早 18 号、新陆早 19 号、新陆早 20 号、新陆早 21 号、新陆早 22 号、新陆早 23 号、新陆早 24 号、新陆早 25 号、新陆早 26 号、新陆早 27 号、新陆早 28 号、新陆早 29 号、新陆早 30 号、新陆早 31 号、新陆早 32 号、新陆早 33 号、新陆早 34 号、新陆早 35 号、新陆早 36 号、新陆早 37 号、新陆早 38 号、新陆早 39 号、新陆早 40 号、新陆早 41 号、新陆早 42 号、新陆早 45 号、新陆早 46 号、新陆早 47 号、新陆早 48 号、新陆早 49 号、新陆早 51 号、新陆中 36 号、18-3、拉玛干 77、锦棉 1 号、锦棉 2 号、锦棉 4 号、锦棉 5 号、绿早 254
北部特早熟	18	黑山棉、辽棉 4 号、辽棉 5 号、辽棉 7 号、辽棉 8 号、辽棉 10 号、辽棉 16 号、辽棉 18 号、辽棉 19 号、辽棉 23 号、晋中 169、晋中 200、晋棉 5 号、晋棉 6 号、晋棉 8 号、晋棉 9 号、晋棉 14 号、晋棉 24 号
国外引进	10	kk1543、斯字棉 2B、岱字棉 15、岱字棉 16、贝尔斯诺、乌干达 3 号、石选 87、99M4、99M7、99M8

### 1.2 试验方法

分别于 2012、2013 年将 172 份材料种植在河南科技学院棉花育种试验田。随机区组设计,单行区,3 次重复;行长 5 m,行距 1 m,每行 14~16 株,大田

常规管理。对 2 年所有材料,分别考查株高、主茎长、主茎节数、总果枝数、有效果枝数、果枝长、果枝节数、苗期(出苗期~现蕾期)、蕾期(现蕾期~开花期)、花铃期(开花期~吐絮期)、生育期(出苗期~吐

絮期)、果枝始节、始节高度、单株籽棉产量、单株皮棉产量、单铃质量、衣分、衣指和籽指等 19 个农艺性状。将 2 年 6 次重复各性状的平均值作为该性状的最终表型值。

1.3 数据处理

利用 Excel 2010 进行数据基本整理。利用 SPSS 19.0 软件将各性状标准化后,采用降维的方法进行主成分分析,基于主成分特征向量矩阵和各农艺性状的标准化值,根据公式  $F=Z \cdot A$  计算各供试材料的主成分得分值,其中  $F$ 、 $Z$ 、 $A$  分别为主成分得分值矩阵、标准化矩阵和特征向量矩阵<sup>[18,20]</sup>。利用 NTSYSpc 2.10e 软件将数据标准化后,计算欧式(Euclid)遗传距离,以类平均法(UPGMA)对 172 份材料进行聚类<sup>[21]</sup>。

2 结果与分析

2.1 供试材料农艺性状的描述统计

172 份供试材料 19 个农艺性状的描述统计结果列于表 2。由表 2 可知,供试材料在 19 个农艺性状上均表现较大差异,变异系数由大到小依次为皮棉产量、籽棉产量、果枝长、始节高度、有效果枝数、果枝节数、果枝始节、主茎长、株高、衣指、单铃质量、衣分、籽指、主茎节数、苗期、蕾期、花铃期、总果枝数、生育期。其中,皮棉产量变异系数最大(29.92%),其次为籽棉产量(26.37%),生育期变异系数最小(5.08%),说明棉花的产量受基因型和环境影响最大,生育期受基因型和环境影响最小。

2.2 主成分分析

对 19 个农艺性状数据标准化后进行主成分分析,获得 19 阶遗传相关矩阵及其特征值和特征向量(表 3)。前 5 个主成分特征值大于 1,累计贡献率达到 80.253%,可反映大部分遗传差异信息。表 4 列出了各主成分中排名前 3 的代表性材料及其主成分得分值。得分值越高,材料受此主成分的影响越大。由表 3 和表 4 看出,第 1 主成分的特征值为 5.856,贡献率为 30.822%;特征向量除籽指外均为正值,其中果枝长、生育期和果枝始节的特征向量较大,分别为 0.759、0.816 和 0.799,说明第 1 主成分得分值高的材料表现为果枝长、生育期长、果枝始节位高,代表性材料有荆 8891、鲁棉 21 号和鑫秋 1 号。第 2 主成分的特征值为 3.450,贡献率为 18.157%;有效果枝数、籽棉产量、皮棉产量和衣指的特征向量较大

表 2 供试材料 19 个农艺性状的描述统计

Table 2 Descriptive statistics of 19 agronomic traits of experimental materials

性状	最小值	最大值	平均值	标准差	变异系数
株高(cm)	74.46	131.25	106.41	11.31	10.63
主茎长(cm)	57.91	103.26	82.62	8.84	10.70
主茎节数	11.70	15.83	13.71	0.92	6.71
总果枝数	12.70	16.83	14.71	0.92	6.25
有效果枝数	7.35	14.40	10.60	1.38	13.02
果枝长(cm)	24.98	59.24	42.72	7.43	17.39
果枝节数	3.90	7.60	5.34	0.66	12.36
苗期(d)	30.00	40.75	35.12	2.31	6.58
蕾期(d)	20.50	28.00	24.01	1.56	6.50
花铃期(d)	40.00	53.75	45.93	2.90	6.31
生育期(d)	94.50	118.00	105.08	5.34	5.08
果枝始节	4.40	7.45	5.90	0.71	12.03
始节高度(cm)	8.99	21.95	15.96	2.70	16.92
籽棉产量(g)	22.96	119.64	71.45	18.84	26.37
皮棉产量(g)	11.29	46.97	26.40	7.90	29.92
单铃质量(g)	3.96	7.42	5.93	0.55	9.27
衣分(%)	28.70	42.55	36.64	2.90	7.91
衣指(g)	4.40	8.58	6.60	0.69	10.45
籽指(g)	9.79	13.79	11.41	0.88	7.71

且为正值,株高、主茎长、果枝长、果枝节数、苗期、蕾期、花铃期、生育期和始节高度的特征向量为负值,说明第 2 主成分得分值高的材料产量高、生育期短、株型较好,代表性材料有百棉 5 号、冀棉 958 和岱字棉 15。第 3 主成分的特征值为 2.607,贡献率为 13.719%;株高、主茎长、主茎节数和总果枝数的特征向量较大且为正值,说明第 3 主成分得分值高的材料植株高,主茎节数和总果枝数较大,代表性材料有晋棉 9 号、新陆早 48 号和新陆早 33 号。第 4 主成分的特征值为 1.995,贡献率为 10.500%;衣分和籽指的特征向量较大,其中前者为负值,后者为正值,说明第 4 主成分得分值高的材料衣分低而籽指高,代表性材料有岱红岱、乌干达 3 号和晋棉 14 号。第 5 主成分的特征值为 1.341,贡献率为 7.056%;特征向量为正值且较大的性状有株高、主茎长、始节高度、单铃质量、衣指和籽指,特征向量为负值且绝对值较大的性状有主茎节数、总果枝数、有效果枝数、果枝长和果枝节数,说明第 5 主成分得分值高的

表 3 19 个农艺性状的主成分分析

Table 3 Principal component analysis of 19 agronomic traits

项目	主成分				
	1	2	3	4	5
株高	0.558	-0.363	0.575	-0.233	0.351
主茎长	0.524	-0.280	0.649	-0.200	0.331
主茎节数	0.233	0.346	0.782	0.311	-0.247
总果枝数	0.233	0.346	0.782	0.311	-0.247
有效果枝数	0.566	0.578	0.264	-0.053	-0.210
果枝长	0.759	-0.358	-0.024	0.000	-0.218
果枝节数	0.400	-0.566	0.099	0.036	-0.343
苗期	0.751	-0.211	-0.336	0.004	-0.146
蕾期	0.531	-0.454	-0.131	0.059	0.066
花铃期	0.616	-0.033	-0.284	0.467	-0.175
生育期	0.816	-0.246	-0.344	0.264	-0.146
果枝始节	0.799	-0.139	-0.152	0.131	-0.052
始节高度	0.612	-0.480	0.167	-0.261	0.359
籽棉产量	0.509	0.723	-0.083	-0.035	0.007
皮棉产量	0.559	0.724	-0.096	-0.194	-0.028
单铃质量	0.439	0.421	-0.289	0.429	0.359
衣分	0.439	0.386	-0.179	-0.681	-0.039
衣指	0.495	0.524	-0.199	-0.248	0.358
籽指	-0.025	0.097	0.019	0.761	0.556
特征值	5.856	3.450	2.607	1.995	1.341
贡献率(%)	30.822	18.157	13.719	10.500	7.056
累计贡献率(%)	30.822	48.978	62.697	73.197	80.253

材料植株较高,铃质量较高,衣指和籽指较高,代表性材料有新陆早 28 号、新陆早 11 号和新陆早 27 号。

### 2.3 聚类分析

对 172 份陆地棉品种(系)的 19 个性状数据标准化后,进行了聚类分析。在遗传距离为 5.62 时可将所有材料分为 10 个类群,不同类群材料农艺性状的平均值见表 5。第 I 类群包括 4 个材料:kk1543、晋棉 6 号、锦棉 2 号和黑山棉,其中,晋棉 6 号和黑山棉来自北部特早熟棉区,锦棉 2 号来自西北内陆棉区,kk1543 为前苏联早期品种。该类群材料植株较矮,果枝长度短,单铃质量偏低,生育期较短,衣分和衣指较低。第 II 类群包括 47 个材料,分别是新陆早 1 号、新陆早 5 号、新陆早 2 号、新陆早 4 号、新陆早 51 号、新陆早 6 号、新陆早 9 号、新陆早 37 号、新

表 4 各主成分中排名前 3 的代表性材料及其主成分得分值

Table 4 Representative materials and their principal component scores of the top-3 in each principal component

代表性材料	主成分得分值				
	1	2	3	4	5
荆 8891	<b>13.378</b>	-2.106	-1.775	2.203	1.375
鲁棉 21 号	<b>11.214</b>	2.355	-2.248	-1.123	0.195
鑫秋 1 号	<b>10.755</b>	3.342	3.259	-1.956	0.450
百棉 5 号	6.084	<b>6.998</b>	-0.732	0.621	1.606
冀棉 958	7.237	<b>6.913</b>	0.255	1.941	0.995
岱字棉 15	2.143	<b>6.646</b>	-0.517	-0.616	-0.868
晋棉 9 号	-3.170	-3.215	<b>7.052</b>	3.238	-0.894
新陆早 48 号	-0.016	-2.302	<b>6.256</b>	-3.696	0.037
新陆早 33 号	-0.678	5.756	<b>5.445</b>	0.218	0.545
岱红岱	-2.643	-4.298	-0.988	<b>7.087</b>	-0.746
乌干达 3 号	3.588	-6.919	-2.114	<b>5.168</b>	1.830
晋棉 14 号	-3.968	-4.478	4.637	<b>4.823</b>	-0.795
新陆早 28 号	-3.834	-0.353	-2.815	-3.040	<b>3.496</b>
新陆早 11 号	0.712	-4.236	1.765	-0.176	<b>3.322</b>
新陆早 27 号	-7.140	-0.307	-0.798	-2.320	<b>3.002</b>

加粗数字为排名前 3 的主成分得分值。

陆早 38 号、新陆早 7 号、晋中 200、新陆早 46 号、拉玛干 77、新陆早 8 号、中棉所 13、18-3、新陆早 42 号、新陆早 10 号、新陆早 40 号、系 9、新陆早 20 号、新陆早 13 号、中棉所 27、岱字棉 16、新陆早 19 号、新陆早 34 号、新陆早 36 号、新陆早 21 号、中棉所 22、新陆早 22 号、新陆早 32 号、新陆早 45 号、新陆早 47 号、新陆早 15 号、新陆早 18 号、新陆早 23 号、新陆早 24 号、新陆早 25 号、苏棉 16、新陆早 35 号、新陆早 27 号、新陆早 28 号、新陆早 39 号、新陆早 31 号、新陆早 11 号、新陆早 26 号、中棉所 19。绝大部分为西北内陆棉区的新陆早系列,所有性状表现均适中。第 III 类群包括 2 个材料,新陆早 3 号和新陆早 49 号,表现为植株较高,果枝长度最长,果枝节数最多,籽棉产量和皮棉产量较低。第 IV 类群仅有 1 个材料新陆早 48 号,表现为植株最高,总果枝数最多,果枝节数较少,生育期短,始节高度最高,单铃质量较低,衣分最高。第 V 类群包括 30 个材料,来自西北内陆和黄河流域 2 个棉区,分别是新陆早 12 号、中棉所 31、辽棉 18 号、晋棉 5 号、中棉所 33、中棉所 26、百棉 2 号、豫棉 5 号、辽棉 10 号、辽棉 16



表 5 不同类群材料农艺性状的平均值

Table 5 Averages of agronomic traits of clustered materials

性状	类群									
	I	II	III	IV	V	VI	VII	VIII	IX	X
株高(cm)	93.68	112.89	121.54	131.03	90.49	108.45	109.68	107.21	103.58	91.28
主茎长(cm)	73.96	87.12	94.09	101.88	70.60	86.13	85.55	82.55	76.95	69.06
主茎节数	13.72	13.34	14.44	15.02	13.29	14.89	14.40	13.21	12.72	14.53
总果枝数	14.72	14.34	15.44	16.02	14.29	15.89	15.40	14.21	13.72	15.53
有效果枝数	8.98	9.94	10.30	10.19	10.04	10.66	12.12	10.50	8.00	7.93
果枝长(cm)	32.51	42.53	52.89	37.39	33.76	37.81	45.76	50.13	47.78	42.32
果枝节数	5.20	5.47	7.39	4.94	4.71	5.08	5.25	5.83	5.65	6.06
苗期(d)	33.06	34.35	33.25	35.50	33.23	32.83	36.09	37.91	38.75	33.75
蕾期(d)	23.06	24.31	24.88	22.00	22.54	22.88	24.05	25.33	26.63	25.00
花铃期(d)	44.06	43.94	44.00	43.00	43.98	45.06	47.49	49.31	45.88	53.75
生育期(d)	100.19	102.68	102.13	100.50	99.77	100.67	107.61	112.62	111.25	112.50
果枝始节	5.03	5.57	5.44	6.45	5.32	5.48	6.33	6.57	7.38	5.80
始节高度(cm)	12.75	17.62	16.95	19.87	12.45	14.82	16.31	16.92	18.19	15.67
籽棉产量(g)	51.30	61.16	53.88	50.69	69.73	61.44	94.89	66.72	45.54	39.77
皮棉产量(g)	17.01	22.16	19.87	19.87	25.43	20.80	36.42	24.96	15.75	12.48
单铃质量(g)	4.84	5.62	5.56	4.70	5.93	6.11	6.21	6.20	5.14	5.98
衣分(%)	30.69	36.35	36.50	39.25	36.46	33.35	38.35	37.28	34.27	29.77
衣指(g)	4.76	6.40	6.24	6.42	6.57	6.38	7.07	6.72	5.99	6.02
籽指(g)	10.75	11.18	10.88	9.94	11.47	12.71	11.37	11.32	11.47	13.54

号、中棉所 30、中棉所 64、中棉所 37、辽棉 4 号、辽棉 5 号、豫棉 12 号、锦棉 4 号、辽棉 7 号、新陆早 17 号、新陆早 30 号、中棉所 42、中棉所 50、中棉所 58、中 1707、中棉所 24、中棉所 36、中棉所 10 号、中棉所 20、豫棉 7 号、中棉所 14。表现为植株最矮,果枝长度较短,果枝节数最少,生育期最短。第 VI 类群由 12 个材料构成,分别是新陆早 29 号、新陆早 33 号、石选 87、晋棉 24 号、贝尔斯诺、中棉所 18、辽棉 8 号、绿早 254、锦棉 5 号、黔农 465、晋棉 9 号、晋棉 14 号。表现为总果枝数较多,生育期较短,单铃质量较高。第 VII 类群包括 43 个材料,分别是中棉所 12、石远 321、江苏棉 1 号、鲁棉 1 号、鲁棉研 29、中棉所 23、冀棉 12、鄂棉 3 号、鄂沙 28、陕 1155、PPM8、新陆早 16 号、中棉所 40、新陆中 36 号、泗棉 2、苏棉 9 号、赣棉 8 号、徐州 142、豫棉 21 号、川棉 56、鑫秋 1 号、新陆早 41 号、百棉 5 号、晋棉 45 号、冀棉 958、国欣棉 3 号、邯郸 802、中棉所 35、晋棉 13 号、晋棉 29、

陕 2365、邯郸 885、苏棉 1 号、晋棉 36 号、中植棉 2 号、岱字棉 15、辽棉 19 号、百棉 1 号、鲁棉 21 号、鲁棉 28 号、泗棉 4、辽棉 23 号、豫棉 9 号。大部分来自于黄河流域棉区,表现为有效果枝数最多,籽棉产量和皮棉产量最高,单铃质量最高,衣分和衣指最大。第 VIII 类群包括 30 个材料,来自中国四大棉区以及国外引进,区域差异较大,具有生育期最长,单铃质量较高的特点,分别是中棉所 15、湘棉 3、盐棉 48、中棉所 25、苏棉 10 号、中棉所 34、湘棉 10 号、洞庭 1 号、锦棉 1 号、汾无 195、陕棉 4 号、鄂棉 14、中棉所 17、豫棉 1 号、钱江 9 号、泗棉 3、冀棉 7 号、晋棉 8 号、冀棉 1 号、斯字棉 2B、99M7、晋中 169、乌干达 3 号、鄂荆 1 号、荆 8891、鲁棉 4 号、鲁棉 10 号、鲁棉 6 号、中棉所 8 号、苏棉 6 号。第 IX 类群包括 2 个材料,苏棉 12 号 and 美国的 99M4,表现为总果枝数最少,生育期长,果枝始节最高、籽棉产量和皮棉产量较低。第 X 类群仅有 1 个材料岱红岱,表现为植株

较矮,有效果枝数最少,生育期长,籽棉产量和皮棉产量最低、衣指最高。

### 3 讨论

主成分分析是将多个复杂的原始指标转化为几个具有代表性的互不相关的综合指标(主成分),每个主成分都是原始指标的线性组合,保留了原始指标的大部分信息,从而减小分析的复杂性<sup>[4]</sup>。Panthee 等<sup>[22]</sup>收集了尼泊尔的 179 份大蒜材料,调查 8 个农艺性状,主成分分析提取的 4 个主成分可以解释超过 86% 的总变异。赵德新等<sup>[23]</sup>对 55 个茄子种质材料的 18 个形态学性状进行了主成分分析,将 18 个性状简化为 6 个主成分,代表 81.071% 的遗传变异。本研究利用主成分分析将棉花 19 个农艺性状简化为 5 个主成分,累计贡献率达 80.253%,反映了大部分遗传变异信息。当前,中国人多地少,粮棉争地矛盾日益突出。为协调粮棉发展,在确保粮食生产安全的前提下,培育早熟、高产的棉花品种至关重要。由本研究主成分分析结果可以看出,第 2 主成分得分值高的材料生育期短、产量高。因此,在进行早熟、高产育种时,可考虑第 2 主成分得分较高的材料,如百棉 5 号、冀棉 958 和岱字棉 15。

聚类分析发现,一些生态区域相同,亲缘关系较近的材料被聚在一起。例如来自河南省的百棉 2 号和豫棉 5 号在遗传距离为 1.91 处被聚在一起,来自北部特早熟棉区的辽棉 18 号和晋棉 5 号在遗传距离为 2.14 处被聚在一起,新陆早系列的新陆早 15 号、18 号、23 号和 24 号在遗传距离为 2.87 处被聚在一起。说明本研究利用包括株高在内的 19 个农艺性状对陆地棉种质资源进行聚类,具有一定可行性。同时,从聚类的 10 个类群中发现,一些来自不同生态区的材料被划分在同一类群。如第 I 类群的 4 个材料中,2 个来自于北部特早熟棉区,1 个来自于西北内陆棉区,1 个为前苏联品种;第 VIII 类群包含了来自中国四大棉区和国外引进的材料。说明陆地棉材料间的遗传差异与地理来源关系不大,亲本材料应从不同的类群中选择,这与其他学者的研究结果一致<sup>[11,14,23-24]</sup>。聚类的 10 个类群中,第 V 类群生育期最短、株高最矮(代表性材料有中棉所 50、中棉所 58、新陆早 17 号、锦棉 5 号),第 VII 类群籽棉产量和皮棉产量最高(代表性材料有百棉 1 号、百棉 5 号、冀棉 958、锦棉 1 号、岱字棉 15)。根据主成分分

析结果,在棉花早熟、高产育种时,应从第 V、VII 类群和第二主成分得分值高的材料中选择优异亲本材料。此外,结合材料系谱图发现,同系列的陆地棉材料常选用同一个优良的品种作为亲本材料,如新陆早 42 号和新陆早 51 号均以新陆早 10 号作为母本,这可能是导致中国陆地棉遗传基础狭窄的一个重要原因<sup>[25]</sup>,因此需不断扩大不同地区不同系列品种间的交流和加强国外引种力度,提高中国棉花遗传多样性。值得说明的是,本研究的聚类分析是基于表型数据,由于表型性状容易受环境因素的影响,只有严格控制环境因素,才能得到准确的聚类结果<sup>[11]</sup>。结合表型数据与分子标记数据评价种质资源<sup>[26-29]</sup>,将更有助于探明种质间的亲缘关系,提高亲本选配的预见性。

### 参考文献:

- [1] 刘文欣,孔繁玲,郭志丽,等.建国以来我国棉花品种遗传基础的分子标记分析[J]. 遗传学报,2003,30(6):560-570.
- [2] CHEN G, DU X M. Genetic diversity of source germplasm of upland cotton in China as determined by SSR marker analysis[J]. Acta Genetica Sinica,2006,33(8):733-745.
- [3] 张小娟,何团结,陆徐忠,等.陆地棉 SSR 核心引物筛选及 95 份骨干种质的遗传多样性分析[J]. 棉花学报,2011,23(6):529-536.
- [4] 何亮.主成分分析在 SPSS 中的应用[J]. 山西农业大学学报,2007,6(5):20-22.
- [5] 娄丽娜,王 辉,王 成,等.腌渍萝卜种质资源的遗传多样性[J]. 江苏农业学报,2013,29(6):1421-1426.
- [6] 赵一洲,李正茂,刘福才,等.粳稻盐粳 188 航天诱变 SP<sub>2</sub> 代的性状变异与选择[J]. 江苏农业科学,2013,41(6):58-61.
- [7] 陈培峰,王建平,黄 健,等.太湖地区香稻品种稻米品质性状相关和聚类分析[J]. 江苏农业学报,2013,29(1):1-7.
- [8] 孙菲菲,王 夏,王 强,等.南京地区白菜种质资源遗传主成分与聚类分析[J]. 江苏农业科学,2014,42(3):106-109.
- [9] 康美玲,田忠景,张倩倩.利用醇溶蛋白电泳图谱分析不同玉米品种的遗传多样性[J]. 江苏农业科学,2013,41(10):70-72.
- [10] 阮燕晔,郭 瑞,崔震海,等.利用 SSR 技术分析辽宁省 32 个骨干玉米自交系的遗传多样性[J]. 江苏农业科学,2013,41(11):29-32.
- [11] 周丽艳,郭振清,马玉玲,等.春小麦品种农艺性状的主成分分析与聚类分析[J]. 麦类作物学报,2011,31(6):1057-1062.
- [12] 王绘艳,史雨刚,马昊翔,等.30 份春小麦品系主要农艺性状的聚类分析[J]. 中国农学通报,2014,30(18):60-64.
- [13] 李培富,杨淑琴,马宏伟.宁夏水稻主要农艺性状的主成分及聚类分析[J]. 中国农学通报,2006,22(12):162-166.
- [14] 李向华,常汝镇.中国春大豆品种聚类分析及主成分分析[J].

- 作物学报,1995,24(3):325-332.
- [15] 罗 珊,康玉凡,濮绍京,等.黑河地区 55 份大豆品种资源农艺性状和营养成分的聚类分析[J]. 大豆科学,2009,28(3):421-425.
- [16] 胡兴雨,陆 平,贺建波,等.黍稷农艺性状的主成分分析与聚类分析[J]. 植物遗传资源学报,2008,9(4):492-496.
- [17] BROWN J S. Principle component and cluster analysis of cotton cultivar variability across the US cotton belt[J]. Crop Science, 1991(31): 915-922.
- [18] 孙长发,赵 晖,陈荣江.棉花新品种产量品质性状的聚类分析与综合评价[J]. 西北农业学报,2010,19(4):77-81.
- [19] 许乃银,陈旭升,狄佳春,等.长江流域棉花纤维品质的区域特征研究[J]. 棉花学报,2003,15(4):221-226.
- [20] 陈荣江,张万琴,王文峰,等.棉花数量性状的因子分析与品种的聚类分析[J]. 福建农林大学学报:自然科学版,2009,38(3):225-230.
- [21] 张 鹏,张海洋,郑永战,等.芝麻种质资源因子分析及聚类分析[J]. 中国油料作物学报,2008,30(1):71-78.
- [22] PANTHEE D R, KC R B, REGMI H N, et al. Diversity analysis of garlic (*Allium sativum* L.) germplasms available in Nepal based on morphological characters[J]. Genetic Resources and Crop Evolution, 2006(53): 205-212.
- [23] 赵德新,孙治强,任子君,等.茄子形态学性状主成分分析及聚类分析[J]. 河南农业大学学报,2009,43(4):393-397.
- [24] 陈华萍,王照丽,魏育明,等.四川小麦地方品种农艺性状与品质性状的聚类分析[J]. 麦类作物学报,2006,26(6):29-34.
- [25] 郭志军,赵云雷,陈 伟,等.陆地棉 SSR 标记遗传多样性及其与农艺性状的关联分析[J]. 棉花学报,2014,26(5):420-430.
- [26] ZHANG Y X, ZHANG X R, CHE Z, et al. Genetic diversity assessment of sesame core collection in China by phenotype and molecular markers and extraction of a mini-core collection[J/OL]. BMC Genetics, 2012, 13: 102 [2012-11-15]. <http://www.biomedcentral.com/1471-2156/13/102>.
- [27] LI G S, ZHANG L J, BAI C K. Chinese Cornus officinalis: genetic resources, genetic diversity and core collection[J]. Genetic Resources and Crop Evolution, 2012(59): 1659-1671.
- [28] BAI C K, WEN M M, ZHANG L J, et al. Genetic diversity and sampling strategy of Scutellaria baicalensis germplasm resources based on ISSR[J]. Genetic Resources and Crop Evolution, 2013(60): 1673-1685.
- [29] 耿立格,王丽娜,张 磊,等.河北省绿子叶黑豆种质资源表现型和 ISSR 标记遗传多样性分析[J]. 植物遗传资源学报, 2010,11(3):266-270.

(责任编辑:张震林)