

主成份分析在谷子育种中的应用

刘晓辉

(吉林省农科院作物所)

摘 要

本文首次将主成份分析运用于谷子育种中,实验表明,用该法评价组合及亲本与育种实践是吻合的。证明主成份分析运用于谷子育种中是切实可行的,随着计算机的应用,此法用于杂种后代选择也将是有益的。

关键词 谷子 育种 主成份分析 组合评价

谷子是主要的旱粮作物之一。现代育种要求谷子新品种既要有优良的品质,更要具有较高的生产能力。大量的杂交育种实践证明,合理地选配亲本是杂种后代能否有更多的机会出现理想性状组合的重要环节,也是育种家育种成败的关键。遗传学研究表明:在一定范围内双亲的基因型差异越大,其后代的分离幅度也就越广,从中得到优良的机会也就越多,所以育种者常把基因型差异的大小作为杂交育种选择亲本配制组合的原则之一。由于谷子产量等经济性状多为数量性状,受微效多基因控制,又受环境影响较大,而且性状间又有着复杂而微妙的相互联系,这样就增加了杂交育种的难度,因此有关学者寻找适宜的方法去简化那些复杂的问题,自然想到用简化的指标综合反映原来复杂指标的信息,这种将多个相关指标化为少数不相关指标的统计方法称为主成份分析,运用主成份分析进行组合评价,在谷子育种中尚未见报道,本文在谷子育种中应用主成份分析评价其组合优劣,进行早期淘汰组合为深入谷子杂交育种做一些探索性研究。

材料与 方法

本试验在吉林省农科院育种圃进行。利用1988年配制的材料,1989年种植20个组合,5个亲本品种,随机排列,3次重复,3行区,行长2米,行距60厘米。每个组合及品种成熟时,随机取样20株,考查了株高、穗长、码数、千粒重、穗粒数、穗粒重。本文以25个杂交组合及品种(见表1)为素材进行统计分析。详细统计方法请见参考资料^[1,2]。

表1 参 试 材 料 及 编 号

代 号	名 称	代 号	名 称	代 号	名 称
1	龙谷23号	10	南繁1号	19	(铁 ₁ ×7506R)×龙谷23
2	龙谷23×7723	11	7723×(铁 ₁ ×7506R)	20	(铁 ₁ ×7506R)×泌洲黄
3	7723×龙谷23	12	(铁 ₁ ×7506R)×7723	21	泌洲黄×(铁 ₁ ×7506R)
4	泌洲黄 _R	13	铁谷 ₁ ×7506R	22	(铁 ₁ ×7506R)×南繁1号
5	7723×泌洲黄 _R	14	南繁1号×龙谷23	23	南繁1号×(铁 ₁ ×7506R)
6	泌洲黄 _R ×7723	15	龙谷23×南繁1号	24	南繁1号×泌洲黄 _R
7	7723	16	龙谷23×泌洲黄	25	泌洲黄 _R ×南繁1号
8	7723×南繁1号	17	泌洲黄×龙谷23		
9	南繁1号×7723	18	龙谷23×(铁 ₁ ×7506R)		

结果及分析

一、方差分析

首先对供试材料的 6 个性状的考种结果进行方差分析(见表 2)。结果表明,6 个性状的品种间差异均达 0.01 显著水平,说明每个性状间差异主要是品种自身差异提供的,因此对这 6 个性状作进一步分析。

表 2 方差分析结果

变因	株高	穗长	码数	千粒重	穗粒数	穗粒重
基因型	152.410	39.706	690.050	0.4126	6442891	33.748
机 误	26.169	5.336	109.320	0.0316	782640	8.296
F	5.824	7.440	6.312	13.0450	8.232	4.067

注:各性状均达 0.01 显著水平。

表 3 遗传相关矩阵

	株高	穗长	码数	千粒重	穗粒数	穗粒重
R =	1	-0.691	-0.320	0.548	0.339	0.303
	-0.696	1	0.898	0.042	0.104	-0.117
	-0.320	0.898	1	0.420	0.375	0.027
	0.548	0.042	0.420	1	0.752	0.293
	0.339	0.104	0.375	0.752	1	0.491
	0.303	-0.117	0.027	0.293	0.491	1

二、主成份分析

在选择遗传差异显著的 6 个数量性状的基础上,按 $r_{gij} = \frac{cov_{gij}}{\sqrt{\sigma_{g_i}^2 \cdot \sigma_{g_j}^2}}$ 计算出每两个性状间的遗传相关系数,结果给出一个 6 阶遗传相关矩阵(见表 3)。

通过 Jacobi 法进行相似变换,将上述遗传相关矩阵转换成相应的对角矩阵,这样就使得原来相互存在相关关系的 6 个性状因子转变为彼此独立无关因子。然后计算其特征根(λ_i)和累积百分率,以及各特征根相应的特征向量(L_j)。为排除作用较小,干扰较大的综合指标,提高分析精度,从 6 个特征根 $\lambda_1, \lambda_2, \dots, \lambda_6$ 中选取较大的特征根及相应的特征向量,使其累积百分率在 85% 以上,即 $\sum_{i=1}^n \lambda_i / \sum_{j=1}^n \lambda_j > 0.85$ 。据此只选用 3 个特征根和相应的特征向量(见表 4)。

表 4 入选特征根和特征向量

	λ_1	λ_2	λ_3	
特征根	2.47	2.335	0.783	分量
累积贡献率	0.4117	0.8009	0.9315	来源
特征向量	0.399	-0.439	-0.297	株高
	-0.040	0.650	0.057	穗长
	0.203	0.598	-0.120	码数
	0.572	0.078	-0.392	千粒重
	0.564	0.120	0.058	穗粒数
	0.391	-0.083	0.859	穗粒重
	L_1	L_2	L_3	
主成份名称	产量因子	穗部因子	穗粒重因子	

表 4 中各特征根的大小代表各综合指标遗传方差的大小,各特征的累积百分率代表各综合指标对遗传方差贡献的百分率。特征向量表示在各综合指标中,供试材料各性状对综合指标贡献的大小。从表 4 的分析结果看出,第一主成份的特征向量除穗长为较小的负值(-0.040)外,其余均为正值,其中以千粒重、穗粒数、穗粒重的正值较大,这些性状都是构成子实产量的重要因子,三因子任何一个增大,均能大大提高生产力,故我们把第一主成份称为产量因子。第二主成份的特征向量,主要是穗长、码数、穗粒数最大,一般来说穗越大,码越多,粒也多,三者合一,定是一个高产大穗材料,因为穗的结构主要决定

于三者,因此将第二主成份称之为穗部因子。第三主成份以穗粒重最大为 0.859,故称为穗粒重因子。

根据特征根 $\lambda_1, \lambda_2, \lambda_3$ 和相应的特征向量 L_1, L_2, L_3 及亲本各性状的标准化基因型值,可计算出各品种的第一、二、三主成份值:

$$\tilde{g}_{ij} = \frac{1}{\sqrt{\lambda_j}} \sum_{k=1}^n L_{kj} g_{ik}, \tilde{g}_{ij} \text{ 表示第 } i \text{ 个品种的第 } j \text{ 个主成份, } g_{ik} \text{ 为第 } i \text{ 个品种第 } k \text{ 个性状的标准化基因型, } L_{kj} \text{ 为第 } j \text{ 个向量的 } k \text{ 分量。}$$

根据主成份评价各组合的优劣,从子实产量考虑,第一主成份值越大越好,随着产量的提高,码数、千粒重、穗粒数和穗粒重均增加,从穗部结构来考虑,第二主成份值越大越好,随着穗长和码数的增多,株高明显降低。穗粒重是构成产量的重要因素,因此第三主成份值越大越好。

表5 人选组合的主成份值

组合代号	主成份			组合代号	主成份		
	\tilde{g}_1	\tilde{g}_2	\tilde{g}_3		\tilde{g}_1	\tilde{g}_2	\tilde{g}_3
6	1.2500	1.2220	2.1830	19	0.6960	0.2800	0.7236
9	0.5770	0.3666	-0.6705	20	0.1054	-0.1289	1.7920
11	0.3412	-0.1251	1.1289	22	0.6850	0.1809	0.4410
12	1.1469	0.6270	3.3270	24	0.4059	1.2627	0.4886
17	0.0970	0.9797	1.3001	25	0.5630	1.2366	1.5248

按前述公式,计算出25个组合及品种的第一、二、三主成份值 $\tilde{g}_1, \tilde{g}_2, \tilde{g}_3$,再根据上述标准,作“主成份”筛选,对25个品种及组合进行选留时,以1个标准差为标准,可淘汰8个品种及组合,这8个多为亲本(即:品种),这说明杂种优于亲本。以半个标准差为标准,又可淘汰7个组合,这是经第一次淘汰,选留的组合中进行的第二次淘汰。经第一、二次淘汰最后可选留的优势组合共10个(见表5)。作为强化优势组合,并预测强优势组合的方向是:中矮秆、长穗、多码、粒多、大粒和重穗。

三、主成份分析评价

为了进一步明确各主成份所携带的产量信息大小,我们计算了三个主成份值与各性状之间的相关系数(见表6),结果表明:穗粒重与第一、二、三主成份均达极显著水平,相关值

表6 各性状与主成份的相关系数

性状	\tilde{g}_1	\tilde{g}_2	\tilde{g}_3
株高	0.3199	-0.6959	-0.3256
穗长	0.6717**	0.7976**	0.1698
码数	0.7110**	0.7647**	0.0124
千粒重	0.1539	0.3403*	-0.5674**
穗粒数	0.4898**	0.1890	0.9167**
穗粒重	0.7153**	0.4061**	0.8819**

分别为0.7153、0.4661、0.8819。穗长与第一、二主成份达极显著水平 $r=0.6717$ 及 $r=0.7976$ 。码数与第一、二主成份达极显著水平,相关系数分别为0.7116、0.7647。千粒重与第二主成份达显著水平 $r=0.3403$,穗粒数与第一、第三主成份也达极显著水平,相关系数分别为0.4898及0.9467。证明各主成份值携带较多的生物学及产量性状信息。

品种间主成份差异在很大程度上反映着它们的遗传差异,选择遗传差异大的亲本杂交,后代产生广泛的变异,出现各种不同类型的优良个体也多,从而有利于后代的选择及优势组合的选留。

经主成份分析后,淘汰的组合和入选的优良组合与实际育种过程中所选(下转第65页)

A STUDY ON THE MODEL OF RATIONAL FERTILIZATION IN MIDDLE- AND LOW-YIELDING FIELD OF SPRING WHEAT IN MOUNTAIN AREA OF JINGYU COUNTY

Sun Yuliang Huang Jie, Wu Changqing

(*Jingyu County Institute of Agricultural Sciences, Jilin Province*)

ABSTRACT

In this paper, the problem of rational fertilizer application in middle-and low-yielding field of spring wheat in mountain area of Jingyu County was studied by using optimum regression model 311-B. The quadratic regression equation between spring wheat and fertilization of N, P_2O_5, K_2O was established and the optimum nutrient quantity, optimum yield and effects of single fertilizer and their interaction were investigated.

Key words: Mountain area of Jingyu County, Middle-and low-yielding field, Spring wheat, Optimum yield.

=====

(上接第13页)

留的 F_3 代优势组合完全吻合,这说明在谷子育种中的早期世代利用主成份分析淘汰不良组合,选留优势组合是可以尝试的,这样既减少育种程序,又提高了育种效率。

参 考 文 献

- (1)刘来福:作物数量性状的遗传距离及其测定,《遗传学报》,1979,6(3):349~355。
- (2)毛盛贤:冬小麦数量性状遗传差异及其在作物育种上的应用,《遗传》,1979,1(5):26~30。
- (3)明道绪:主成份分析,《遗传距离》,1983,四川农学院印,2:16~21。

APPLICATION OF THE PRINCIPAL COMPONENT ANALYSIS IN MILLET BREEDING

Liou Xiaohui

(*Crop Institute, Jilin Academy of Agricultural Sciences*)

ABSTRACT

This paper explored the application of principal component analysis to millet breeding program. It was shown that the results from evaluation of crosses and parents by the analysis method were consistent with those from practical breeding, indicating the possibility of application of principal component analysis. It is also helpful to use this method in the selection scheme of millet crosses, with computerizing data.

Key words: Millet, Breeding, Principal component analysis, Cross evaluation.