

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

罗珊¹,康玉凡¹,濮绍京²,张丽¹,李振华¹,肖伶俐¹,李永华¹

(¹中国农业大学芽菜研究中心,北京 100193;²北京农学院红小豆研究中心,北京 102206)

摘要:收集黑河地区 55 份大豆品种资源材料,以豆芽厂原料大豆 S135 为对照。采用 SPSS15.0 软件对大豆品种的 15 个农艺性状及营养成分进行相关分析、主成分分析和非系统聚类分析。结果表明:单株产量与生育期和单株荚数呈极显著正相关($P < 0.01$),与营养生长期和百粒重呈显著正相关($P < 0.05$),与底荚高呈极显著负相关($P < 0.01$);粗脂肪与粗蛋白呈极显著负相关($P < 0.01$),与粗纤维、营养生长期呈显著负相关($P < 0.05$)。在主成分分析中,选取方差累计贡献率为 84.49% 的前 5 个主成分来评价黑河地区大豆品种资源。进一步通过非系统聚类分析,将参试的 55 份品种资源分为 5 类,其中 5 个品种与对照品种 S135 为第一类群,从农艺性状和大豆的营养成分来考虑,第一类适合作芽用大豆品种;且 5 个类群性状之间的差异较明显,这有利于育种亲本材料的选择,同时为进一步选育和筛选优良的芽用大豆品种提供理论依据。

关键词:大豆;相关分析;主成分分析;聚类分析

中图分类号:S565.1

文献标识码:A

文章编号:1000-9841(2009)03-0421-05

Cluster Analysis of Agronomic Characters and Nutrition Composition of 55 Soybean Accessions from Heihe Area

LUO Shan¹, KANG Yu-fan¹, PU Shao-jing², ZHANG Li¹, LI Zhen-hua¹, XIAO Ling-li¹, LI Yong-hua¹

(¹ Bean Sprouts Research Center of China Agricultural University, Beijing 100193; ² Azukibean Research Center of Beijing University of Agriculture, Beijing 102206, China)

Abstract: To select soybean varieties suitable for sprouts production, 55 soybean accessions from Heihe area including control variety S135 were used, and 15 agronomic characters and nutrition composition were analyzed with the method of correlation analysis and principal component analysis (PCA) and K-means Cluster (KCA) of SPSS15.0. Correlation analysis indicated that single-plant yield correlated significantly positive with growth period and pods per plant ($P < 0.01$) and vegetative period and 100-seed weight ($P < 0.05$), single-plant yield correlated significantly negative with bottom-pod height ($P < 0.01$). Crude fat correlated significantly negative with crude protein ($P < 0.01$) and crude fibre and vegetative period ($P < 0.05$). PCA results showed that 5 Eigenvalues of cumulative variance proportion 84.49% were selected for evaluation of soybean varieties in Heihe. By the KCA, 55 soybean varieties was classified into 5 clusters, five varieties and S135 belonged to first clusters, which could be used to grow soybean sprouts. The selected 6 soybean varieties would be useful for breeding and selecting good soybean sprouts varieties.

Key words: Soybean varieties; Correlation analysis; Principal component analysis; K-means cluster analysis

以日本豆芽工厂化生产为先导,中国、韩国、美国以及西欧国家也逐渐开始推广工厂化豆芽生产,这促使国际豆芽产业的形成。目前,我国传统家庭作坊式的豆芽生产已被日产 60~150 t 的程序化、机械化、自动化、可调控的现代生产方式,即豆芽工厂化生产所取代。现代大豆芽产业的健康持续发展需要优质专用的芽用大豆品种、规模化的优质芽用大

豆原料基地建设与生产为重要支撑^[1]。

韩国和日本十分重视芽用大豆品种选育,一般芽用大豆品种要求百粒重 10 g 左右、发芽率高、抗病、豆芽生长快、芽长,同时要求脂肪含量低、蛋白质和碳水化合物含量高、味道佳、色泽黄而有光泽、耐贮藏。韩国已育成 Pureunkong、Iksannamulkong、Pungsannamulkong 等一批适于发豆芽的大豆高产新

收稿日期:2009-01-15

基金项目:国家农业行业专项基金资助项目(nyhyzx07-017)。

作者简介:罗珊,女(1980-),硕士研究生,研究方向为种子科学与工程。E-mail:luoshanlsh@163.com。

通讯作者:康玉凡,博士,教授。Email:yfkang@cau.edu.cn。

品种。美国也育成 IL1、IL2、Mercury、Minnatto 等豆芽和纳豆兼用的大豆专用品种^[2]。目前我国用于生产大豆芽的原料大豆,包括不同大小、籽粒颜色的品种;而小粒专用品种育种侧重于纳豆用,芽用大豆品种选育还有待加强。传统的作为芽用大豆加工品质则要求小粒大豆具有高蛋白质、低脂肪的特点^[3]。而现在我国工厂化生产豆芽要求的大豆百粒重一般在 14.0 ~ 18.0 g 左右,属于中小粒品种^[4]。

以我国有机绿色的大豆生产基地——黑龙江黑河地区的 55 个大豆品种为材料,运用相关分析、主成分分析和非系统聚类分析进行初步鉴定与评价黑河地区的大豆资源,旨在筛选出优良的芽用大豆核心种质资源材料,为我国工厂化大豆芽产业快速发展提供科学的理论和技术支持。

1 材料与方法

以黑河地区 55 份大豆品种为材料,编号为 S81 ~ S135,其中 S135 的百粒重为 17.21 g,为我国工厂化生产豆芽专用大豆品种。55 个大豆品种的籽粒颜色全为黄色,其种脐有黄色和黑色两种,百粒重范围在 9.33 ~ 27.06 g。试验材料在黑河地区红色边疆农场种植。随机区组设计,2 次重复,行长 5 m,单行区,株行距 0.1 × 0.5 m。于 2008 年 5 月 20 日播种,采用机器开沟,人工双粒点播,出苗后 10 d 定苗。生育期间进行物候期调查,收获前每小区随机取 10 株进行室内考种。考种性状包括:生育期、

株高、底荚高、主茎节数、有效分枝数、单株荚数等 11 个性状。收获后每个品系混合取样,进行品质分析。蛋白质、脂肪含量分析分别采用国家标准 GB2905-1982(谷类、豆类作物种子粗蛋白质测定方法)和国家标准 GB2906-1982(谷类、油料作物种子粗脂肪测定方法)^[5],粗纤维的测定采用酸碱洗涤法^[6]。分析数据使用软件 SPSS15.0 和 Excel2003,在进行大豆品种性状间的相关分析,采用两尾测验,用($P < 0.01$)和($P < 0.05$)两种水平进行;在主成分分析中,选择方差累计贡献率 $\geq 84.49\%$ 的前 5 个主成分作为从不同方面来表征总体遗传方差的主要因素。在进行非系统聚类分析时,将原始数据进行标准化处理,采用 K-means Cluster 方法进行聚类分析。

2 结果与分析

2.1 农艺性状和营养成分的品种间差异

各性状不同品种间的变异情况见表 1。15 个性状的方差分析结果表明,除生殖生长期、每荚粒数、分枝数和粗纤维外,其它性状品种间的差异均达到极显著。从变异系数上可以看出,各性状在遗传上的变异大小情况,分枝数、底荚高、单株产量、单株荚数和百粒重的变异系数较大,而粗蛋白/粗脂肪比值、主茎节数和粗纤维等的变异系数次之,生育期的变异系数最小。分枝数变异系数较大的原因是,参试的大多数品种没有分枝,个别品种有分枝,因此其平均值较小,标准差较大。

表 1 各性状品种间的变异情况表
Table 1 Statistics of the agronomic traits

性状 Character	变异幅度 Range	均值 Mean	标准差 SD	变异系数 CV/%	F 值 F
生育期 Growth period/d	98.00 - 120.0	110.07	4.055	3.68	5.33**
营养生长期 Vegetative period/d	61.00 - 80.00	71.07	3.399	4.78	8.57**
生殖生长期 Reproductive period/d	37.00 - 40.00	39.00	2.373	6.08	0.21
株高 Plant height/cm	84.60 - 90.80	95.41	10.081	10.57	5.85**
主茎节数 Node number of main stem	13.70 - 30.60	16.63	2.243	13.49	37.48**
底荚高度 Height of lowest pod/cm	10.75 - 38.50	23.78	6.618	27.83	11.10**
单株荚数 Pods per plant	14.60 - 34.30	23.65	4.691	19.83	9.10**
每荚粒数 Seed number per plant	2.10 - 4.12	2.63	0.261	9.93	1.93
单株产量 Seed yield per plant/g	5.26 - 17.08	10.10	2.405	23.81	8.53**
分枝数 Branches number	0 - 0.800	0.06	0.139	243.74	1.01
百粒重 100-seed weight/%	9.33 - 27.06	18.68	2.850	15.26	4.94**
粗蛋白 Crude protein/%	33.39 - 44.75	37.48	2.587	6.90	9.29**
粗脂肪 Crude fat/%	17.18 - 26.53	22.55	1.954	8.66	12.12**
粗纤维 Crude protein/crude/%	2.78 - 5.53	3.87	0.505	13.06	1.73
粗蛋白/粗脂肪 Rate of crude protein and crude fat	1.28 - 2.57	1.68	0.246	14.66	16.24**

2.2 大豆品种性状间的相关分析

从表 2 可以看出,单株产量与生育期和单株荚数呈极显著正相关,与营养生长期和百粒重呈显著正相关,与底荚高呈极显著负相关,这与韩秉进和张彩英等的研究结果一致^[7-8]。底荚高度与株高呈极显著正相关,与单株荚数和单株产量呈极显著负相关,因此要获得较高的产量,首先要提高单株荚数,增加百粒重,适当降低结荚部位。单株荚数与生育期和营养生长期呈显著正相关,与底荚高呈极显著

负相关,与百粒重呈显著负相关。百粒重与单株产量呈显著正相关,与单株荚数呈显著负相关。粗脂肪与粗蛋白、粗蛋白/粗脂肪比值呈极显著负相关,与粗纤维、营养生长期呈显著负相关。因此根据育种目标,可以针对性地选择亲本。比如要选育优良的芽用大豆品种,可以选择粗蛋白含量高、产量高和其它农艺性状优良的品种作母本,选择早熟和粗脂肪含量相对低的品种作父本。

表 2 大豆品种性状间的相关矩阵
Table 2 Correlations matrix of soybean traits

	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	L15
L1	1														
L2	0.811**	1													
L3	0.547**	-0.046	1												
L4	-0.199	-0.217	-0.030	1											
L5	0.197	0.161	0.106	0.144	1										
L6	0.054	0.073	-0.012	0.440**	0.138	1									
L7	0.282*	0.282*	0.077	-0.080	0.194	-0.428**	1								
L8	0.150	0.181	-0.003	0.100	0.023	0.199	-0.038	1							
L9	0.388**	0.323*	0.201	-0.230	0.082	-0.345**	0.630**	0.014	1						
L10	0.031	0.023	0.021	-0.191	-0.074	-0.035	-0.280*	-0.153	0.279*	1					
L11	0.023	0.139	-0.159	-0.067	0.088	-0.013	0.040	-0.044	-0.071	-0.067	1				
L12	-0.234	-0.339*	0.085	0.139	-0.013	-0.122	-0.053	-0.139	0.039	0.246	-0.560**	1			
L13	0.081	0.097	-0.001	-0.036	0.097	0.038	0.007	-0.102	-0.086	-0.189	0.214	-0.311*	1		
L14	0.155	0.262	-0.109	-0.127	0.027	0.048	0.050	0.081	-0.093	-0.235	0.843**	-0.903**	0.307*	1	
L15	0.032	-0.003	0.059	0.214	0.032	0.049	0.097	0.121	0.146	-0.155	0.082	0.045	0.004	0.018	1

*代表在 0.05 的水平上显著; **代表在 0.01 的水平上显著; L1: 生育期; L2: 营养生长期; L3: 生殖生长期; L4: 株高; L5: 主茎节数; L6: 底荚高度; L7: 单株荚数; L8: 每荚粒数; L9: 单株产量; L10: 百粒重; L11 粗蛋白; L12 粗脂肪; L13 粗纤维; L14 粗蛋白/粗脂肪; L15: 分枝数。

* means significantly at the level of 0.05; ** means significantly at the level of 0.01. L1: Growth period, L2: Vegetative period, L3: Reproductive period, L4: Plant height; L5: Node number of main stem, L6: Height of lowest pod, L7: Pods per plant, L8: Seed number per plant, L9: Sing-plant yield, L10: 100-seed weight, L11: Crude protein, L12: Crude fat, L13: Crude fiber, L14: Rate of crude protein and crude fat, L15: Branches per plant.

2.3 大豆品种性状的主成分分析

从方差分析的结果来看,品种之间的分枝数、生殖生长期、每荚粒数和粗纤维的差异未达到显著水平,在进行主分析时可以将这四个变量剔除,用剩余的 11 个指标进行主成分分析及聚类分析。主成分分析结果如表 3 所示。

负荷量是指变量与主成份的相关系数,它的平方表示该主成份解释该变量的方差比例。其绝对值越大,表明该变量与该主成份的相关性越高^[9]。由表 3 可知,在第一主成分中,粗蛋白、粗蛋白/粗脂肪的比值、营养生长期和生育期的负荷量较大,单株荚数和单株产量的负荷量次之。第一主成分的方差贡献率为 26.99%,表明第一主成分大的品种,具有粗蛋白含量高、单株荚数多、单株产量高、百粒重小、营养生长期和生育期长等特点。第一主成分值大的品

种适合作芽用大豆。芽用大豆品种的特点一般要求是百粒重为中粒或中小粒,粗蛋白含量高。在第二主成分中,单株产量、单株荚数、百粒重,粗脂肪的负荷量较大,生育期和营养生长期的负荷量次之,其方差贡献率为 22.61%,表明第二主成分大的品种,具有单株产量高,单株荚数多,百粒重大,粗脂肪含量高,生育期相对较长的特点,第二主成分值大的品种适合作高油大豆。在第三主成分中,底荚高的负荷量最大,其次是株高、主茎节数、生育期和营养生长期,其方差贡献率为 14.19%,这表明第三主成分大的品种,具有植株高大,底荚结荚部位高且生育期长的特点。在第四主成分中,单株荚数的负荷量最大,其次是株高和主茎节数,方差贡献率为 12.96%,这表明第四主成分大的品种具有单株荚数多,植株高大,生育期短且早熟的特点。在第五主成分中,百粒

重的负荷量最大,其次是主茎节数、单株产量、粗蛋白和株高,其方差贡献率为 7.73%,这表明第五主

成分大的品种具有百粒重大、粗蛋白含量高、植株高大和生育期短且极早熟等特点。

表3 主成分分析的各性状的特征根和特征向量

Table 3 PCA eigenvalues and eigenvector of the agronomic traits

主成份 Principal component	第一主成份 First principal component	第二主成份 Second principal component	第三主成份 Third principal component	第四主成份 Fourth principal component	第五主成份 Fifth principal component
特征根 Eigenvalue	2.97	2.49	1.56	1.43	0.85
贡献率 Cumulative/%	26.99	22.61	14.19	12.96	7.73
累计贡献率 Cumulative variance proportion/%	26.99	49.61	63.80	76.76	84.49
性状 Characters	特征向量 Eigenvector		特征向量 Eigenvector		
	1	2	3	4	5
生育期 Growth period	0.66	0.37	0.47	-0.15	-0.26
营养生长期 Vegetative period	0.73	0.26	0.42	-0.16	-0.29
株高 Plant height	-0.36	-0.33	0.51	0.32	0.26
主茎节数 Node number of main stem	0.19	0.01	0.51	0.43	0.38
底荚高度 Height of lowest pod	-0.13	-0.45	0.71	-0.29	0.08
单株荚数 Pods per plant	0.43	0.46	-0.09	0.70	-0.03
单株产量 Sing-plant yield	0.42	0.75	-0.06	0.08	0.37
百粒重 100-seed weight	0.10	0.50	-0.04	-0.70	0.47
粗蛋白 Crude protein	0.60	-0.55	-0.25	0.04	0.33
粗脂肪 Crude fat	-0.73	0.52	0.08	0.09	0.04
粗蛋白/粗脂肪 Rate of crude protein and crude fat	0.75	-0.62	-0.20	-0.02	0.09

2.3 大豆品种性状聚类分析

为了进一步评价黑河地区大豆品种资源,对不同性状品种间达到显著差异的 11 个性状进行非系统聚类分析。聚类结果如表 4 所示。第 I 类的总体特征为生育期相对较长且中晚熟,主茎节数多,单株荚数多,百粒重较小,脂肪含量低,另外营养生长期长,有利于营养成分的积累。对于黑龙江省黑河地区第五积温带来说,营养生长期长还能缓解该区的前期干旱的情况^[10]。第一类包括 6 份品种资源,其中有一份为对照品种 S135,从农艺性状和大豆的营养成分考虑,这 6 份品种适宜作芽用大豆。第 II 类包括 8 份品种资源,总体特征为生育期和营养生长期短且早熟,粗脂肪含量高,百粒重、底荚结荚位、粗蛋白/粗脂肪比值和粗蛋白含量偏低,这一类品种属于高油大豆。第 III 类只有 1 份品种资源,总体特征为生育期和营养生长期长且极晚熟,百粒重较大,单株产量、粗蛋白/粗脂肪比值和粗蛋白含量相对较高;植株矮小,结荚部位低,主茎节数和单株荚数偏低。第 IV 类包括 19 份品种资源,总体特征为生育期较短且早熟,植株高大,底荚结荚部位较高,单株荚数相对较多。第 V 类包括 21 份品种资源,总体特征为生育期短且极早熟,植株和底荚结荚部位较高,单株荚数、单株产量和粗蛋白含量偏低。

表4 参试品种资源各类群的特征

Table 4 The characters of every cluster

性状 Characters	I	II	III	IV	V
生育期 Growth period	114.00	104.88	120.00	110.37	110.19
营养生长期 Vegetative growth period	74.00	67.25	80.00	70.84	71.48
株高 Plant height	83.02	89.44	51.90	100.81	98.40
主茎节数 Node number of main stem	17.40	15.20	14.40	16.70	17.00
底荚高度 Height of lowest pod	18.83	15.98	10.75	22.27	30.14
单株荚数 Pods per plant	30.23	22.92	21.80	25.42	20.54
单株产量 Sing-plant yield	11.55	8.98	17.08	11.08	8.91
百粒重 100-seed weight	15.48	15.63	29.19	17.04	16.80
粗蛋白 Crude protein	37.75	36.33	38.67	38.13	37.19
粗脂肪 Crude fat	21.26	23.46	21.49	22.71	22.48
粗蛋白/粗脂肪 Rate of crude protein and crude fat	1.79	1.58	1.80	1.69	1.67

3 讨论

通过对黑河地区 55 份大豆品种资源 15 个农艺性状及营养成分的主成份分析和聚类分析,结果表明,在主成份分析中,由于所选的遗传群体和研究的性状不同,主成份构成的综合性状也有差异。分析结果表明,所选前 5 个主成分的方差累计贡献率为 84.49%,说明这 5 个主成分能较好地代替所考查的各个性状的综合特征,并利用所选的主成分对品种资源进行评价和选择。这与赵银月等的研究结果一致^[11]。第一主成份值大的品种,具有粗蛋白含量高、单株荚数多、单株产量高、百粒重小等特点;第二主成份值大的品种,具有单株产量高,单株荚数多,百粒重大,粗脂肪含量高,生育期相对较长的特点;第三主成份值大的品种,具有植株高大,底荚结荚部位高且生育期长的特点;第四主成份值大的品种,具有单株荚数多,植株高大且早熟的特点;第五主成份值大的品种,具有百粒重大、粗蛋白含量高、植株高大、生育期短且极早熟等特点。因此,在杂交育种选配亲本时,可以选择主成份互补的材料。比如,要选育农艺性状好品质优良且早熟的芽用大豆品种,可以选择第一主成份值大的品种作母本,第四主成份值大的品种作父本,进行杂交育种。

从聚类分析结果来看,参试的 55 个大豆品种被分成了 5 类,总体情况是相同的群体分到了一类,但也有少数相同地区的品种聚不到同一类群中的现象,这表明品种类型间的遗传变异与地理分布是有一定关系的。由于各类群间的性状遗传差异较大,故选配亲本时应在类间根据性状间的差异进行选择。

东北地区是我国大豆的主产区,而又以黑河地区的大豆品质为佳。芽用大豆品种黑河 17^[12]、黑河 20^[13]等在黑河地区曾经有大面积推广,并出口韩国和日本等国家。但由于产量低,这些品种种植面积越来越小。最近几年,随着我国豆芽产业的形成以及对优质芽用大豆品种的需求日益增加,这就迫切需要选育一批优良的芽用大豆品种做支撑。

参考文献

- [1] 肖伶俐,康玉凡,陶礼明,等. 不同大豆品种芽用特性比较[J]. 大豆科学,2008,27(6):23-26. (Xiao L L, Kang Y F, Tao L M, et al. Comparison analysis on characteristics of soybean growing sprouts for different soybean varieties[J]. Soybean Science,2008,27(6):23-26.)
- [2] 花登峰,赵团结,张黎萍,等. 小粒专用大豆品种遗传改良研究进展[J]. 杂粮作物,2005,25(5):311-313. (Hua D F, Zhao T J, Zhang L P, et al. Study advances of genetic improvement in small seed soybean[J]. Rain Fed Crops,2005,25(5):311-313.)
- [3] 郑伟. 小粒大豆新品种筛选试验[J]. 中国种业,2007,11:58-59. (Zheng W. Filtration experimentation of small seed soybean new varieties[J]. Journal of China Seed,2007,11:58-59.)
- [4] 王连铮,郭庆元. 现代中国大豆[M]. 北京:金盾出版社,2007:8. (Wang L Z, Guo Q Y. Soybean of modern time in China[M]. Beijing:Golden Shield Press,2007:8.)
- [5] 李为喜,刘三才,邱丽娟,等. 新收集大豆种质资源主要品质鉴定与评价[J]. 植物遗传资源学报,2005,6(1):75-79. (Li W X, Liu S C, Qiu L J, et al. Identification and evaluation of quality traits in newly collected soybean germplasm[J]. Journal of Plant Genetic Resources,2005,6(1):75-79.)
- [6] 张丽英. 饲料分析及饲料质量检测技术[M]. 北京:中国农业大学出版社,2003(Zhang L Y. Feed analysis and feed quality examination technology[M]. Beijing:China Agricultural University Press,2003)
- [7] 韩秉进,潘相文,金剑. 大豆农艺及产量性状的主成分分析[J]. 大豆科学,2008,27(1):67-73. (Han B J, Pan X W, Jin J. Principal component analysis of agronomic and yield-related traits in soybean[J]. Soybean Science,2008,27(1):67-73.)
- [8] 张彩英,段会军,常文锁,等. 大豆遗传资源农艺性状的鉴定和筛选[J]. 中国农学通报,2004,20(5):115-121. (Zhang C Y, Duan H J, Chang W S, et al. Identification and screening of agronomic traits in soybean genetic resources[J]. Chinese Agriculture Science Bulletin,2004,20(5):115-121.)
- [9] George, Darren. Mallery, Paul. SPSS for Windows step-by-step = : SPSS 13.0[M]. Beijing:World Book Publishing Company,2006.
- [10] 李殿祥. 大豆主要农艺性状灰色关联度分析[J]. 农业与技术,2006,26(4):53-55. (Li D X. Gray correlation analysis on main agriculture traits of soybean[J]. Agriculture & Technology,2006,26(4):53-55.)
- [11] 赵银月,耿智德,保丽萍. 云南省大豆地方品种资源的主成分分析及聚类分析[J]. 湖南农业大学学报(自然科学版),2007,33. (Zhao Y Y, Geng Z D, Bao L P. Principal component analysis and cluster analysis of local soybean varieties of Yunnan[J]. Journal of Hunan Agricultural University (Natural Sciences),2007,33(8):120-122.)
- [12] 张雷闰,洪睿鹿,文成梁,等. 黑河 17 号的特征特性及高产栽培技术[J]. 耕作与栽培,2005,2:61. (Zhang L Y, Hong R L, Wen C L, et al. Cultivating technique and characters of soybean variety 17[J]. Farming and Cultivating,2005,2:61.)
- [13] 马启慧. 2000 年黑龙江省审定推广的大豆品种[J]. 大豆科学,2000,19(2):193-196. (Ma Q H. Soybean varieties were approved in Heilongjiang province in 2000. [J]. Soybean Science,2000,19(2):193-196.)