

改进列联表法在大豆育种上的应用研究

张 琪

(商丘市农林科学院,河南 商丘 476000)

摘要:应用改进列联表法对1990~2004年大豆育种历史资料进行了具体整理和分析。以河南省近年来生产上大面积推广种植的5个大豆品种的株高、单株荚数、单株粒数、百粒重、生育期和单株粒重6个性状作为选择标准,建立了单株选择用表,并在大豆育种上应用,育成了商豆号系列大豆新品种。

关键词:列联表法;线性综合指数;单株选择;大豆

中图分类号:S565.1

文献标识码:A

DOI:10.11861/j.issn.1000-9841.2014.06.0940

The Application of Improved Contingency Table Method in Soybean Breeding

ZHANG Qi

(Shangqiu College of Agricultural Science, Shangqiu 476000, China)

Abstract: The historical data of soybean breeding during 1990-2004 was collated specifically and analyzed by using the improved contingency table method. According to the 6 main traits (plant height, pods per plant, seeds number per plant, 100-seed weight, growth period and grain weight per plant) of the 5 soybean varieties which were large-area planted in recent years in Henan, the individual selection table was established, and applied in the soybean breeding, and bred Shangdou series new soybean varieties.

Key words: Contingency table method; Linear comprehensive index; Plant selection; Soybean

单株选择是建立在优胜劣汰基础上的一种选择方法。在育种选择中,一般采用逐项选择法、独立淘汰水平法和指数选择法。1942年Hazel和Lush在理论上比较了这3种选择方法的相对选择效率,认为指数选择法最为有效^[1]。继Smith在小麦株系选择上使用选择系数后^[2],人们随后对该法在玉米、棉花、大豆等作物上的应用进行了探讨^[3-5],这些研究都极大地丰富了单株选择的理论。然而,指数选择法的试验设计复杂,计算繁琐,且结论的产生往往滞后于育种过程,因此,对指导遗传育种工作有一定的局限性。1990年王福亭等^[6]试用改进列联表法对小麦单株选择进行了探讨,取得了比较满意的效果。但该法在大豆上的应用,至今鲜见报道。因此,本文应用改进列联表法对1990~2004年的大豆育种历史资料进行了具体整理与分析,建立了单株选择用表,试图为大豆早期世代的单株选择寻求一种简单实用的方法,以丰富大豆育种理论与实践。

1 材料与方法

1.1 材料

数据资料取自河南省商丘市农林科学院大豆研究室1990~2004年共115个有代表性的大豆杂

交组合,从 $F_2 \sim F_5$ 代经田间选择及室内考种,共64780个单株。主要研究性状有自变量因子株高(x_1)、单株荚数(x_2)、单株粒数(x_3)、百粒重(x_4)和生育期(x_5)及因变量单株产量(Y)。

1.2 统计原理

首先,将自变量 x 和因变量 Y 按一定标准,分别划分成 a 和 b 个级别(本研究 $a=b=3$)。从育种历史资料中统计自变量 x 和因变量 Y 在各个级别出现的频率($i=1,2,\dots,a; j=1,2,\dots,b$),得到 $a \times b$ 列联表(表1)。

其次,为了判断自变量 x 与因变量 Y 的相关性,通常用列联表进行 χ^2 测验^[7]。

其公式为:

$$\chi^2 = \sum_{i=1}^a \sum_{j=1}^b \frac{n_{ij} - \frac{n_i n_j}{N}}{\frac{n_i n_j}{N}}$$

式中: n_{ij} 为自变量 x 第 j 级在因变量 Y 第 j 级出现的实际频率; $\frac{n_i n_j}{N}$ 为理论频率。因为 $\frac{n_i}{N}$ 为因变量 Y 在第 j 级出现的频率,故理论频率可理解为:当自变量 x 落在第 i 级时,在因变量 Y 第 j 级出现的可能频率。

收稿日期:2014-03-20

基金项目:现代农业产业技术体系建设专项(CARS-004);河南省重点科技计划(2102110013,112102110014)。

作者简介:张琪(1958-),男,副研究员,主要从事大豆育种与栽培技术研究。E-mail:zhangqi8052@163.com。

表 1 a × b 列联表
Table1 a × b contingency table

x 变量 x variable	Y 变量 Y variable				Σ
	1	2	...	b	
1	n ₁₁	n ₁₂	...	n _{1b}	n ₁
2	n ₂₁	n ₂₂	...	n _{2b}	n ₂
:	:	:	...	:	:
a	n _{a1}	n _{a2}	...	n _{ab}	n _a
Σ	n ₁	n ₂	...	n _b	N

其三,计算随机系数 $C = \sqrt{\frac{x^2}{N + x^2}}$, C 值越大,说明自变量 x 与因变量 Y 的相关程度越高,反之亦然。

其四,定义列联比 $\theta_{ij} = \frac{n_{ij}}{n_i} + \frac{n_{ij}}{n_j}$,式中 $\frac{n_{ij}}{n_i}$ 表示自变量 x 为第 i 级时,所选单株在因变量 Y 第 j 级出现的概率; $\frac{n_{ij}}{n_j}$ 表示因变量 Y 第 j 级时,所选单株在自变量 x 第 i 级出现的概率。显然,列联比越大,所选单株在自变量 x 第 i 级因变量 Y 第 j 级出现的可能性越大。因此 θ_{ij} 可作为选择单株的依据。

其五,因各个自变量因子 x 对因变量 Y 的贡献并不相同,故以随机系数 $C_g(g = 1, 2, \cdots, m)$ 作为每一因子列联比的加权值,进行线性综合,得线性综合指数 $P_j = C_1\theta_{1j} + C_2\theta_{2j} + \cdots + C_m\theta_{mj}(j = 1, 2, \cdots,$

$b)$ 。式中, m 表示自变量 x (或随机系数 C)的个数, θ_{ij} 表示对因变量 Y 第 j 级来说,自变量 x_1 落入 i_1 级的列联比, P_j 共含有 m 个加项,与此类推。

最后,在根据考种数据选留单株时,取线性综合指数 P_j 最大的那一级 Y 作为判断单株等级的第一可能性, P_j 级次大的 Y 级作为第二可能性。

2 结果与分析

2.1 大豆单株选择用表的构建

2.1.1 确定选择标准 根据近几年来河南省大面积推广的 5 个大豆品种(豫豆 8 号、豫豆 14、豫豆 16、豫豆 22 和中豆 20)在商丘常规育种条件下的单株各性状表现情况,以 $x \pm 1ax$ 为依据,将单株产量 Y 与诸多因子分为 3 级(表 2)。其中株高(x_1)和生育期(x_5)则是根据多年育种实践经验而拟定。

表 2 各性状选择标准
Table 2 The selection criteria of each character

变量 Variable	级别 Levels		
	1	2	3
单株产量(Y) Yield per plant/g(Y)	>20.3	14.2 ~ 20.3	<14.2
株高(x_1) Plant height(x_1)/cm	75 ≤ x_1 < 85	85 ~ 95	>95 或 <75
单株荚数(x_2) Pods per plant(x_2)	>46.2	30.4 ~ 46.2	<30.4
单株粒数(x_3) Seeds per plant(x_3)	>98.7	61.3 ~ 98.7	<61.3
百粒重(x_4) 100-seed weight(x_4)/g	>19.8	14.8 ~ 19.8	<14.8
生育期(x_5) Growth period(x_5)/d	95 ≤ x_5 < 105	105 ~ 110	>110 d 或 <95

2.1.2 建立各性状列联表 依据上述标准,将田间选择及室内考种 6 480 个单株,按性状建成了 3 × 3 列联表。表 3 为株高 x_1 的列联表,表内左上为频率 n_{ij} ,右下为列联比 θ_{ij} 。同理,可得其余性状列联表(略)。

2.1.3 计算 x^2 与随机系数 C 根据列联表中数字代入相关公式可得表 4。经 x^2 测验,5 个自变量分别达 1% 显著水平,表明单株产量 Y 确实与这些性状有关。其相关程度依次为单株粒数(x_3) > 单株荚数(x_2) > 百粒重(x_4) > 生育期(x_5) > 株高(x_1)。

表3 株高(x_1) 3×3 列联表
Table 3 The 3×3 contingency table of high(x_1)

x 变量 x variable	Y 变量 Y variable			Σ
	1	2	3	
1	737	3125	573	4435
		0.92	1.38	0.80
2	127	868	174	1169
		0.24	1.12	0.35
3	117	650	109	876
		0.25	0.88	
Σ	981	4643	856	6480

表4 各性状 χ^2 测验结果
Table 4 The χ^2 test results of characters

因子统计量 Factor statistics	x_1	x_2	x_3	x_4	x_5
χ^2	28.34 **	1176.32 **	1931.57 **	397.18 **	180.09 **
C	0.07	0.39	0.48	0.24	0.16

$\chi^2_{0.01} = 13.28$; ** 表示达 1% 显著水平。

$\chi^2_{0.01} = 13.28$; ** shows significant difference at 0.01.

2.1.4 构建单株选择用表 用列表中的数据与随机系数 C 计算,即可得加权列联比值。如性状株高(x_1) 在表 5 中, x_1 的 2 级落入 Y 的 2 级中的列联比值为 $0.07 \times 1.12 = 0.08$,其余性状的加权列联比值的计算与此类同,且构成单株选择用表 5。其中表 5 的 P_{Yj} 的值就是其所在列的各个加权列联比值的总和(与 P_j 意义大同小异)。如 $P_{Y1} = 0.06 + 0.02 +$

$0.02 + 0.39 + 0.28 + 0.00 + 0.35 + 0.36 + 0.05 + 0.12 + 0.17 + 0.07 + 0.14 + 0.05 + 0.07 = 2.15$,与此类推。由此可知,各级加权列联比之和 $P_{Y1}:P_{Y2}:P_{Y3}$ 的分布为 $2.15:4.29:2.01 \approx 1:2:1$ 。也就是说,二级单株分布较多,一级和三级的分布较少,大致呈正态分布的趋势。

表5 单株选择用表
Table 5 Plant selection table

因子 Factors	出现级别 Levels	Y 各级的列联比 The contingency ratios of Y		
		Y_1	Y_2	Y_3
x_1	1	0.06	0.09	0.06
	2	0.02	0.08	0.03
	3	0.02	0.06	0.02
x_2	1	0.39	0.34	0.00
	2	0.28	0.55	0.35
	3	0.00	0.32	0.22
x_3	1	0.35	0.34	0.00
	2	0.36	0.76	0.22
	3	0.05	0.26	0.54
x_4	1	0.12	0.27	0.03
	2	0.17	0.31	0.16
	3	0.07	0.41	0.16
x_5	1	0.14	0.21	0.12
	2	0.05	0.16	0.06
	3	0.07	0.13	0.04
$P_{Yi}(\Sigma)$		2.15	4.29	2.01

2.1.5 单株的筛选与取舍 在选留单株时,可根据所在级别查阅表 5,并进一步求出线性综合指数 P_j ,然后以 P_2 为基准,当 $P_1 > P_3$ 时,若 $P_1/P_2 > 0.5$,则

该单株为 1 级;当 $P_1 < P_3$ 时,若 $P_3/P_2 > 0.5$,则该单株为 3 级,否则便为 2 级单株。1、2 级单株保留,3 级单株淘汰。

2.2 在大豆育种上的具体应用

1997 年,我们配制了杂交组合商 85225/郑 92029,以后各世代选留了一部分单株,以系谱号 9708-3-15 第 8 株所形成的品系为例,说明该法的具体应用(表 6)。

该品系株高(x_1)为 76.8 cm,单株荚数(x_2)为 43.6 个,单株粒数(x_3)85.1 个,百粒重(x_4)21.5 g,全生育期(x_5)104 d。据表 6 各性状的选择标准可知,它们分别属于 1 级、2 级、2 级、1 级、1 级。根据

各性状级别,查单株选择用表 5 中的加权列联比值,将这些列联比值求和,即为线性综合指数 P_j 。本例中: $P_1 = C_1\theta_{11j} + C_2\theta_{22j} + \cdots + C_5\theta_{55j} = 0.06 + 0.28 + 0.36 + 0.12 + 0.14 = 0.96$,同理 $P_2 = 1.88, P_3 = 0.78$ 。因 $P_1 > P_3, P_1/P_2 = 0.96/1.88 = 0.51 > 0.5$,所以该单株形成的品系平均产量为 $2\ 761.8\text{ kg}\cdot\text{hm}^{-2}$,比对照种豫豆 16($2\ 456.5\text{ kg}\cdot\text{hm}^{-2}$)提高 12.4%,其他性状如抗倒伏和抗花叶病毒病等均优于豫豆 16,说明该法应用于育种实践是可行的。

表 6 改进列联表应用示例
Table 6 Application examples of improved contingency table

因子 Factors	系谱号 The pedigree number			品系 Strains
	9708-3	9708-3-15	9708-3-15-8	
株高 Plant height(x_1)	78.9	82.1	78.5	76.8
单株荚数 Pods per plant(x_2)(x_2)	44.8	43.3	42.8	43.6
单株粒数 Kernels per plant(x_3)	91.6	86.1	83.8	85.1
百粒重 100-seed weight(x_4)	22.6	22.3	21.8	21.5
生育期 Growth period(x_5)	105	103	104	104
P_1	0.96	0.96	0.96	0.96
P_2	1.88	1.88	1.88	1.88
P_3	0.78	0.78	0.78	0.78
单株等级 Plant level	1	1	1	1
世代 Generation	F ₂	F ₃	F ₄	F ₅

为进一步证实改进列联表法对大豆育种的指导意义,又采用最近两年在河南省推广面积较大的商豆 951099 进行验证(表 7)。

该品种系谱号为 9030-2-2-6,由河南省商丘市农林科学院选育而成,2000 和 2001 年参加国家黄淮海夏大豆区试(南片)及生产试验,平均产量为

$2\ 927.5\text{ kg}\cdot\text{hm}^{-2}$,比对照种中豆 20($2\ 522\text{ kg}\cdot\text{hm}^{-2}$)增产 13.82%,并且籽粒商品性好,耐肥抗倒,抗花叶病毒病,适应性广。于 2003 年通过国家农作物品种审定委员会审定。这一实例进一步验证了该法对大豆育种的指导意义。

表 7 商豆 95199 印证示例
Table 7 The application examples of Shangdou 951099

因子 Factors	系谱号 The pedigree number					
	9030-2	9030-12	9030-2-2	9030-2-10	9030-2-2-6	9030-2-2-3
x_1	83.2	81.8	80.5	78.7	78.5	82.6
x_2	61.7	58.9	56.8	52.6	54.6	55.9
x_3	138.2	116.2	123.9	109.5	119.7	123.2
x_4	16.3	15.6	15.3	15.1	15.6	15.2
x_5	105	105	103	104	104	105
P_1	1.11		1.11		1.11	
P_2	1.25		1.25		1.25	
P_3	0.34		0.34		0.34	
单株等级 Plant level	1	1	1	1	1	1
世代 Generation	F ₂		F ₃		F ₄	

3 结论与讨论

改进列联表法克服了以往仅凭直观印象选种的随意性,使单株选择有了量化标准。由于这种方法是建立在多年育种历史资料的基础上,其置信度

较大,判断结果更加客观。与综合选择指数法相比,其原理易懂,计算简便,易于育种工作者掌握和运用。应用单株选择用表,我们选育出了国审商豆 1099、商豆 6 号和省审商豆 14 等大豆新品种,证明该法可作为一种新的大豆单株选择的量化方法指

导大豆育种实践。

构建大豆单株选择用表,选择标准的合理确定是十分重要的。要进一步培育出优于现有推广品种的新品种,选择标准太低或太高,都会使育种工作失之偏颇,最终不能付诸实现。本文所采用的标准,根据河南省近年来大面积推广的 5 个大豆品种,在本地常规育种条件下单株性状表现的情况而确定,这种选择标准的确定比较合理。另外,在性状的选取上,由于人们在田间选择时对抗病性已足够注意,因此,在最后取舍单株时,这些性状可以不再考虑。

由于大豆数量性状受环境影响较大,本文所建单株选择用表并不是各生态区都适用。育种工作者可根据当地试验的历史资料,计算出适合本地气候条件及其他生态环境的单株选择用表,以增加选择的准确性。

最后需要指出的是,大豆不同世代诸性状的遗传力不尽相同,因而选择策略也应有所变动。由于本文在建立改进列联表时没有考虑不同世代遗传力对选择效果的影响,这一问题有待进一步研究。

(上接第 939)

细胞坏死使 LDH 释放入组织液中,通过测定 LDH 可以帮助了解组织器官受损状况^[10],与空白对照组相比,PFOA 低、中、高剂量组的 LDH 的活性均表现出不同程度的上升,说明 PFOA 可能加速了细胞坏死,而 GEN 干预 PFOA 低、中、高剂量组 LDH 的活性均明显低于 PFOA 低、中、高剂量组,这在一定程度上可能说明 GEN 可以在一定程度上减少细胞的坏死。在本试验中,MDA 和 NO 的含量虽然在一定程度上随着 PFOA 的剂量加大有着不同程度的上升,但 MDA 和 NO 的含量在同剂量的 PFOA 组和 GEN 干预组中没有明显的差异性,可能是 GEN 的干预浓度比较低,还不足以引起氧化损伤的本质性变化。这个原因可能需要更进一步地进行探讨,其具体机制应经深入研究才能得到证实。

参考文献

- [1] Luccisano A E, Campbell J L, Butenhoff J L, et al. Evaluation of placental and lactational pharmacokinetics of pfoa and pfos in the pregnant, lactating, fetal and neonatal rat using a physiologically based pharmacokinetic model[J]. *Reproductive Toxicology*, 2012, 33(3):468-490.
- [2] Harada K, Nakasanishi S, Sasaki K, et al. Particle size distribution and respiratory deposition estimates of airborne perfluorooctanoate and perfluorooctanesulfonate in Kyoto area, Japan. *Bull*[J]. *Environmental Contamination and Toxicology*, 2006, 76(3):306-310.
- [3] 金一和, 丁梅, 翟成, 等. 长江三峡库区江水和武汉地区地面水中 PFOS 和 PFOA 污染现状调查[J]. *生态环境*, 2006, 15(3):

参考文献

- [1] Hazel L N, Lush J L. The efficiency of three methods of selection [J]. *Journal of Heredity*, 1942, 33:393-399.
- [2] Smith H F. A discriminant function for plant selection [J]. *Annals of Eugenics*, 1936, 7:240-250.
- [3] Robinson H F, Comstock R E, Harvey P H. Genotypic and phenotypic correlation in corn and their implication in election [J]. *Agronomy Journal*, 1951, 43:282-287.
- [4] Manning H L. Yield improvement from selection index technique with cotton [J]. *Heredity* 1956, 10:303-322.
- [5] 马育华, 盖钧镒. 江淮下游地区大豆地方品种的初步研究 (三) 数量性状的表型、遗传型相关, 选择指数, 及其育种意义 [J]. *作物学报*, 1979(4):1-12. (Ma Y H, Gai J Y. Preliminary study on the local soybean varieties in lower Yangtze and Huai valleys III. Phenotypic and genotypic correlations, selection indices, and their implications in soybean breeding [J]. *Acta Agronomica Sinica*, 1979(4):1-12.)
- [6] 王福亭, 郭瑞林. 改进列联表法在小麦育种上的应用研究 [J]. *作物学报*, 1991, 17(1):11-17. (Wang F T, Guo R L. Application of improved contingency table method in maize breeding [J]. *Acta Agronomica Sinica*, 1991, 17(1):11-17.)
- [7] 马育华. 试验统计 [M]. 北京: 中国农业出版社, 1982:320-338. (Ma R H. *Experimental statistics* [M]. Beijing: China Agriculture Press, 1982:320-338.)
- [8] 486-489. (Jin Y H, Ding M, Zhai C, et al. An investigation of the PFOS and PFOA pollution in Three Gorges Reservoir areas of the Yangtze River and surface water of Wuhan areas [J]. *Ecology and Environment*, 2006, 15(3):486-489.)
- [4] Valachovicova T, Slivova V, Sliva D. Cellular and physiological effects of soyflavonoids [J]. *Mini-Review in Medicinal Chemistry*, 2004, 4(8):881-887.
- [5] 陶利, 李禾. 大豆异黄酮的药理作用 [J]. *解放军药学报*, 2011, 27(4):360-362. (Tao L, Li H. Pharmacological action of soybean isoflavone [J]. *Pharmaceutical Journal of Chinese People's Liberation Army*, 2011, 27(4):360-362.)
- [6] Klaunig J E, Babich M A, Baetcke K P, et al. PPAR α agonist-induced rodent tumors: Modes of action and human relevance [J]. *Critical Reviews in Toxicology*, 2003, 33(6):655-780.
- [7] Lau C, Anitole K, Hodes C, et al. Per fluoroalkyl acids: A review of monitoring and toxicological findings [J]. *Toxicological Sciences*, 2007, 99(2):366-394.
- [8] Atlante A, Bobba A, Marra E, et al. Transport and metabolism of L-lactate occur in mitochondria from cerebellar granule cells and are modified in cells undergoing low potassium dependent apoptosis [J]. *Biochimica et Biophysica Acta*, 2007, 176(7):1285-1299.
- [9] 张粟, 冉云, 王玲, 等. 茶多酚对乙醇所致原代肝细胞损伤的保护作用研究 [J]. *营养学报*, 2009, 31(4):379-383. (Zhang L, Ran Y, Wang L, et al. Study of protective effects on ethanol induced primary hepatocytes damage by tea polyphenol [J]. *Acta Nutrimenta Sinica*, 2009, 31(4):379-383.)
- [10] 范远景, 张玲, 张东吟, 等. 三羟异黄酮保护急性酒精性肝损伤机制研究 [J]. *营养学报*, 2013, 35(3):273-282. (Fan Y J, Zhang L, Zhang D Y, et al. Study on the mechanism of genistein on protection against alcohol-induced acute liver injury [J]. *Acta Nutrimenta Sinica*, 2013, 35(3):273-282.