

主成分分析在春大棚黄瓜高产育种中的应用

赵庆媛 张文峰

(吉林省蔬菜花卉研究所, 长春 130031)

提 要 笔者应用主成分分析方法分析了近年来适合于春大棚 18 份黄瓜材料的 7 个数量性状。根据主成分分析原理, 确定 3 个主分量性状。按主分量值的大小, 利用主分量性状的综合性及相互独立性进行方向性综合选择, 选出了综合性状优良的杂交组合及亲本材料。实践证明, 用该法评价春大棚黄瓜组合及亲本材料与育种实践是一致的。

关键词 黄瓜; 数量性状; 主成分分析

多年来, 大量的育种实践表明, 无论是常规育种还是杂种一代优势利用, 有效的选择至关重要。传统的单一性状选择存在着许多弊端, 要研究多个数量性状, 又因为指标较多增加了分析的复杂性, 且在许多情况下, 不同指标之间存在着不同程度的相关, 影响选择的效果。因此, 人们寻求新的方法并进行多性状的综合选择。主成分分析法即是将多个相关指标简化为少数不相关指标的分析方法。本文根据对春大棚黄瓜 18 份材料 7 个数量性状的分析, 研究讨论主成分分析法在春大棚黄瓜高产育种中的应用。

1 材料和方法

试验于 1996 年在吉林省蔬菜花卉研究所保护地黄瓜课题的春大棚试验地进行。试材为保护地黄瓜课题 1990 年以来配制的较优组合 12 份和亲本材料 6 份, 随机区组设计, 3 次重复, 小区面积 7.1 m^2 。除第一雌花节位每小区调查 20 株外, 其余性状均为全小区实测实收。早期单株产量、早期单瓜重、早期单株结瓜数均为 6 月 10 日前的实测数字; 单株产量、单瓜重、单株结瓜数由 7 月 15 日前小区收获量计算而得。统计方法请见参考资料[1, 2]。

2 结果及分析

2.1 主成分分析

2.1.1 主成分性状的特征值、贡献率及累积贡献率的分析

首先对 18 份材料的 7 个数量性状调查结果进行方差分析。结果表明, 18 份材料 7 个性状的品种间差异均达 0.05 以上显著水平, 说明这 7 个性状间的差异主要是品种自身差异造成的, 因此, 对这 7 个性状作进一步分析。

主成分性状的特征值及贡献率分别代表着各主分量的方差值及主分量的贡献百分率, 这是选择主分量的依据, 见表 1。

表 1 主分量性状的特征值、贡献率及累积贡献率

主 分 量	F ₁	F ₂	F ₃	F ₄	F ₅	F ₆	F ₇
特征值	3.910	1.637	0.864 6	0.383 1	0.163 8	0.033 3	0.008 57
贡献率(%)	55.900	23.400	12.4	5.5	2.3	0.5	0.1
累积贡献率(%)	55.900	79.200	91.6	97.1	99.4	99.9	100.0

表 1 的结果表明,在构成所有的主分量性状中,遗传信息主要集中在前几个主分量中,其中前 3 个主分量就包含了总信息量 91.6%的遗传信息。通常不需要全部的主成分,按照累积贡献率大于 85%的原则,用前几个主成分就够了。因而把前 3 个主分量作为选择手段进行综合选择。

2.1.2 主分量中各性状载荷

构成主分量各性状的载荷,体现了该性状在主分量中相对贡献率的大小。

表 2 主分量中各性状载荷(特征向量)

构成主分量的性状		主分量 I (F ₁)	主分量 II (F ₂)	主分量 III (F ₃)
第一雌花节位	(X ₁)	−0.102 2	0.691 0	0.122 0
早期单株产量	(X ₂)	−0.477 4	−0.009 1	−0.219 3
早期单瓜重	(X ₃)	−0.419 4	0.057 4	0.550 2
早期单株结瓜数	(X ₄)	0.000 2	0.701 8	−0.097 5
单株产量	(X ₅)	−0.490 3	−0.012 1	−0.239 4
单瓜重	(X ₆)	−0.402 6	−0.153 9	0.523 5
单株结瓜数	(X ₇)	−0.428 1	−0.052 0	−0.541 7

主分量 I 包含的信息量占总信息量的 55.9%,为最大。从各构成主分量性状载荷看,单株产量最大,其次是早期单株产量、单株结瓜数,因而,可以把主分量 I 看作是产量主分量。主分量 II 体现总信息量的 23.4%,其中早期单株结瓜数、第一雌花节位载荷较大,从其体现的性状载荷可以把它看作是早期单株结瓜数主分量。主分量 III 占总信息量的 12.4%,早期单瓜重、单株结瓜数、单瓜重载荷较大,实际上是反映了单瓜重的情况,因此,可以把主分量 III 看成是单瓜重的主分量。

2.2 独立地进行方向性选择

2.2.1 选择的方向

表 3 各性状间的相关分析

性状	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇
X ₁	1.000 0						
X ₂	0.124 9	1.000 0					
X ₃	0.293 1	0.704 9	1.000 0				
X ₄	0.612 0	0.026 9	−0.000 2	1.000 0			
X ₅	0.168 4	0.942 1	0.681 4	−0.002 8	1.000 0		
X ₆	−0.009 1	0.617 4	0.837 4	−0.145 6	0.673 6	1.000 0	
X ₇	0.084 3	0.864 8	0.426 2	−0.037 1	0.940 4	0.457 9	1.000 0

主分量Ⅰ中载荷最大的为单株产量(X_5)和早期单株产量(X_2)。表3的结果表明,早期单株产量(X_2)与单株产量(X_5)呈极显著正相关,而载荷较大的单株结瓜数(X_7)和早期单瓜重(X_3)也与单株产量呈极显著正相关,而且载荷均为负值。从高产育种这个目标出发,按主分量Ⅰ进行负向选择,就可获得单株产量高的材料。

主分量Ⅱ中,早期单株结瓜数(X_4)、第一雌花节位(X_1)分量载荷明显大于其它分量载荷。早期单株结瓜数(X_4)与单株产量呈极小负相关,第一雌花节位(X_1)与单株产量呈较小正相关。因此,必须在主分量Ⅰ选择的基础上,再根据主分量Ⅱ进行正向选择,即选择该主分量值高的材料,就可获得早期单株结瓜数和第一雌花节位适宜的高产品种。

对于主分量Ⅲ,早期单瓜重(X_3)、单株结瓜数(X_7)、单瓜重(X_6)的分量载荷较大,因单株结瓜数(X_7)、早期单瓜重(X_3)在主分量Ⅰ中已进行了选择,故在主分量Ⅲ选择中不再放在重要位置。单瓜重(X_6)分量载荷为正值,同单株产量(X_5)呈极显著正相关,故按正方向对主分量Ⅲ进行选择,就可达到预期选择的目的。

2.2.2 选择的结果

利用主分量法进行综合选择时,必须考虑各材料的主分量值的大小,表4所列为本试验选用的18份材料的主分量值。

表 4 待选材料的主分量值

材 料	主分量Ⅰ	主分量Ⅱ	主分量Ⅲ	材 料	主分量Ⅰ	主分量Ⅱ	主分量Ⅲ
吉选 2 号	-1.389 7	0.840 7	-0.979 2	9001	4.778 5	-0.488 6	1.485 1
吉选 1 号	-0.155 4	0.614 6	-0.402 6	9002	-0.525 1	-1.205 5	2.016 4
吉选 6 号	-1.534 5	0.968 6	-1.510 2	9003	-1.539 6	-0.110 3	0.228 6
吉选 4 号	0.130 3	-1.175 2	0.447 9	9006	0.886 9	-0.727 7	0.305 7
和平农场密刺	-1.152 6	0.152 4	0.261 3	9201	-0.814 7	0.319 5	0.751 2
山东密刺	-0.470 1	-0.115 4	1.075 5	9203	-1.181 2	0.748 6	-0.451 6
8806	0.502 9	-0.677 5	1.290 8	9204	-2.455 9	1.122 6	0.527 5
8807	2.600 1	-1.537 1	-0.626 1	9401	-1.310 2	0.421 1	-0.185 9
8808	-1.662 3	1.681 4	-0.074 9	9402	3.443 3	-0.955 5	-2.632 0

依据主分量Ⅰ进行负向选择,入选的材料有9204、8808、9003、吉选6号和吉选2号;在入选材料中根据主分量Ⅱ进行正向选择,入选的材料有8808、9204、吉选6号和吉选2号。

在入选的材料中根据主分量Ⅲ的正向选择,首先为9204和8808,其次为吉选6号。

根据主成分分析法,认为在所有的18份供试材料中以9204和8808最好。如果按照常规的方法,主要以单株产量作为衡量杂种一代优势及亲本的标准,但因单株产量遗传力较低,又存在遗传与环境互作较大,高产还必须建立在其构成优良性状的基础之上。9204和8808的评价既注重了其单株产量,而且也注重了其它重要的构成性状,因而,他们的总体结构较好,在实践中具有更加稳定可靠的高产性。在近几年的生产实践中,9204和8808两组合及亲本材料吉选6号在生产及科研中的表现是很理想的。

3 讨 论

农作物本身是一个极其复杂的有机统一体,其农艺性状往往是互相联系着的,当我们对某一性状进行选择时,另外一个或几个性状也跟着发生变化,因此,对单一性状选择时难于保证选择的效果。如果指标过多,很难选留去舍,若将各性状分散于不同世代逐个选择,往

往顾此失彼，很难把多个优良性状综合一体。所以，无论是杂种一代优势利用和亲本材料选择，都应当从多个角度去进行，即进行多性状综合选择。

主成分分析是在不损失过多遗传信息基础上的一种降维思想，使育种者在育种过程中可以把多个单一性状的注意力集中到少数几个主分量性状分析上。在构成主成分时，应本着多维空间椭球各轴互相垂直的思想。对于正态分布来讲，即要求 $COV(F_i, F_j) = 0$ ，且主成分间是互不相关的，因而利用主分量性状可以独立地进行方向性选择，简单易行，提高了选择的效果。

本试验研究仅就这 18 份材料进行的分析。随着计算机的广泛应用，相信主成分分析方法在农作物育种上的应用将越来越受到育种者的重视。

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
2 明道绪·主成分分析·遗传距离, 1983
3 刘晓辉·主成分分析在谷子育种中的应用·吉林农业科学, 1993, (2)
4 张尧庭, 齐开泰·多元统计分析引论·北京: 科学出版社, 1982

The Application of Principal Component
Analysis in Spring Protective Cucumber Breeding

ZHAO Qingyuan et al.

(Research Institute of Jilin Province Vegetable and Flower, Changchun 130031)

Abstract By the method of principal component analysis, seven quantitative characters of 18 cucumber varieties that are adapted to spring protective cultivation have been studied. Three principal components characters were determined according to the principle. The comprehensive and independent characters were applied to oriented comprehensive selection. And we get optimum hybrids and parents materials. According to the experimental results, the estimation of genetic components and parents materials of spring protective cucumber is agreed with usual breeding practice.

Key words Spring protective cucumber, Principal component analysis, Oriented comprehensive selection; Quantitative character

(责任编辑: 张 瑛)