

大豆杂交亲本主要数量性状遗传距离 的测定及其聚类分析*

李爱萍 赵志安 刘雨明

(四平地区农业科学研究所)

合理地选配亲本是大豆杂交育种工作的关键环节。双亲遗传差异的大小是选择亲本及配制组合的重要依据之一。以往的大豆杂交育种工作多根据表型或地理上的差异来选配组合。由于育种中选择的性状多数是受微效多基因控制的数量性状,这些性状受环境影响较大,并且彼此间有一定相关性,这就增加了按遗传差异选配组合的困难。同时,遗传学研究指出,地理差异与遗传差异之间并无直接联系^(1,2)。因此,依照以往的原则所进行的亲本选型的实际效果不够理想。在配制组合中存在相当程度的盲目性。由于组合过多,各单位受人力、地力所限,每一组合入选材料数目相对减少,从而影响筛选和鉴定效果。

为了寻求科学合理的选配亲本方法,近年来人们应用“多元分析”的方法,以“遗传距离”为尺度定量地测定亲本的遗传差异,即用“遗传距离”作为育种中亲本选配的一个参数。在我国已被应用在小麦、水稻等作物指导亲本选配及预测杂种优势^(3,4,5,7),在大豆方面尚无报道。本研究的目的在于试图利用这一较新的方法在前人研究基础上测定大豆杂交亲本主要数量性状的遗传距离,进而对亲本进行聚类分析,以便指导选配组合,减少亲本选型的盲目性,提高大豆杂交育种效果。

一、材料和方法

(一) 试验材料、设计及考察性状

本试验结合引种鉴定进行,以产量性状为主,收集品种(品系)共28个(见表7)。

田间试验采用随机区组法,三次重复,双行区,3米行长,行距65厘米,株距10厘米。收获时从小区中部随机取样10株考种。

考察7个性状:株高,分枝数,主茎节数,每株荚数,每株粒数,单株粒重,百粒重。

表1 方差、协方差分析模式

变异来源	自 由 度	方差	理 论 方 差 组 成	协 方 差	理 论 协 方 差 组 成
总	$Tn-1$				
重 复	$T-1$				
品 种	$n-1$	V_1	$6\epsilon^2 + 16g^2$	CoV_1	$6e_{ij} \cdot 16g_{ij}$
机 误	$(T-1)(n-1)$	V_2	$6\epsilon^2$	CoV_2	$6e_{ij}$
遗传方差	6^2g		$(V_1 - V_2) / T$	遗传协方差 $6g_{ij}$	$(CoV_1 - CoV_2) / T$

V_1 —品种方差;

CoV_2 —机误协方差;

V_2 —机误方差;

$6e_{ij} - i \cdot j$ 两性状机误协方差;

CoV_1 —品种协方差;

$6g_{ij} - i \cdot j$ 两性状基因型协方差。

* 本文试验数据由省农科院大豆所计算机组陈光、宋艳丽,作物所靳树凯,省计算中心李春贵、方臣等同志协助计算,谨表谢意。

(二) 统计分析方法及运算程序

1、对28个品种(品系)7个数量性状进行方差分析,淘汰F值不显著性状。然后计算选留性状中每两个性状的协方差,按表1模式进行方差、协方差分析。

2、计算每两性状遗传相关系数。

计算公式:

$$\gamma_{gij} = \frac{\sigma_{gij}}{\sqrt{\sigma^2_{gi} \cdot \sigma^2_{gj}}}$$

其中 γ_{gij} —i·j两性状遗传相关系数; σ_{gij} —i·j两性状遗传型协方差;
 σ^2_{gi} —i性状遗传型方差; σ^2_{gj} —j性状遗传型方差。

3、列出遗传相关矩阵。

4、求上述遗传相关矩阵的特征根及特征向量。

5、求出每品种主成份值,进行主成份分析。

6、计算遗传距离(D^2_{ij})

7、对参加试验的28个品种(品系)进行聚类分析。

二、试验结果分析

1、28个大豆品种(品系)7个性状的方差分析结果见表2。由表2看出,7个性状的方差分析结果F测验均极显著,故7个性状全部入选。然后进行每两性状的协方差分析,结果见表3。

表2 七个性状方差分析结果

方差来源	自由度	株高	分枝	主茎节数	单株荚数	单株粒数	单株粒重	百粒重
总	83	174.60	2.19	5.48	65.13	299.21	11.96	6.46
区组	2	24.03	2.51	4.03	169.84	702.76	44.30	0.43
品种	27	471.88	5.75	13.83	131.02	554.49	22.81	18.52
机误	54	31.53	0.41	1.36	28.31	156.63	5.34	0.65
F值	品种	**	**	**	**	**	**	**
	机误	14.97	14.02	10.17	4.63	3.54	4.27	28.49

表3 各项协方差

协方差性状	CoV1			CoV2			CoV3			CoV4		
	CoV1	CoV2	$\sigma_{gi \cdot j}$	CoV1	CoV2	$\sigma_{gi \cdot j}$	CoV1	CoV2	$\sigma_{gi \cdot j}$	CoV1	CoV2	$\sigma_{gi \cdot j}$
1·2	6.33	0.072	2.086	2·4	14.9719	0.5835	4.7961	3·7	-4.27	0.22	-1.50	
1·3	71.54	3.69	22.62	2·5	4.67	0.95	1.24	4·5	200.89	47.96	50.98	
1·4	41.95	-0.14	14.03	2·6	6.58	0.25	2.11	4·6	44.24	9.45	11.59	
1·5	47.28	0.84	15.48	2·7	5.94	0.1005	1.95	4·7	-2.02	-0.35	-0.79	
1·6	22.59	1.52	7.02	3·4	8.1974	0.8222	2.45	5·6	59.22	21.04	12.73	
1·7	-3.6148	0.0067	-1.2072	3·5	20.73	4.85	5.29	5·7	-42.17	-0.22	-14.13	
2·3	0.31	-0.0322	0.11	3·6	1.67	0.5987	0.3571	6·7	10.52	0.47	3.35	

*性状序号按表3各性状次序

2、依前公式求出每个性状间的遗传相关系数(见表4)。

表4 七个性状遗传型相关系数

性状	株高	分枝数	主茎节	单株荚数	单株粒数	单株粒重	百粒重
分枝	0.1291						
主茎节	0.9156	0.0417					
单株荚数	0.1979	0.6142	0.2060				
单株粒数	0.0792	0.0805	0.2585	0.7566			
单株粒重	0.2402	0.6552	0.0726	0.8211	0.4579		
百粒重	-0.0408	0.5978	-0.3005	0.1358	-0.4976	0.5693	

3、由7个性状的相关系数可以列出如下7阶遗传相关矩阵(R)。

$$R = \begin{pmatrix} \text{株高} & \text{分枝} & \text{主茎节} & \text{单株荚数} & \text{单株粒数} & \text{单株粒重} & \text{百粒重} \\ 1.0000 & 0.1291 & 0.9156 & 0.1979 & 0.0792 & 0.2402 & -0.0408 \\ 0.1291 & 1.0000 & 0.0417 & 0.6142 & 0.0805 & 0.6552 & 0.5978 \\ 0.9156 & 0.0417 & 1.0000 & 0.2060 & 0.2585 & 0.0726 & -0.3005 \\ 0.1979 & 0.6142 & 0.2060 & 1.0000 & 0.7566 & 0.8211 & 0.1358 \\ 0.0792 & 0.0805 & 0.2585 & 0.7566 & 1.0000 & 0.4579 & -0.4976 \\ 0.2402 & 0.6552 & 0.0726 & 0.8211 & 0.4579 & 1.0000 & 0.5693 \\ -0.0408 & 0.5978 & -0.3005 & 0.1358 & -0.4976 & 0.5693 & 1.0000 \end{pmatrix}$$

4、采用“Jacobi”法将上述实对称矩阵进行正交相似变换($M^{-1}BM$)，使之转换成相应的对角矩阵，主对角矩阵中的主对角线元素即为上述遗传相关矩阵(R)的特征根，对应于每一特征根的正交矩阵M的列向量就是这个特征根的特征向量。求得矩阵R的特征根的大小代表各综合指标遗传方差的大小。特征向量则表示在各综合指标中，供试品种的诸性状对综合指标贡献的大小。然后从7个特征根($\lambda_1 - \lambda_7$)中选取3个较大的特征根及对应的特征向量(特征根入选的标准是，入选的特征根的累计贡献在85%)(表5)。

表5 入选的特征根和特征向量

	λ_1	λ_2	λ_3	分量来源
特征根	2.9990	2.1207	1.4623	
累计贡献率%	42.8	73.1	94.0	
特征向量	0.2492	-0.4166	-0.5368	株高
	0.4294	0.2991	-0.1506	分枝
	0.2100	-0.5433	-0.3856	主茎节
	0.5328	-0.0228	0.2786	单株荚数
	0.3280	-0.3189	0.5651	单株粒数
	0.5284	0.1843	0.0374	单株粒重
	0.1968	0.5529	-0.3770	百粒重
	i_1	i_2	i_3	
主成分名称	荚数因子(第一主成分) / 百粒重因子(第二主成分) / 粒数因子(第三主成分)			

5、主成份分析结果

由表5可见，第一主成份的特征向量以单株荚数的正值最大，其他分量也皆为正值，单株粒重及分枝两性状的正值仅次于单株荚数，说明单株荚数多的品种产量较高。这样的品种应是植株稍高，分枝适中，主茎节多。我们将第一主成份定名为荚数因子。第二主成份向量中以百粒重的正值最大，分枝、单株粒重为正值，株高、主茎节、单株粒数、单株荚数为负值，说明大粒品种多矮秆、主茎节偏少、单株粒数较少。单株荚数虽为负值，但仅为-0.0228，这说明随着粒子增大，荚数减少的并不大。单株粒重分量为0.1843，说明粒大的品种产量并不低。我们称第二主成份为百粒重因子。第三主成份特征向量以单株粒数正值为最大称之为粒数因子。但在分量中，百粒重为负值，随着粒数的增加，百粒重有降低趋势。

将入选的三个较大的特征根和对应的特征向量以及28个品种、7个性状标准化的基因型值代入求主成份值公式：

$$\tilde{g}_{ij} = \frac{1}{\sqrt{\lambda_j}} \sum_{k=1}^n l_{kj} g_{ik} \quad (n=7, j=1, 2, 3)$$

表6 28个品种的三个主成份值

品种编号	\tilde{g}_1	\tilde{g}_2	\tilde{g}_3	品种编号	\tilde{g}_1	\tilde{g}_2	\tilde{g}_3
1	7.86	-3.66	-1.83	15	8.01	-2.16	-2.13
2*	10.41	-4.09	-1.07	16*	8.83	-3.99	-1.50
3*	8.23	-2.28	-3.36	17*	9.84	-4.83	-2.41
4	8.06	-3.01	-2.35	18	8.63	-4.42	-1.33
5	8.04	-3.18	-0.28	19	8.94	-4.21	-3.08
6*	9.43	-0.32	-3.38	20*	9.57	-4.39	-2.09
7*	10.09	-3.58	-3.67	21	8.58	-3.99	-3.12
8*	9.83	-4.67	-4.59	22*	12.80	-2.06	-1.28
9*	9.26	-3.55	-1.94	23*	9.79	-4.70	-2.20
10	8.70	-3.87	-2.82	24	8.74	-4.49	-1.68
11	8.42	-2.53	-2.21	25*	9.42	-3.12	-0.98
12*	9.03	-4.15	-1.05	26	8.49	-3.38	-2.71
13	8.54	-3.33	-1.20	27	8.29	-3.29	-2.20
14	9.29	-4.18	-2.84	28*	9.66	-3.81	-1.21

求出28个品种(品系)的3个主成份值 \tilde{g}_1 、 \tilde{g}_2 、 \tilde{g}_3 (见表6)。

根据上述主成份分析的结果说明，在本试验的条件下，产量较高的品种应该是单株荚数较多、中大粒、中秆、分枝适中、单株粒数较多。依据主成份值的大小来筛选亲本的标准应是第一主成份值越大越好；第二主成份值大小的选择应依选育目标而定。结合四平地区的具体条件来看，靠近东部地区选育目标为有限、亚有限类型，中大粒，故要求第二主成份值应大些；而中西部地区育种目标为亚有限、无限类型，粒中等太小，则要求第二主成份值可稍小些。但是从第二主成份分析结果来看，百粒重增加，单株产量还是随之增加的，因此无论选育哪种类型，百粒重都不能太小，所以第二主成份值的选择标准应大于第

二主成份值的平均值。第三主成份值的选择应在第一、第二主成份值选定的基础上适当考虑，因为粒数增加百粒重有降低趋势。按此原则，我们从28个品种（品系）中筛选出14份材料，作为配制杂交组合首先考虑的亲本。这些入选的材料大多是单株荚数、粒数较多的中大粒品种（品系）。荚数、粒数、百粒重的平均值分别为50.74、118.50、17.83（表7）。

表7 供试材料主要数量性状考种结果

编 号	性 状 品 种	株 高 (cm)	分 枝 (个)	主茎节 (个)	单株荚数 (个)	单株粒数 (个)	单株粒重 (克)	百粒重 (克)
1	九农9	70.16	1.36	18.26	40.86	107.73	14.40	14.06
2*	长7413-1	79.90	1.33	20.73	56.10	142.40	22.83	16.96
3*	开育8	74.30	1.23	17.96	41.86	88.56	16.90	20.60
4	长7417-1	73.20	1.43	17.63	41.86	102.30	16.43	17.70
5	吉林13	58.60	1.40	15.53	45.53	121.93	14.93	13.66
6*	公7514-2	64.40	5.30	15.86	44.86	90.33	19.90	24.70
7*	Century	95.90	3.40	21.80	50.13	108.86	20.33	18.96
8*	开育3	114.80	1.63	23.63	44.76	108.76	20.50	18.46
9*	公7622-4	79.40	1.30	19.23	48.23	120.23	19.50	17.20
10	四7834-1	90.90	0.26	19.40	43.53	118.66	17.80	17.63
11	九7617-23-4-1-1	70.10	0.56	17.06	41.23	106.46	19.13	19.20
12*	吉林16	95.60	3.03	22.60	45.26	98.90	15.30	16.86
13	九农5	69.70	0.73	17.16	46.36	117.93	17.73	15.66
14	A100	92.70	1.53	20.86	48.36	113.03	18.36	16.33
15	合丰23	62.60	0.86	16.40	38.86	101.93	17.70	19.30
16*	公7703-4	79.90	1.13	18.56	44.80	125.60	17.96	14.86
17*	四7833-2	93.90	1.36	22.60	51.30	127.63	19.76	15.63
18	吉林17	73.60	0.96	20.20	45.06	126.10	16.63	13.60
19	四7831-2	93.40	2.33	20.76	42.00	113.33	16.83	15.76
20*	公7604-2	88.60	1.63	20.90	53.43	122.66	18.13	15.50
21	九7312-12-1-1-1	92.30	0.66	19.90	40.50	108.13	17.73	16.76
22*	铁丰18	84.90	6.30	18.16	70.30	144.10	28.80	21.60
23*	四7832-2	90.70	1.96	22.10	50.00	130.53	19.16	15.33
24	吉林18	78.60	1.13	20.56	46.10	123.70	16.33	14.00
25*	九农2	69.20	1.56	17.56	52.56	123.93	20.30	16.60
26	F70-707	84.20	1.36	18.26	42.26	104.46	17.23	16.76
27	九7602-26-3-1-1	76.90	0.40	17.80	41.53	107.90	17.86	16.83
28*	九农6	79.10	0.83	18.96	56.73	126.56	19.96	16.33
平 均		81.35	1.65	19.30	46.93	115.21	18.52	17.03

*为筛选出来的优良材料。

6、主要数量性状的遗传距离及聚类分析

以各品种的三个主成份值组成一个主成份向量，每两个主成份向量间的几何距离即为这两个品种间的遗传距离，计算公式为：

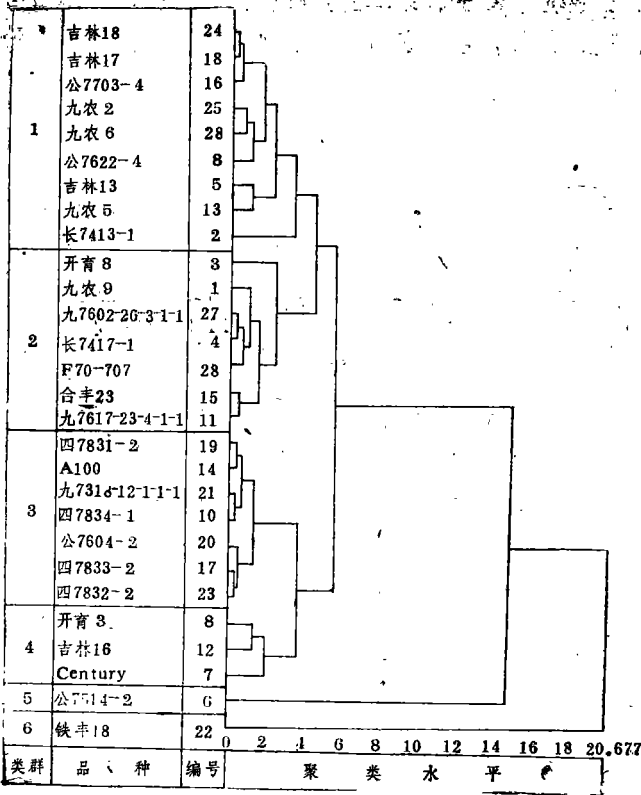
$$D^2_{ij} = \sum_{k=1}^{n'} (g'_{ik} - g'_{jk})^2$$

个28品种每两个品种间的遗传距离，共 $C_{28}^2 = 378$ 个。用类平均距离法聚类，共聚成6类（见图）。各类包括的品种及类内、类间距离列于表8。

从表8可见，同一类群内品种间的遗传距离较小，不同类群间的品种的遗传距离较大。选择亲本一般不应在类内选，应在不同类群间选择。选择的标准应该是遗传距离大于、等于总平均的遗传距离（本试验的总平均遗传距离计算得6.09）。依据上述原则，该试验满足这一条件的类群有I和IV、I和V、I和VI、II和IV、II和V、II和VI、III和V、III和VI、IV和V、IV和VI、V和VI。

三、讨 论

通过主成份分析从供试品种（品系）中选出的14个杂交亲本多



28个大豆品种的聚类图

表8 各类群间、类内的遗传距离及各类中的品种

类 群	I	II	III	IV	V	VI	品种数	品 种 编 号
I	2.01						9	2, 5, 13, 16, 18, 24, 25, 28
II	4.24	1.29					7	1, 3, 4, 11, 15, 26, 27
III	3.72	4.12	1.07				7	10, 14, 17, 19, 20, 21, 23
IV	11.82	7.63	3.34	1.63			3	7, 8, 12
V	17.14	9.71	31.29	15.65	0		1	6
VI	17.98	23.90	20.24	22.87	18.91	0	1	22

是近年来在生产上推广面积较大的优良品种，或是我省1981~1983年大豆联合区域试验中新审定的品种，还有的是目前我省一些育种单位拟参加下一轮区域试验的新品系。这些入选的品种或品系共同的特点是多荚、多粒、中大粒类型，主茎节数较多，株高中等，分枝适中或较多，单株粒数也较多，其他性状也都较好。由此可以看出，应用这一较新的数学方法，通过主成份分析及主成份值鉴别入选的材料与育种中常规选择亲本的结果基本符合，所以可以用来评价大豆品种资源，尤其是在进行大量品种多个性状的品种资源研究方面，由于它可以通过少数指标（几个主成份）来反映多个性状的主要遗传信息，从而使筛选的目标明确、集中，这就提高了评价品种资源及筛选优良杂交亲本的效率。

应用多元分析的方法以遗传距离大小将供试品种（品系）聚成6类，计算类内品种间

遗传距离及类群间品种的遗传距离的结果表明，类内品种间的遗传距离较小，不同类群间的遗传距离较大，这与前人在稻麦等作物的研究结果相一致。

聚类后分析品种系谱关系表明，遗传距离的大小在一定程度上反映出亲缘关系，表现在遗传距离较小的类内品种亲缘较近的为多数，而遗传距离较大的类间品种亲缘较远的为多数，但以遗传距离为依据的划类并不完全取决于亲缘关系的远近。例如，有些品种亲缘较近，但并没有划在一类，象公7574—2由（铁丰18×公6602—3）杂交育成，但公7514—2与铁丰18被划在第V与第VI两类；又如九农9号（由九农2号×荆山朴）育成，但九农9与九农2也分别划在第I与第II两类等等。这可能是有些品种虽然近缘，但由于基因型不同，遗传差异仍很大而被划在不同的类间。故遗传距离能较精细地刻划出基因型不同而引起的遗传差异。

遗传距离的大小与亲缘关系有否固定的相关关系，这种相关程度多大？有待进一步深入研究。用主成份分析评价筛选出的亲本，以遗传距离选配规则为指导并结合一般亲本选配原理（如性状互补等）首配组合的优劣，需在育种实践中进一步验证。

由于本试验收集品种（品系）较少，类型尚不丰富。所以第V和第VI类中分别只包括一个品种，当更多品种聚类时，这两类将包括更多的品种。

参 考 文 献

- (1) Bhatt, G.M., 1970, Aust.J. Agric. Res., 21: 1—7.
- (2) 毛盛贤等: 1979, 遗传 1 (5): 26—30.
- (3) 徐静斐等: 1981, 安徽农业科学, <水稻数量遗传论文专辑> 65—71.
- (4) 赵安常: 1981, 安徽农业科学, <水稻数量遗传论文专辑> 71—77.
- (5) 李成荃等: 1981, 安徽农业科学, <水稻数量遗传论文专辑> 77—83.
- (6) 刘来福: 1979, 遗传学报 6 (3): 349—355.
- (7) 潘佩萱: 1983, 安徽农业科学 1: 28—32.
- (8) 李杭萍等: 1983, 我国东北与美国加拿大OOO—IV成熟组的大豆育成品种系谱结构比较, 黑龙江八一农垦大学(油印本)。