

应用聚类分析和主成分分析法评价洋葱新品种

朱新明, 常国军, 何丹, 赵海霞, 马超

(甘肃省酒泉市农业科学研究院 甘肃 酒泉 735000)

摘要: 对酒泉市引进的27个洋葱新品种(系)叶片形态、叶色、生育期、产量等15个因子聚类分析, 按其鳞茎颜色、产量、品质综合分为5类。在此基础上, 对15个因子变量观测值进行主成分分析, 前4个公因子对所考察性状变异的累计方差贡献率达76.849%, 较好的反映了所有性状大部分信息。根据品种在各公因子上的得分值, 对27个品种综合评价, 综合得分按聚类分析结果分类排序, 确定建议示范推广品种。

关键词: 洋葱; 品种; 聚类分析; 主成分分析; 综合评定

中图分类号: S633.2 **文献标识码:** A **文章编号:** 1001-1463(2014)10-0025-04

doi: 10.3969/j.issn.1001-1463.2014.10.010

Evaluation of New Onion Varieties Using Cluster Analysis and Principal Component Analysis Methods

ZHU Xin-ming, CHANG Guo-jun, HE Dan, ZHAO Hai-xia, MA Chao

(Jiuquan Institute of Agricultural Sciences, Jiuquan Gansu 735000, China)

Abstract: Using the methods of cluster analysis methods, 15 factor of introducing 27 new onions varieties were selected, such as leaf shape, leaf color, growth period, yield etc, according to its bulb color, yield, quality comprehensive divided into five categories. On this basis, the method of principal components analysis is adopted to analyze factor variables for 15 observations, the results shows that cumulative total variance contribution rate amounted to 76.849% of previous 4 common factors, which reflected the most of the characteristic information; according to the scores of all varieties in all common factor, assessing the characteristic of 27 varieties, which sorted by comprehensive score and the result of cluster analysis, this results is used to determine to popularize and apply varieties.

Key words: Onion; Varieties; Clustering analysis; Principal component analysis; Comprehensive evaluation

酒泉市是全国重要的洋葱产区^[1], 种植面积近 13 333.33 hm²。近年, 各种子企业引进了大量的品种, 给农民选择品种带来困难。2013 年, 酒泉市农业科学研究院对酒泉市引进的 27 个洋葱品种应用系统聚类分析和主成分分析结合的方法进行了综合评价, 现将结果报道如下。

1 材料与方法

1.1 供试材料

参试品种(系)27个, 为新红早丰、卡丽布、庄园 1 号、保罗、黄皮 02、红运紫星、欢呼 6064、卡钦、红盾 801、紫钰、红天、新红福、千里马、新丰田黄金、新丰田红珠、新丰田黄贵人、新丰

收稿日期: 2014-01-14

作者简介: 朱新明(1966—), 男, 甘肃酒泉人, 高级农艺师, 主要从事种子学研究工作。联系电话: (0)13993706139。

执笔人: 常国军

836.4 元/hm²。说明缓释包衣尿素在河西平川灌区玉米保护性耕作中具有推广价值。

参考文献:

- [1] 翟军海, 高亚军, 周建斌. 控释/缓释肥料研究概述[J]. 干旱地区农业研究, 2002, 20(3): 45-48.
- [2] 崔静, 谷思玉, 李菊梅, 等. 缓/控释肥养分释放特性的评价方法概述[J]. 化肥工业, 2009, 36(1): 23-27.
- [3] 杜建军, 廖宗文, 宋波, 等. 包膜控释肥养分释放特性评价方法的研究进展[J]. 植物营养与肥料学报,

2002, 8(1): 16-21.

- [4] 何刚, 张崇玉, 王玺, 等. 包膜缓释肥料的研究进展及发展前景[J]. 贵州农业科学, 2010, 38(6): 141-145.
- [5] 侯翠红. 控制释放肥料养分释放特性的研究[J]. 磷肥与复肥, 1998(4): 6-8.
- [6] 张向波, 聂春柏, 吴方, 等. 5种除草剂对保护性耕作春玉米苗期田间杂草的防效[J]. 甘肃农业科技, 2012(6): 27-29.

(本文责编: 郑立龙)

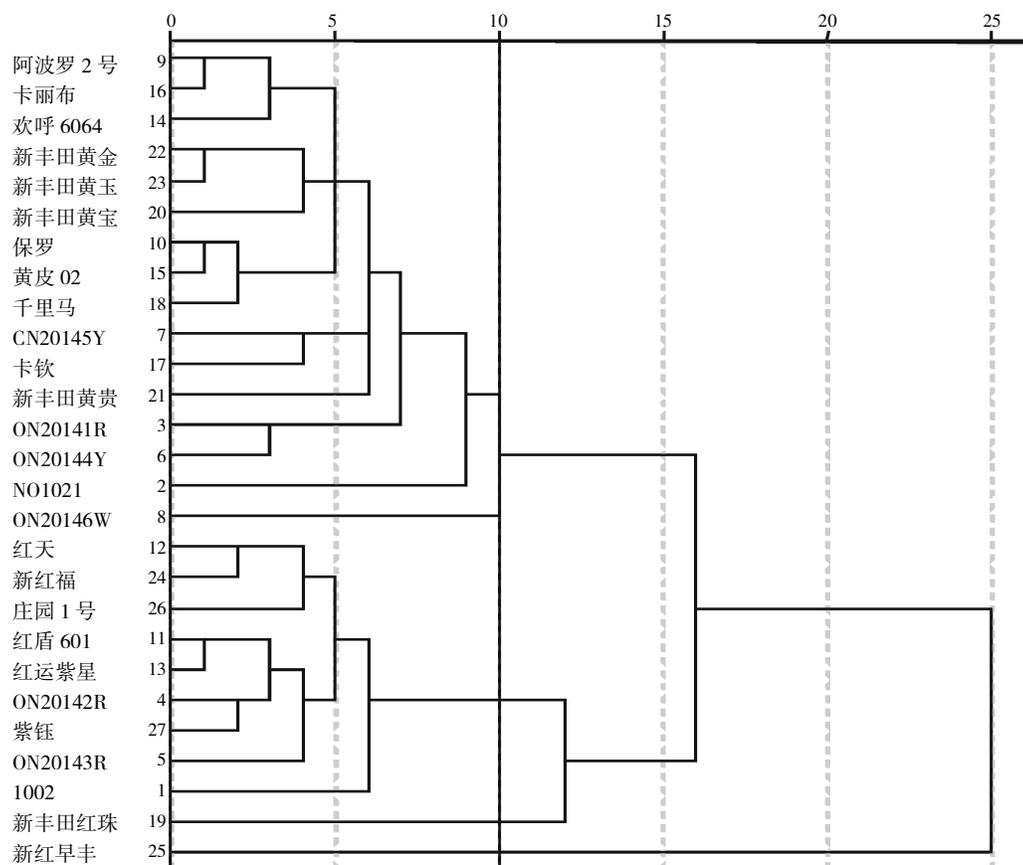


图 1 27 个品种(系)欧氏距离聚类

田黄玉、新丰田黄宝、阿波罗 2 号、1002、NO102、ON20141R、ON20142R、ON20143R、ON20144Y、ON20145Y、ON20146W，均由酒泉市洋葱经销商提供。

1.2 试验方法

试验采用随机区组排列，3 次重复，小区面积 4.0 m^2 ($2.5 \text{ m} \times 1.6 \text{ m}$)，每小区 9 行，行距 15 cm，株距 16 cm，小区间距 25 cm，重复间间距 50 cm，折合 $35.16 \text{ 万株/} \text{hm}^2$ 。试验覆黑色地膜育苗移栽。管理同当地大田，小区收获计产。

对品种的叶片形态、叶色、生长天数、鳞茎纵径、横径、球形、皮色、亮度、着色皮层数、鳞茎皮脱落度、硬度、收口、单球重、干皮厚度、产量 15 个变量的观测值进行系统聚类分析和主成分分析^[2]，聚类分析按欧氏距离进行聚类。聚类分析和主成分分析用统计分析软件 IBM SPSS19.0 的分类和降维两个分析模块^[3]。在聚类分析和主成分分析的基础上，对品种的特征特性进行综合评定。

2 结果与分析

1.2 聚类分析

从图 1 分析，归入同一类的品种特征特性相似性较高，彼此的遗传距离较小，具有相近的鳞

茎皮色、生育期、品质及产量，而不同类群之间的品种特征特性具有较大的差异。根据洋葱品种特征特性，将类群间区分的欧氏距离定为 10，分为 5 类，符合洋葱品种类群区分习惯，能较好的反映洋葱品种不同类群的特征特性。洋葱品种在酒泉通常按鳞茎皮色分为铜葱（俗称黄葱）、红葱和白葱品种。按生育期分为早熟、中晚熟、晚熟品种。分类结果见表 1。

从表 1 看出，第 1 类共有 15 个品种(系)，鳞茎皮色为铜色和深铜色，生育期 112 ~ 124 d，属于晚熟品种，单球重平均 281.74 g；第 2 类 1 个品种，鳞茎皮色为白色，生育期 114 d，属于晚熟品种，单球重 275.7 g；第 3 类共 9 个品种(系)，鳞茎皮色为红色或深红色，生育期 104 ~ 117 d，属于中晚熟品种，单球重平均 202.40 g；第 4 类和第 5 类品种各 1 个，鳞茎皮色为深红色和浅红色，生育期为 94 d 和 84 d，属于早熟品种，单球重分别为 250.7、199.3 g。

1.3 主成分分析

主成分分析因子方差提取见表 2。提取的公因子方差从表 2 的结果可知，15 个变量的共性方差除收口为 0.476 外，其余均大于 0.5，故提取的公

表 1 各类群内品种与表现

类群	品种(系)数 (个)	品种(系)名称	品种(系)特征
1	15	卡丽布、黄皮02、保罗、新丰田黄玉、阿波罗2号、ON20145Y、卡钦、ON20141R、ON20144Y、欢呼6064、千里马、新丰田黄贵人、新丰田黄宝、NO1021、新丰田黄金	植株叶色多数为绿色,生育期112~124 d。鳞茎球形或长圆型,皮色为铜色或深铜色,皮层着色3~4层,单球重213.5~336.1 g,平均281.74 g。
2	1	ON20146W	植株叶色绿色,生育期114 d。鳞茎球形,皮色为白色,单球重平均275.7 g。
3	9	1002、ON20142R、ON20143R、红盾801、红天、红运紫星、庄园1号、紫钰、新红福	植株叶色深绿,生育期104~117天。鳞茎球形或长圆型,皮色红色或深红色,皮层着色5~8层,单球重160.4~243.1 g,平均202.4 g。
4	1	新丰田红珠	植株叶色深绿色,生育期94 d。鳞茎扁圆形,皮色为深红色,皮层着色7层,单球重平均250.7 g。
5	1	新红早丰	植株叶色为绿色,生育期84 d。鳞茎球形,皮色为浅红色,单球重平均199.3 g。

表 2 公因子方差

成分	初始	提取	成分	初始	提取
叶态	1	0.786	着色皮层	1	0.865
叶色	1	0.753	皮脱落度	1	0.803
生长天数	1	0.853	硬度	1	0.847
纵径	1	0.762	收口	1	0.476
横径	1	0.808	单球重	1	0.78
球形	1	0.714	干皮厚	1	0.728
皮色	1	0.81	产量	1	0.822
亮度	1	0.72			

因子能够很好地反映原始变量的主要信息。

如果主成分分析中所提取的多个主成分的特征值贡献率能达到 70% 以上, 就可以用这几个主成分对事物的属性进行概括性分析, 基本可以得出影响事物性质的主要因素^[4]。从表 3 看出, 提取前 4 个主成分累计贡献率已达 76.849%, 可以概括不同洋葱品种主要特征特性的 15 个性状的绝大部分信息。

表 3 27 个洋葱品种(系)主成分相关矩阵的特征值

成分	初始特征值			提取特征值		
	特征值	贡献率 (%)	累积贡献率 (%)	特征值	贡献率 (%)	累积贡献率 (%)
1	7.267	48.447	48.447	7.267	48.447	48.447
2	1.939	12.927	61.374	1.939	12.927	61.374
3	1.225	8.164	69.538	1.225	8.164	69.538
4	1.097	7.311	76.849	1.097	7.311	76.849
5	0.915	6.101	82.950			
6	0.718	4.789	87.739			
7	0.593	3.954	91.693			
8	0.398	2.650	94.343			
9	0.335	2.234	96.577			
10	0.209	1.392	97.969			
11	0.114	0.760	98.729			
12	0.080	0.534	99.263			
13	0.063	0.423	99.686			
14	0.030	0.200	99.886			
15	0.017	0.114	100.000			

表 4 为因子得分系数矩阵, 由此可得最终因子得分公式:

$$F_1=0.261Zx_1-0.678Zx_2+\dots+0.796Zx_{14}+0.865Zx_{15}$$

$$F_2=0.193Zx_1+0.519Zx_2+\dots+0.26Zx_{14}-0.088Zx_{15}$$

$$F_3=-0.632Zx_1+0.015Zx_2+\dots-0.125Zx_{14}-0.251Zx_{15}$$

$$F_4=-0.53Zx_1-0.156Zx_2+\dots-0.103Zx_{14}+0.053Zx_{15}$$

$Zx_1、Zx_2、Zx_3、\dots、Zx_{15}$ 为叶态 x_1 、叶色 x_2 、生长天数 x_3 ……产量 x_{15} 经过标准化处理的值, 数据标准化是指 Z 标准化。

第 1 主成分产量、单球重、硬度系数较大, 其次是干皮厚度、皮脱落度、纵径、横径的系数较大, 可归纳为产量因子; 第 2 主成分生育期系数较大, 为生育期因子; 第 3 主成分纵径、收口、横径较大、为球型形态因子; 第 4 主成分球型系数较大, 为商品率因子。

表 4 27 个洋葱品种(系)主要性状的主成分相关矩阵成分矩阵

性状	x	主成分1	主成分2	主成分3	主成分4
叶态	x_1	0.261	0.193	-0.632	-0.530
叶色	x_2	-0.678	0.519	0.015	-0.156
生长天数	x_3	0.571	0.603	0.371	-0.161
纵径	x_4	0.717	-0.117	0.447	-0.184
横径	x_5	0.726	-0.397	0.351	-0.017
球形	x_6	-0.324	0.198	-0.128	0.744
皮色	x_7	-0.748	0.433	0.103	-0.229
亮度	x_8	0.740	0.396	0.011	0.127
着色皮层	x_9	-0.764	0.516	0.060	0.106
皮脱落度	x_{10}	0.747	0.424	-0.077	0.243
硬度	x_{11}	0.834	0.299	-0.219	0.122
收口	x_{12}	0.456	0.322	0.404	-0.02
单球重	x_{13}	0.855	-0.029	-0.196	0.103
干皮厚度	x_{14}	0.796	0.260	-0.125	-0.103
产量	x_{15}	0.865	-0.088	-0.251	0.053

表 5 因子得分矩阵

品种(系)名称	类群	F_1	F_2	F_3	F_4	综合得分	排名
NO1021	1	0.005 62	-0.1315 4	-0.908	1.200 1	-0.07	15
ON20141R	1	0.028 49	-0.363 07	1.041 6	0.123 06	6.09	14
ON20144Y	1	0.485 96	-0.232 19	0.582 24	0.661 99	30.14	9
ON20145Y	1	0.896 41	0.049 5	1.144 58	1.517 54	64.51	2
阿波罗2号	1	0.633 17	-0.529 48	-0.843 38	-0.236 19	15.22	12
保罗	1	1.288 27	0.133 75	-0.062 42	-0.917 92	56.92	5
欢呼6064	1	0.869 71	-1.080 5	-0.238 48	-0.479 95	22.71	11
黄皮02(CK)	1	1.206 87	0.247 48	-0.802 5	-0.595 49	50.76	6
卡丽布	1	0.522 79	-0.362 67	-0.494 63	-0.363 16	13.95	13
卡钦	1	0.882 67	0.210 35	1.409 86	0.531 2	60.88	4
千里马	1	1.255 27	0.347 68	0.024 52	0.045 3	65.84	1
新丰田黄宝	1	0.727 43	0.895	-1.798 76	-0.858 06	25.85	10
新丰田黄贵人	1	1.181 89	0.484 17	0.645 63	-0.638 38	64.12	3
新丰田黄金	1	0.854 15	0.067 38	-0.66 94	1.137 31	45.1	7
新丰田黄玉	1	0.754 52	0.402 48	-1.712 76	0.454 69	31.1	8
ON20146W	2	0.677 33	-1.842 66	2.352 66	0.002 65	28.22	
1002	3	-1.000 01	0.864 28	0.682 01	0.020 24	-31.56	2
ON20142R	3	-1.289 51	0.490 63	-0.060 74	0.097 34	-55.91	8
ON20143R	3	-1.012 23	1.202 48	1.135 55	-0.341 61	-26.72	1
红盾801	3	-0.966 77	0.675 81	0.352 74	-1.240 26	-44.29	7
红天	3	-1.092 59	0.962 56	-0.204 19	-0.003 93	-42.19	6
红运紫星	3	-0.820 79	0.155 98	1.005 83	-1.446 31	-40.11	5
新红福	3	-1.180 63	0.577 74	0.561 26	0.860 82	-38.85	4
庄园1号	3	-0.639 95	0.943 23	-1.042 78	-0.898 56	-33.89	3
紫钰	3	-1.481 06	0.072 82	-0.179 24	-0.802 94	-78.15	9
新丰田红珠	4	-1.348 93	-0.491 25	-0.900 21	3.201 23	-55.65	
新红早丰	5	-1.438 07	-3.749 96	-1.021 01	-1.030 71	-134.02	

1.4 综合得分评价

对 27 个品种(系)进行综合评价,对 4 个公因子 F_1 、 F_2 、 F_3 、 F_4 得分进行加权求和,权数取方差贡献值。

$$\text{综合得分} = 48.477F_1 + 12.927F_2 + 8.164F_3 + 7.311F_4$$

综合得分结果按聚类分析的 5 个类群分别进行排序,结果见表 5。从表 5 看出,在第 1 类皮色为铜色的 15 个晚熟品种组中,综合得分高的品种(系)有千里马、ON20145Y、新丰田黄贵人、卡钦、保罗;第 3 类皮色为红色或深红色的 9 个中晚熟红葱品种组中,得分较高的为 ON20143R、1002、庄园 1 号和新红福;第 2、4、5 类各只有 1 个品种,不做排序。第 4、5 类的两个品种虽然各独立成为 1 类,但外在表现和生育期基本接近,属早熟红葱品种。其中新丰田红珠得分较高。以上综合得分评定结果与品种在田间的表现吻合。

3 小结与讨论

1) 对 27 个洋葱新品种(系)的叶片形态、叶色、生育期、产量等 15 个因子进行聚类分析,按鳞茎颜色、产量、品质综合分为 5 类。在此基础上,对 27 个品种的 15 个因子变量观测值进行主成分分

析,前 4 个公因子对所考察性状变异的累计方差贡献率达 76.849%,较好的反映了所有性状大部分信息。

2) 根据各品种(系)在各公因子上的得分值,对 27 个品种(系)综合评价,综合得分按聚类分析结果分类排序。皮色为铜色的 5 个晚熟品种(千里马、ON20145Y、新丰田黄贵人、卡钦、保罗)、4 个中晚熟红葱品种(ON20143R、1002、庄园 1 号和新红福)、早熟红葱品种新丰田红珠以及晚熟白葱品种 ON20146W 综合得分较高,可作为进一步示范推广的品种。

参考文献:

- [1] 蒲兴秀,马令卉.酒泉市洋葱生产现状与发展对策[J].甘肃农业科技,2004(11):32-34.
- [2] 唐守正.多元统计分析方法[M].北京:中国林业出版社,1984:20;61.
- [3] 杜强,贾丽艳.SPSS 统计分析从入门到精通[M].北京:人民邮电出版社,2013:248;260.
- [4] 辛秀璐,于凤瑶,周顺启,等.黑龙江省近期审定大豆品种的聚类分析和主成分分析[J].浙江农业学报,2010(4):806-810.

(本文责编:陈珩)