

甜玉米自交系的模糊聚类分析

党拥华 徐占宏 李克祥 李开俊

(吉林市农科院,吉林 132101)

提 要 本文应用模糊聚类分析对 15 份优良甜玉米自交系进行聚类分析,结果表明:甜粒、白甜、油玉米及黄甜 4 份材料分类系统较近,而与 42-1 甜,60GS 甜,97-1 甜,98-1 甜自交系分类系统较远,此结果为甜玉米的选育及开发利用提供了科学依据。

关键词 甜玉米;自交系;模糊聚类分析

随着生活水平的提高,人们的膳食习惯也在由温饱型转向营养型。被国外喻为水果、蔬菜的甜玉米,在我国也走进了千家万户。它除青食外,还被制成甜玉米子粒罐头,玉米笋罐头,速冻玉米等销往国内大城市及国外,而且供不应求,因此甜玉米品种的选育,尤其作为育种基础材料的自交系就更为重要。探索自交系间在系统进化中所处的地位及彼此间的相互关系能为选系提供科学依据,所以对甜玉米自交系进行聚类分析,显得十分必要。

1 试验材料及来源

本试验是 1991 年在吉林市农科院玉米试验田中进行,从 93 份甜玉米自交系材料中筛选出 15 份优良的材料。调查记载的性状为:生育期、株高、穗位、穗长、行数、行粒数、百粒重和单穗粒重。

2 统计方法和步骤

2.1 确定论域 V ,列出自交系的各性状表,对表中的原始数据进行标准化处理

论域 $V = \{U_1, U_2, \dots, U_{15}\}$,采用极差法对数据进行标准化处理^[1]。公式为

$$U_{ij} = \frac{U_{ij} - \min\{U_{ij}\}}{\max\{U_{ij}\} - \min\{U_{ij}\}} \quad \max\{U_{ij}\} \text{ 表示第 } j \text{ 列原始数据最大值。} \min\{U_{ij}\} \text{ 表示第 } j \text{ 列原始数据最小值。}$$

处理后数值列入表 1。

表 1 15 份甜玉米自交系标准化处理后性状表

编号	自交系名称	生育期	株 高	穗 位	穗 长	行 数	行粒数	百粒重	单穗粒重
U_1	14 混甜	1.0000	0.5072	0.8118	0.2143	1.0000	0.0667	0.5993	0.3556
U_2	31-4GS 甜	0.7778	0.6957	0.8706	0.5179	1.0000	0.4667	0.7417	0.4746
U_3	42-1 甜	0.4444	0.4029	0.9647	0.0000	0.6667	0.1333	0.7450	0.0000
U_4	60GS 甜	0.7222	0.0724	0.5529	0.0536	1.0000	0.0000	0.0000	0.3333
U_5	63 甜	0.8333	0.3478	0.8471	0.2321	1.0000	0.1333	0.9305	0.9683
U_6	97-1 甜	0.1667	0.0145	0.0000	0.2321	0.3333	0.4000	0.8874	0.0952
U_7	98-1 甜	0.1111	0.1797	0.4235	0.0000	1.0000	0.4000	0.6987	0.0476
U_8	103-2 甜	0.6667	0.6087	0.8000	0.0179	0.6667	0.5333	0.1987	0.2222
U_9	甜玉米	0.3333	0.8551	0.4000	0.8750	0.0000	0.7333	0.8113	0.5714
U_{10}	油玉米	0.0000	0.7826	0.7647	0.5179	0.3333	0.7067	0.8444	0.5238
U_{11}	冬甜玉米	0.1667	0.0000	0.1529	1.0000	0.3333	0.2200	0.7947	0.9683
U_{12}	超 甜	0.2778	0.3188	0.4471	0.9643	0.3333	1.0000	0.8576	1.0000
U_{13}	甜 粒	0.1667	0.6377	0.8000	0.8036	0.3333	0.5333	1.0000	0.4603
U_{14}	黄 甜	0.1111	1.0000	0.6353	0.6607	0.3333	0.2800	0.9238	0.4286
U_{15}	白 甜	0.1111	0.7971	1.0000	0.9464	0.3333	0.5200	0.8113	0.3016

2.2 建立模糊相似矩阵 R

采用算术平均最小值^[1],计算 R 中各元素 r_{ij} ,公式为:

$$r_{ij} = \frac{\sum_{k=1}^m \min(U_{ik}, U_{jk})}{\frac{1}{2} \sum_{k=1}^m (U_{ik} + U_{jk})} \quad (m=1, 2, \dots, 8)$$

U_{ik} 为样本 U_i 的第 K 个性状。

U_{jk} 为样本 U_j 的第 K 个性状。

根据计算结果,构成相似矩阵 R^*

	U_1	U_3	U_5	U_7	U_9	U_{11}	U_{13}	U_{15}							
U_1	1														
	0.86	1													
U_3	0.76	0.73	1												
	0.75	0.66	0.57	1											
U_5	0.86	0.84	0.74	0.68	1										
	0.45	0.52	0.51	0.27	0.50	1									
U_7	0.65	0.68	0.71	0.59	0.64	0.64	1								
	0.76	0.79	0.75	0.68	0.68	0.42	0.62	1							
U_9	0.54	0.72	0.51	0.33	0.57	0.51	0.49	0.56	1						
	0.63	0.80	0.61	0.37	0.65	0.58	0.57	0.65	0.83	1					
U_{11}	0.46	0.57	0.44	0.33	0.62	0.64	0.48	0.36	0.68	0.63	1				
	0.54	0.67	0.53	0.38	0.68	0.57	0.54	0.53	0.82	0.76	0.82	1			
U_{13}	0.66	0.80	0.64	0.41	0.68	0.62	0.58	0.68	0.79	0.89	0.70	0.82	1		
	0.63	0.75	0.61	0.41	0.65	0.61	0.57	0.60	0.79	0.86	0.67	0.70	0.88	1	
U_{15}	0.63	0.78	0.66	0.39	0.62	0.57	0.57	0.66	0.81	0.87	0.68	0.76	0.90	0.86	1

* R 具有对称性,故只列下三角阵。

2.3 将农业模糊相似矩阵 R 改造成农业等价矩阵

R 通过自乘为 $R^2=R \cdot R$, R^2 自乘为 R^4 ,如此 R^K 自乘,直至 $R^{2^K}=R^K$ 时止, R^K 即为所求的农业模糊等价关系。经过计算 $R^8=R^{2 \times 8}$,则 R^8 为模糊等价矩阵。

	U_1	U_3	U_5	U_7	U_9	U_{11}	U_{13}	U_{15}							
U_1	1														
	0.86	1													
U_3	0.76	0.76	1												
	0.75	0.75	0.75	1											
U_5	0.86	0.86	0.76	0.75	1										
	0.64	0.64	0.64	0.64	0.64	1									
U_7	0.71	0.71	0.71	0.71	0.64	1									
$R^t =$	0.79	0.79	0.76	0.75	0.79	0.64	0.71	1							
U_9	0.80	0.80	0.76	0.75	0.80	0.64	0.71	0.79	1						
	0.80	0.80	0.76	0.75	0.80	0.64	0.71	0.79	0.83	1					
U_{11}	0.80	0.80	0.76	0.75	0.80	0.64	0.71	0.79	0.82	0.82	1				
	0.80	0.80	0.76	0.75	0.80	0.64	0.71	0.79	0.82	0.82	0.82	1			
U_{13}	0.80	0.80	0.76	0.75	0.80	0.64	0.71	0.79	0.83	0.89	0.82	0.82	1		
	0.80	0.80	0.76	0.75	0.80	0.64	0.71	0.79	0.83	0.88	0.82	0.82	0.88	1	
U_{15}	0.80	0.80	0.76	0.75	0.80	0.64	0.71	0.79	0.83	0.89	0.82	0.82	0.90	0.88	1

2.4 取不同的 λ 值,按 λ -截矩阵,进行聚类分析

取 $\lambda=1$,根据 R_λ ,15份甜玉米自交系各为一类。

取 $\lambda=0.90$,聚 $\{U_{13}, U_{15}\}$ 为一类,其它自交系各为一类。

取 $\lambda=0.89$,聚 $\{U_{10}, U_{13}, U_{15}\}$ 为一类,其它自交系各为一类。

取 $\lambda=0.88$,聚 $\{U_{10}, U_{13}, U_{14}, U_{15}\}$ 为一类,其它自交系各为一类。

取 $\lambda=0.86$,聚 $\{U_1, U_2, U_5\}$ 为一类, $\{U_{10}, U_{13}, U_{14}, U_{15}\}$ 为一类,其它自交系各为一类。

取 $\lambda=0.83$,聚 $\{U_1, U_2, U_5\}$ 为一类, $\{U_9, U_{10}, U_{13}, U_{14}, U_{15}\}$ 为一类,其它自交系各为一类。

取 $\lambda=0.82$,聚 $\{U_1, U_2, U_5\}$ 、 $\{U_3\}$ 、 $\{U_4\}$ 、 $\{U_6\}$ 、 $\{U_7\}$ 、 $\{U_8\}$ 各为一类,其它自交系为一类。

取 $\lambda=0.80$,聚 $\{U_3\}$ 、 $\{U_4\}$ 、 $\{U_6\}$ 、 $\{U_7\}$ 、 $\{U_8\}$ 各为一类,其它自交系为一类。

取 $\lambda=0.79$,聚 $\{U_3\}$ 、 $\{U_4\}$ 、 $\{U_6\}$ 、 $\{U_7\}$ 各为一类,其它自交系为一类。

取 $\lambda=0.76$,聚 $\{U_4\}$ 、 $\{U_6\}$ 、 $\{U_7\}$ 各为一类,其它自交系为一类。

取 $\lambda=0.75$,聚 $\{U_6\}$ 、 $\{U_7\}$ 各为一类,其它自交系为一类。

取 $\lambda=0.71$,聚 $\{U_6\}$ 为一类,其它自交系为一类。

取 $\lambda=0.64$,聚15个自交系为一类。

2.5 作聚类分析谱系图

为了直观表达分类结果,作出模糊聚类分析谱系图,见下图。

3 小结及讨论

3.1 自交系亲缘关系的远近和彼此分类差别,可以作为组配杂交种的重要依据^[2]。

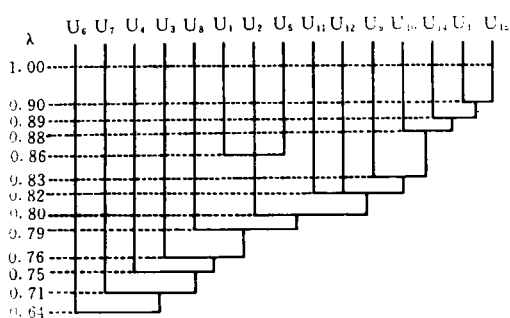
从模糊聚类分析的结果中可以看出 U_{13} (甜粒), U_{15} (白甜), U_{10} (油玉米), U_{14} (黄甜),4个自交系分类系统较近,进行组配杂交种优势会很小。让不同类群间杂交能使其性状互补,同时能丰富自交系的遗传基础,为甜玉米的选育提供科学依据。

3.2 由于模糊聚类分析只采用甜玉米数

量性状作为聚类的依据,没有考虑质量性状,因此子粒颜色不同,也可能分类很近。如: U_{15} (白粒)与 U_{13} (黄粒)分类系统较近。

3.3 在模糊聚类分析中, λ 取值常是解决问题的关键, λ 值不同,分类结果也不同, λ 既即要有一定的专业知识,又要根据等价矩阵中的实际相关系数来确定 λ 值,使其更加符合客观实际。

3.4 在进行聚类分析时,模糊相似矩阵一定要变换为等价矩阵,才能进行聚类,等价矩阵是聚类的基础。



模糊聚类分析谱系图

参 考 文 献

- 1 郭瑞林. 农业模糊学. 河南科学技术出版社, 1991, 107-125
- 2 曹端生等. 玉米自交系聚类分析. 黑龙江农业科学, 1989, (2): 25-28