

# 主成分分析在油用向日葵 杂交种选育中的应用

李素萍,安玉麟,李巧枝,门果桃,于海峰

(内蒙古农业科学院 作物所,内蒙古 呼和浩特 010031)

**摘要:**文章应用主成分分析方法,对 18 个油葵杂交种的 12 个性状进行分析,分析的结果与品种的实际表现基本一致。结果证明,由于各遗传性状对品种的重要性不同,运用主成分分析法对向日葵杂交种进行多性状综合评判,比单纯依靠产量结果对参试品种进行评判更具有科学性。这种方法可作为品种比较试验中对参试种进行综合评判的一种有效方法。

**关键词:**向日葵杂交种;遗传性状;主成分分析

中图分类号:S565.535.1

文献标识码:A

文章编号:1000-7091(2004)S1-0001-05

在向日葵育种过程中,为了选育具有高产、抗病、质优、株型理想等性状优良的向日葵杂交种,往往要对向日葵杂交种的诸多性状加以研究,从中选出理想的向日葵新品种。主成分分析法,就是将这些诸多性状进行适当数学处理,使之成互不相关的变量,并选择其中影响较大的几个,然后逐个地对它们进行分析,使育种者能从繁杂的试验数据中透析事物的本质。

## 1 材料和方法

试验材料为 2002 年内蒙古自治区油用向日葵区域试验的 18 个品种,试验在内蒙古农科院试验地进行,试验采用随机区组设计,3 次重复,小区长 7.36m,宽 4.20m,行距 0.70m,株距 0.23m,小区面积为 30.91m<sup>2</sup>。每小区种 6 行,每行种 32 株,共计 192 株。种植密度为 4141 株/667m<sup>2</sup>。试验方法:测定株高、茎粗、叶片数、花盘直径、单头籽粒重、百粒重、单盘粒数、结实

表 1 品种代号及名称

编号	供试品种	编号	供试品种
N1	C17	N10	巴葵杂 2 号
N2	S98	N11	巴葵杂 4 号
N3	S40	N12	内葵杂 3 号
N4	S46	N13	LG2002
N5	CJ01	N14	12244
N6	康地 115	N15	F51
N7	KWS305	N16	F52
N8	KWS301	N17	G101(CK1)
N9	KWS307	N18	S3ICK2)

收稿日期:2004-11-11

基金项目:农业部“948”项目(编号 2003-T13)

作者简介:李素萍(1963-),女,巴彦淖尔人,副研究员,在读硕士,主要从事向日葵遗传育种研究工作。

率、籽仁率、籽实含油率、生育期、产量共 12 个性状。统计分析方法:首先以各性状小区平均数进行方差分析,经 F 测验后舍去不显著的性状,然后应用 DPS 数据处理系统进行主成分分析,建立相关矩阵表(表 2),求出公因子对所有变量的公因子的总贡献率,从而把一些错综复杂的多因子归结为几个较少的综合因子。

## 2 结果分析

### 2.1 性状的相关矩阵

通过计算 12 个性状间的相关系数得到相关矩阵表(表 2)。

表 2 相关矩阵

相关系数	X(1)	X(2)	X(3)	X(4)	X(5)	X(6)	X(7)	X(8)	X(9)	X(10)	X(11)	X(12)
生育期(d)	1	0.84683	0.77538	0.58983	0.37487	0.15154	0.10356	0.55054	-0.38773	-0.04916	0.06069	0.15253
株高(cm)	0.84683**	1	0.82223	0.64752	0.39766	0.40205	0.11759	0.57904	-0.38214	-0.08188	0.13259	0.40236
茎粗(cm)	0.77538**	0.82223**	1	0.37676	0.53282	0.35161	0.20642	0.38635	-0.45922	-0.06266	-0.12757	0.35259
叶片(片)	0.58983*	0.64752**	0.37676	1	0.02363	0.20978	-0.1278	0.65388	-0.2168	-0.05643	0.13565	0.21093
盘茎(cm)	0.37487	0.39766	0.53282*	0.02363	1	-0.07161	-0.0806	0.24537	-0.29638	0.25926	-0.04382	-0.0699
单头籽实重(g)	0.15154	0.40205	0.35161	0.20978	-0.07161	1	0.30098	0.02464	0.20853	-0.53018	0.17593	0.99998
百粒重(g)	0.10356	0.11759	0.20642	-0.12781	-0.08063	0.30098	1	-0.35434	0.11162	0.27388	-0.20405	0.30098
单盘粒数	0.55054*	0.57904*	0.38635	0.65388**	0.24537	0.02464	-0.3543	1	-0.4625	-0.20052	0.06554	0.02579
结实率(%)	-0.38773	-0.38214	-0.45922	-0.2168	-0.29638	0.20853	0.11162	-0.4625	1	-0.14194	0.24113	0.2094
籽仁率(%)	-0.04916	-0.08188	-0.06266	-0.05643	0.25926	-0.53018*	0.27388	-0.20052	-0.14194	1	-0.36896	-0.52974
籽实含油率(%)	0.06069	0.13259	-0.12757	0.13565	-0.04382	0.17593	-0.2041	0.06554	0.24113	-0.36896	1	0.17457
小区产量(kg)	0.15253	0.40236	0.35259	0.21093	-0.0699	0.99998**	0.30098	0.02579	0.2094	-0.52974*	0.17457	1

注:  $v=16$   $r_{0.05}=0.468$ ,  $r_{0.01}=0.590$

由表 2 看出:小区产量与单头籽实重正相关系数最高;株高与生育期、叶片数,茎粗与株高、生育期、盘径、单盘粒数与叶片数、生育期、株高也呈较高的正相关;籽仁率与单盘籽实重、小区产量呈负相关,所以,生产科研上选用单头籽实重、生育期长、茎秆粗壮,叶片多、植株相对高一些的品种或亲本可以提高油用向日葵的产量。

### 2.2 特征根及特征向量

从遗传相关矩阵的 12 个特征根中选取 5 个较大的(累计贡献率达 86.9632%)作为主成分,基本上保存了它们的生物信息。进一步的分析发现,主成分及其变异的大小与它们的信息

大小相平行,因此可望利用主成分分析来研究各性状之间的关系。我们把彼此关联并制约的12个性状简化为独立无关的主成分(表3)。

表3 特征根与特征向量

$\lambda_i$	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$
	4.1876	2.6492	1.7565	0.9570	0.8853
累计贡献率(%)	34.8969	56.9732	71.6107	79.5853	86.9632
生育期(d)	0.4214	-0.1238	0.0395	-0.0229	0.2141
株高(cm)	0.4638	-0.0142	0.0564	0.0106	0.1653
茎粗(cm)	0.4200	-0.0614	0.2495	0.1676	-0.1383
叶片数(个)	0.3391	-0.0316	-0.2327	-0.4566	0.3223
盘茎(cm)	0.2065	-0.2612	0.1606	0.7065	-0.0221
单头籽实重(g)	0.2233	0.5114	0.1351	0.0096	-0.1540
百粒重(g)	0.0246	0.1523	0.6230	-0.2144	0.2704
单盘粒数	0.3336	-0.1717	-0.3624	-0.1877	-0.1154
结实率(%)	-0.2196	0.3560	-0.0051	0.1038	0.4398
籽仁率(%)	-0.1048	-0.3971	0.3708	-0.1006	0.4076
籽实含油率(%)	0.0475	0.2273	-0.3993	0.4019	0.5567
小区产量(kg)	0.2237	0.51102	0.1355	0.0097	-0.1539

由表3可以看出,入选的5个主成分对遗传方差的贡献率分别为34.89%、22.08%、14.64%、7.97%、7.38%,依次称为第一、二、三、四、五主成分。第一主成分 $\lambda_1=4.1876$ ,特征向量中性状组合系数以株高(0.4638)、生育期(0.4214)、茎粗(0.4200)、叶片数(0.3391)较大,可称之为生育期和形态因子,在选择品种和亲本时,该指标应适中,太大则生育期延长,太小则植株偏矮,茎秆偏细,叶片数少,光合产物少,影响产量;第二主成分 $\lambda_2=2.6491$ ,性状组合系数中以小区产量(0.5110)、单盘籽实重(0.5114)和结实率(0.3560)最大,因这3个性状均为重要的产量性状,故称之为产量因子,选择品种或亲本时,该主成分值越大越好; $\lambda_3=1.7565$ ,主要分量来源为百粒重(0.6230)、籽仁率(0.3708),但单盘粒数和籽实含油率的组合系数为较大的负值,即是说,百粒重和籽仁率增加到一定程度,会引起单盘粒数和籽实含油率下降,故此值应适中,不宜过大; $\lambda_4=0.9570$ ,性状组合因子以盘径(0.7065)、籽实含油率(0.4019)为最大,但叶片数为较大的负值,因此,此值应适中; $\lambda_5=0.8853$ ,性状组合系数为籽实含油率(0.5567)、结实率(0.4398)、籽仁率(0.4076),称之为籽粒品质性状,在品种和亲本选育过程中,此值越高越好。

### 2.3 品种的主成分评价

由表4可以看出,参试品种生育期与形态因子主成分( $Y_{11}$ )值以C17、S40、CJ01、康地115、S31居高,说明这5个品种的生育期较长,植株较高,茎秆粗壮,单株叶片多,具有较好的生育期与形态因子;产量因子主成分( $Y_{12}$ )值以C17、12244、KWS307、内葵杂三号4个品种较高,说明这些品种的综合产量性状好,该主成分可以作为向日葵品种选育中的丰产性指标; $0.5000 \geq Y_{13} \geq -1.2895$ ;  $1.0000 \geq Y_{14} \geq -0.5000$ ;  $1.000 \geq Y_{15} \geq -0.5000$ 。根据此标准入选的品种有C17、KWS307、内葵杂三号、12244 4个品种。

### 3 讨论

3.1 对于这 18 个油用向日葵杂交种而言, $\lambda_1$  作为生育期和形态因子,在选择品种和亲本时,该指标应适中,太大则生育期延长,太小则植株偏矮,茎秆偏细,叶片数少,光合产物少,影响产量;第二主成分  $\lambda_2$  越大产量将越高,并且是通过单盘籽实重和结实率来提高产量,故在选择

表 4 18 个品种的主成分值

No	Y(i,1)	Y(i,2)	Y(i,3)	Y(i,4)	Y(i,5)
C17	3.56897	3.48265	-1.28954	0.70901	-0.72106
S98	0.55691	0.769010	0.39678	-1.55050	0.79675
S40	4.25170	-1.86810	-1.79673	-0.39790	0.53464
S46	-1.57166	-0.72743	-0.89595	-0.31972	1.67911
CJ01	2.10024	-3.37646	1.78222	0.49405	-0.23996
康地 115	1.40523	-1.02486	-1.18323	-0.48655	0.35570
KWS305	-2.45692	-2.38223	1.38842	-0.22673	-1.46342
KWS301	0.55966	0.70837	0.59945	0.60781	0.17056
KWS307	0.92517	1.88053	3.28949	-0.76143	0.34606
巴葵杂二号	-2.14795	-0.71394	-0.34535	-0.63821	0.62622
巴葵杂四号	-0.56364	-0.8804	0.14530	3.06912	0.49189
内葵杂三号	-2.12015	1.29652	0.48307	-0.31029	-0.05782
LG2002	-1.40383	0.00370	-0.83120	0.65177	1.06980
12244	-1.97749	2.22035	0.15181	1.04324	-0.47137
F51	-1.99320	0.79746	-0.79802	-0.53349	0.63046
F52	-1.00266	-0.07782	-0.82928	-0.85103	-0.94334
G101(ck1)	-0.81983	-0.39868	-1.81196	-0.30791	-2.39668
S31(ck2)	2.68947	0.29134	1.54472	-0.19122	-0.40754

品种或亲本时,该主成分值越大越好; $\lambda_3$  与百粒重和籽仁率呈正相关,但与单盘粒数和籽实含油率呈负相关,因此  $\lambda_3$  应适中,不宜过大; $\lambda_4$  与盘径和籽实含油率呈正相关,但叶片数为较大的负值,故此值以适中为宜; $\lambda_5$  与籽实含油率、结实率和籽仁率呈正相关,在品种和亲本选育过程中,此值越高越好。

3.2 以本文所研究的 18 个油葵杂交种为例,选育晚熟和产量、含油率高的油葵杂交种,那么  $\lambda_1$ 、 $\lambda_2$ 、 $\lambda_5$  较大, $\lambda_3$ 、 $\lambda_4$  适中。如果选育早熟、高产、优质的油葵杂交种,则以  $\lambda_1$  较小, $\lambda_2$ 、 $\lambda_5$  较大, $\lambda_3$ 、 $\lambda_4$  适中为宜。

3.3 由于主成分值是众多性状生物学信息的凝聚,而各主成分之间又互不关联,这对于各性状的选择和简化油用向日葵杂交种的选育程序是有利的,随着电子计算机在育种现代化方面的应用和普及,这种方法是可行的。

**参考文献：**

- [1] 卫文星. 苎麻品种主成分分析和遗传距离测定及其在杂交育种中的应用[J]. 华北农学报, 1994, 9(3): 29-33.
- [2] 李宗仁. 主成分分析和聚类分析在小麦遗传距离上的综合应用[J]. 青海大学学报, 1995, 13(4): 38-42.
- [3] 印志同. 玉米自交系性状的遗传相关分析和主成分分析[J]. 扬州大学学报, 2001, 4(1): 48-51.
- [4] 曹靖生. 玉米杂交种的主成分分析[J]. 玉米科学, 1994, 2(1): 21-24.
- [5] 安玉麟. 中国向日葵产业发展的问题与对策[J]. 内蒙古农业科技, 2004, (4): 1-4.
- [6] 门果桃, 安玉麟, 郭富国, 等. 油用向日葵部分性状与籽实含油率的相关性研究[J]. 内蒙古农业科技, 2001, (6): 10-12.

## **Application of Principal Component Analysis in Oilseed Sunflower Cross-breeding**

LI Su-ping, AN Yu-lin, LI Qiao-zhi, MEN Guo-tao, YU Hai-feng  
(Crop Institute Inner Mongolia Academy of Agricultural Sciences, Huhhot 010031, China)

**Abstract:** Principal component analysis was conducted to test 12 inheritable characters in 18 oilseed sunflower crossed varieties. The result showed that principal component analysis was an effective and scientific way to compare the synthetical inheritable characters of candidate varieties in variety-comparing test since the method used to analyze multiple characters was more scientific than single yield character analysis.

**Key words:** Sunflower crossed variety; Inheritable character; Principal component analysis