

大豆种子氨基酸组分变异分析

林忠平 彭光子

尹光初

(中国科学院植物研究所)

(黑龙江省农业科学院大豆研究所)

摘 要

对国内外51个大豆品种(系)进行了蛋白质含量和氨基酸组分的测定。并借助于微型电子计算机,按概率统计中计算线性相关系数的方法和灰色系统中计算灰色关联度和判断关联极性的方法加以分析。结果表明:国内各地大豆蛋白变异范围大,改良品质的潜力很大。氨基酸线性相关系数分析表明:我国各地大豆品种中氨基酸含量间变化呈负相关的比呈正相关的多得多,而来自北美的大豆中呈正相关的多于负相关的。两者差异的实质值得研究。本文还讨论了两种含硫氨基酸在变异中的特殊性,以及基于柱层分析测定含硫氨基酸的方法在精度和重复性上有缺陷可能影响分析的结果。

大豆富含蛋白质(约含40%),高蛋白的品系种子蛋白含量可高达55%。而且它的赖氨酸含量较高,可与禾谷类成营养互补(后者赖氨酸含量低)。所以大豆蛋白在改善食品营养中的应用引起了广泛的重视^[1, 2]。然而大豆蛋白中甲硫氨酸和半胱氨酸明显不足,又成为大豆营养的限制因子。因而研究大豆品种资源中种子氨基酸组分是大豆品质改良的重要课题^[3]。

我们对国内外51个大豆品种或品系进行了蛋白质含量和氨基酸组分的测定。并利用概率统计中计算线性相关系数的方法和灰色系统中计算灰色关联度的方法加以分析^[4-8]。

一、材料与方法

来自我国东北、华北、华东、华中、西北和西南各主要大豆产区的16个商品大豆种子由商业部西安油脂化学研究所提供。黑龙江省三江地区的26个大豆品系的种子来自国营八五〇农场和宝清县东明村试验基地。另外有来自北美的9个商品大豆品种由黑龙江省外贸局提供。

粗蛋白含量的测定按微量凯氏法进行。用于氨基酸组分分析的大豆磨成细粉后,用正戊烷脱脂,6NHCl 120℃水解21小时。在日立835型氨基酸分析仪上测定。

数据处理用自编的Fortran语言程序,在ACS-8600高档微型机上进行。

本文于1986年6月19日收到。This paper was received in June 19, 1986.

二、蛋白质含量和氨基酸组分测定结果

蛋白质含量和氨基酸组分测定结果的平均值见表1。

表1 氨基酸含量的平均值、标准差和变异系数

Table 1 Mean value standard deviation and coefficient of variability
of amino acid contents

		A			B			C			D		
		m	s.d	v.c%	m	s.d	v.c%	m	s.d	v.c%	m	s.d	v.c%
天	Asp.	11.81	.2864	2.38	11.70	.4065	3.47	11.55	.2415	2.04	11.60	.3577	3.01
苏	Thr.	3.67	.1172	3.13	3.95	.2571	6.51	3.73	.2195	5.77	3.78	.2095	5.48
丝	Ser.	4.08	.2951	7.08	5.01	.1285	2.65	4.15	.3965	9.42	4.41	.4765	10.73
谷	Glu.	22.61	.5417	2.35	21.47	.5665	2.61	22.42	.4475	1.95	22.16	.8244	3.64
甘	Gly.	4.26	.1535	3.54	4.25	.1057	2.49	4.23	.1189	2.77	4.24	.1399	3.25
丙	Ala.	4.25	.1757	4.06	4.28	.1108	2.59	4.08	.2907	6.83	4.20	.1886	4.38
半胱	Cys.	0.67	.1495	21.86	1.14	.3632	31.47	.82	.1670	19.80	0.87	.3145	36.72
缬	Val.	5.08	.1451	2.80	4.77	.1302	2.73	5.10	.1558	3.01	4.98	.2389	4.73
蛋	Met.	0.86	.1736	19.83	.90	.1954	21.78	.85	.1326	15.72	0.87	.1754	20.01
异亮	Ile.	4.92	.1787	3.56	4.58	.1193	2.60	5.03	.2342	4.61	4.84	.2741	5.60
亮	Leu.	7.57	.3342	4.36	7.90	.2860	3.67	7.63	.2041	2.64	7.70	.3050	3.95
酪	Tyr.	3.07	.2113	6.68	3.48	.2277	6.54	3.10	.1525	4.79	3.21	.02524	7.73
苯丙	Phe.	5.82	.3374	5.70	5.71	.2106	3.68	5.69	.2976	5.15	5.74	.3099	5.13
赖	Lys.	6.27	.2912	4.54	6.24	.0862	1.06	6.36	.2009	3.13	6.29	.2415	3.80
组	His.	2.01	.1769	8.32	2.39	.1467	6.14	2.15	.3269	14.87	2.18	.2340	10.53
精	Arg.	7.30	.3625	4.86	7.91	.5655	7.18	6.99	.1881	2.63	7.40	.4906	6.51
脯	Pro.	3.76	.4676	12.08	4.17	.1559	3.74	4.01	.5314	12.82	3.98	.4360	10.86

A: 三江地区26个品系, B: 我国各地16个品种, C: 北美9个品种, D: 全部15个品种或品系。

A: 26 lines from Sanjiang, B: 16 strains from different areas of China,

C: 9 strains from North America, D: all of 15 strains or lines.

三、数学分析

为了表述方便,我们把原始数据定义为矩阵元素 X_{ij} (i 表示氨基酸, j 表示大豆的品种)。

1. 标准差和变异系数的分析(4) 为了研究各种氨基酸含量在各品系之间变异的幅度,我们计算它们的标准差,见表1中s.d栏。

另外,三江地区26个大豆品种种子粗蛋白含量变量的标准差为3.345;我国各地16

个品种为 2.201; 国外 9 个品种为 1.49。

为了研究各种氨基酸变异的相对量, 比较它们在遗传上的稳定程度, 我们计算了各种氨基酸的变异系数 s/\bar{x} (s 为标准差, \bar{x} 为变异的算术平均值), 见表 1 中 c.v% 栏。

2. 线性相关系数分析^[5] 为了研究各氨基酸组分变化时, 彼此之间有无关联, 我们先分别计算了三江地区 26 个品种、我国各地 16 个品种和国外 9 个品种的线性相关系数。然后, 将这 51 个品种总合在一起计算相关系数, 见表 2。

表 2 51 个大豆品系氨基酸含量变异相关系数矩阵

Table 2 Correlation coefficient between amino acid contents of 51 soybean lines

[Asp] 1.000	$\alpha=0.05$ 时, $ ric > 0.273$ 有意义														
[Thr] -.191 1.000	$\alpha=0.01$ 时, $ ric > 0.354$ 有意义														
[Ser] -.315 .526 1.000															
[Glu] .378 -.464 -.647 1.000															
[Gly] .440 .063 -0.78 .082 1.000															
[Ala] -.326 .020 -.080 -.216 .139 1.000															
[Cys] -.749 .282 .499 -.406 -.481 -.026 1.000															
[Val] .379 -.318 -.737 .396 .517 .288 -.669 1.000															
[Met] -.426 -.105 -.087 -.218 -.090 .430 .188 .118 1.000															
[Ile] .301 -.381 -.622 .387 .489 .279 -.633 .890 .078 1.000															
[Leu] -.019 .221 .212 -.354 .388 .317 .051 .124 .053 .133 1.000															
[Tyr] -.273 .311 .454 -.290 -.292 -.212 .476 -.653 -.098 -.653 .036 1.000															
[Phe] .101 -.258 -.359 .267 .093 .030 -.229 .264 -.061 .195 -.125 -.165 1.000															
[Lys] -.222 .005 -.066 .057 .243 .431 -.139 .406 .245 .507 .271 -.379 -.095 1.000															
[His] -.644 .282 .310 -.536 -.540 .219 .509 -.385 .381 -.447 .073 .399 -.203 -.022 1.000															
[Arg] .272 .214 .412 -.358 -.039 -.366 -.056 -.325 -.099 -.369 -.130 .112 -.253 -.263 .029 1.000															
[Pro] -.328 .104 -.010 -.218 -.527 -.119 .492 -.324 .095 -.416 -.128 .401 -.077 -.509 .490 -.186 1.000															
[Asp][Thr][Ser][Glu][Gly][Ala][Cys][Val][Met][Ile][Leu][Tyr][Phe][Lys][His][Arg][Pro]															

3. 灰色关联度^[6]和聚类分析大豆种子蛋白质的组分及其氨基酸构成受多种因素影响^[3], 而且氨基酸分析费用颇高, 在取样上很难达到充分的代表性, 因此存在着参数信息和关系信息不完全的情况。在这种情况下采用信息部分明确和部分不明确的灰色控制系统的方法来进行分析, 应是比较合理的^[6]。由于品种、环境、地域等方面的差异, 每种氨基酸含量变化的波动情况是各不相同的。有些氨基酸变动情况比较近似, 有些氨基酸变动情况相去甚远。就变异的近似程度而言, 17 种氨基酸中每一种与其它 16 种氨基酸都有一个相关程度大小的问题。我们用关联度的大小表示不同氨基酸之间在变异中的近似程度。

$$\bar{E}_{ij} = \frac{1}{n} \sum_{i=1}^n E_{ij}(I)$$

E_{ij} 称为第 i 个氨基酸与第 j 个氨基酸的关联度。 $E_{ij}(l)$ 称为第 i 个氨基酸与第 j 个氨基酸在 l 点的关联系数。 $i=1, 2, \dots, m; j=1, 2, \dots, m$ 。 n 为矩阵的列数, m 为矩阵的行数。这里,

$$E_{ij}(l) = \frac{\Delta(\min) + \sigma}{\Delta_{ij}(l) + \sigma}; \Delta(\min) = \min_j \{ \Delta_{ij}(\min) \}$$
$$\Delta_{ij}(\min) = \min_j \{ \Delta_{ij}(l) \}, \sigma = \zeta \cdot \Delta(\max)$$
$$\Delta(\max) = \max_j \{ \Delta_{ij}(\max) \}; \Delta_{ij}(\max) = \max_j \{ \Delta_{ij}(l) \}$$
$$\Delta_{ij}(l) = |x_i(l) - x_j(l)|; x_i(l) = x_i l / \bar{x}_i; x_i(l) = x_i l / \bar{x}_i$$

$\bar{x}_i = \frac{1}{n} \sum_{l=1}^n x_{il}, \bar{x}_j = \frac{1}{n} \sum_{l=1}^n x_{jl}$, 计算中分辨系数 ζ 取为 0.1〔7〕。

微型机给出的关联度矩阵 (relation grade matrix) 中每行的数值表示某种氨基酸与其它氨基酸变异的近似程度, 即关联性的强弱, 为了使这一种表示方法一目了然, 按关联度的大小加以排列, 得到排序表 (sorting table), 表中的数字代表本文中各种氨基酸的顺序编号, 如 1 代表天门冬氨酸 (Asp), 2 代表苏氨酸 (Thr)18 代表脯氨酸 (Pro)。

表 3 51 个大豆制品系中氨基酸关联度排序表
Table 3 Amino acid relational grades sorting table of 51 soybean 3 varieties.

[1]	5	11	14	4	10	2	16	12	8	6	15	13	17	9	3	7
[2]	5	11	16	1	14	12	4	10	15	8	6	13	3	9	7	17
[3]	16	2	12	1	5	14	11	10	4	15	8	7	6	13	9	17
[4]	8	10	1	14	11	6	5	13	2	12	15	17	16	9	3	7
[5]	11	1	14	10	2	4	6	8	16	12	13	15	9	17	3	7
[6]	8	4	10	13	11	14	5	1	15	17	2	16	9	12	7	3
[7]	6	2	8	10	15	12	14	11	4	13	17	5	1	16	3	9
[8]	10	4	6	13	11	14	5	1	2	15	12	17	16	9	7	3
[9]	6	14	15	8	10	11	4	5	13	1	17	2	16	12	3	3
[10]	8	4	14	11	5	6	1	13	2	16	12	15	17	9	3	7
[11]	5	14	1	10	4	2	8	6	12	16	15	13	17	9	17	7
[12]	1	2	14	11	5	16	4	15	10	8	13	6	3	9	7	7
[13]	8	4	6	10	14	11	1	5	17	15	12	2	9	16	7	3
[14]	11	5	1	10	4	2	6	8	12	15	16	13	9	17	3	7
[15]	6	14	1	11	2	4	12	5	8	13	10	16	17	9	7	3
[16]	2	1	5	11	12	14	10	4	15	3	8	6	13	9	17	7
[17]	6	4	13	8	15	10	1	11	14	5	9	2	12	16	7	3

1. (Asp) 2. (Thr) 3. (Ser) 4. (Glu) 5. (Gly) 6. (Ala) 7. (Cys) 8. (Val) 9. (Met) 10. (Ile)
11. (Leu) 12. (Tyr) 13. (Phe) 14. (Lys) 15. (His) 16. (Arg) 17. (Pro)

进一步研究排序表和关联度矩阵, 可以看到每种氨基酸与其它各种氨基酸的关联度大小差异的不同, 即有些关联度十分接近, 有些相差甚远。基于这种认识, 我们根据聚类分析的思想〔8〕, 在保持排序表中次序不变的前提下, 对关联度进行分类 (组)。这样就

便于进行各序列关联度的比较。因此，设同一序列中关联度间的“距离”为 d_i 。

$$d_i = \frac{1}{m-2} \sum_{j=1}^m (\bar{E}_{ij} - \bar{E}_{i+1,j}), (j, j+1 \neq i),$$
 这里 d_i 为第 i 个氨基酸序列中关联度

间的距离； m 为原始数据矩阵的行数， E_{ij} 为关联度矩阵中第 i, j 个元素。认定经过排序的表中第 i 个关联度序列中相邻两个关联度之差值小于 d_i ，那么它们就是一类（组）的。由此我们得到一系列分组表（grouping table）。

我们仍然是先分别对三江地区26个品种、国内16个品种和国外9个品种进行关联度分析，而后再将这51

个品种总合起来进行关联度分析，见表4。表4表示三江地区的分组表中天门冬氨酸与其它氨基酸之间的关联度可分为7个等级。

表4 天冬氨酸（Asp）与其它氨基酸关联度分组表
Table 4 Grouping relational grades of Asp and other amino acids

组 别 No	各 组 成 员 Grouping amino acids
1	甘 Gly
2	亮、赖、谷 Leu Lys Glu
3	异亮、苏、精、酪、缬、丙 Ile Thr Arg Trp Val Ala
4	组、苯丙 His Phe
5	脯 Pro
6	甲硫、丝 Met Ser
7	半胱 Cys

说明：天冬氨酸与各种氨基酸的关联度可分为7个等级（组）第一组成员与天冬氨酸关联度最大，第二组次之第三组再次之……。

Note: The relational grades of Asp with other amino acids can be divided into 7 grades (or groups) among them the relational grade of Asp with Gly is the largest, the relational grades between Asp and members in second group (Leu, Lys, and Glu) ranks second……etc.

四、讨论

1. 利用计算标准差的方法考察了种子总蛋白含量

在品种间变动的幅度。来自北美的9种商品大豆蛋白质含量差异很小，我国各地商品大豆蛋白含量差异甚大。说明通过育种程序提高我国大豆品种的蛋白含量，改善品质的潜力是很大的。

2. 表1的资料表明三江、国内和国外三部分数据中氨基酸组分的变化是不一样的。但是谷氨酸含量，其次是精氨酸和脯氨酸含量变化较大则是带有普遍性的。就氨基酸相对含量变化而言（表1），半胱氨酸和甲硫氨酸这两种含硫氨基酸含量的变异是最大的。而这两种氨基酸恰恰是大豆营养的限制性氨基酸。许多作者^[9, 10]认为以离子交换层析为基础的测定含硫氨基酸的方法，包括氨基酸分析仪的测定不但费时、昂贵，而且重复性差。因此在涉及含硫氨基酸的分析不能不考虑到测定方法上存在的问题。

3. 各种氨基酸之间相关性的研究表明下列凡对氨基酸有相当显著的负相关（ $\alpha=0.01$ ）：半胱一天门，谷一丝，缬一丝，异亮一丝，组一谷，组一甘，缬一半胱，异亮一半胱，酪一缬，酪一异亮，脯一赖。下列凡对氨基酸之间有相当显著的正相关系数：丝一苏，缬一甘，组一半胱，异亮一缬。但是如果分别考虑（在 $\alpha=0.01$ ）：“三江”、“国内”和“国外”三部分数据时，我们发现这三部分的大豆，其氨基酸组分的变异是有不少差别的：三江地区与全国各地区大豆相比时，半胱一天门，赖一天门的负相关是共同的。三江地

区与北美大豆相比时,半胱一甘的负相关是共同的。我国各地大豆与北美大豆相比时,酪一异亮的负相关,脯一半胱的正相关是共同的。只是当在稍低置信水平($\alpha=0.05$)考察时,这三部分彼此之间才有较多的共同之处。这可能与三部分大豆遗传背景上的差异有关。此外还应注意北美9个商品大豆中呈正相关的氨基酸(5对)多于负相关的氨基酸(3对);而国内各地16种大豆中呈负相关的氨基酸多达16对,正相关的为5对;三江地区呈负相关的(8对)也略多于呈正相关的(6对)。这是否反映了我国大豆与北美大豆在栽培条件、代谢水平上的差别是值得进一步研究的。

我们用上述方法考察了各种氨基酸含量与粗蛋白含量之间的关系,没有看到显著的相关关系。我们在另外的研究中发现除去种皮之后大豆籽粒中蛋白质的总量与其蛋白质的组分是有关系的^[11]。因此,如果摒弃了种皮对粗蛋白百分含量的影响,那么粗蛋白与氨基酸组分之间可能也存在一定的相关性。

已经知道不同的氨基酸其含量变化的幅度是不同的。那么,就这种变化而言,那些氨基酸变化比较相似,哪些氨基酸的变异情况差别大呢?灰色系统关联度的分析提供了这方面的资料。总的来说,甘一天门,甘一亮,缬一异亮,丙一缬,赖一亮,都是比较相似的。在三江、国内和国外三个地区情况又稍有差别。三个地区的数据分析中有一个比较一致的情况是各种氨基酸与甲硫氨酸及半胱氨酸的关联度都很小,这可能反映了甲硫氨酸和半胱氨酸在植物代谢上的特殊性^[12],也可能反映了对这两种含硫氨基酸的测定方法还存在问题,这是值得进一步研究的。

关于两种氨基酸之间相关关系的分析,我们还利用灰色系统中关联极性的判断来验证上面得到的结果^[13],所得的结果是一致的。

Mosse' 和 Pernollet^[14] 对豆科植物种子氨基酸组分的大量研究资料作了综述和概括。认为豆科植物种子氨基酸组分的变异主要受遗传因素的影响。由于未用计算机处理数据没有得到各种氨基酸之间相互关系的结果,本研究正是在这方面作了初步的研究。

参 考 文 献

- [1] 林忠平、尹光初, 1983, 大豆科学, 2, 232—238.
- [2] 林忠平, 1986, 中国油脂, No1, P52.
- [3] 林忠平、付燕凤、尹光初, 1985, 大豆科学, 4, 327—336.
- [4] 中国科学院计算中心概率统计组编著, 概率统计计算, 1983, 科学出版社.
- [5] 冯士雍, 1985, 回归分析, 科学出版社.
- [6] 邓聚龙, 1985, 灰色控制系统, 华中工学院出版社.
- [7] 郭 洪, 1985, 灰色系统关联度分辨系数, 模糊数学, 5(2), 55.
- [8] 方开泰等, 1980, 聚类分析, 地质出版社.
- [9] Radford R. L., Chavengsaksong Kram C. and Hymowitz, 1977, Crop Sci, 17, 273—277.
- [10] Maranville C. J., Mattern P. J. and Clark, R. B., 1984, Crop Sci. 24, 303—308.
- [11] 林忠平、尹光初、雷勃钧, 1985, 科学通报, 30, 540—543.
- [12] Giovanelli, J., Mudd, S. H. and Datko A. H., 1980, The Biochemistry of Plant. Vol. 5, Amino and Derivatives. ed. by B. J. Mifflin. P453—505, Academic Press.
- [13] 邓聚龙, 1980, 灰色系统——社会与经济, 国防出版社, .
- [14] Mosse' and Pernollet J. C. 1983. Chemistry and Biochemistry of Legumes ed. by S. K. Arora P111—193 Edward Arnold.

ANALYSIS OF AMINO ACID COMPONENT VARIATION OF SOYBEAN PROTEIN

Lin Zhongping, Peng Guangzi

(Institute of Botany, Academia Sinica, Beijing)

Yin Guanchu

(Soybean Institute, Heilongjiang Academy of Agricultural Sciences, Harbin)

Abstract

Protein content and amino components of 51 local and exotic soybean varieties or lines were determined. Data was analysed by linear correlation coefficient methods and degree of grey relation and relational polarity methods in grey system with the help of microcomputer. The results show: local soybean varieties from different places have a wide protein content range, thus the potential of improving soybean quality is very high in breeding programs. The linear correlation coefficient of amino acid content shows that negative relationship between two amino acids is much greater than positive relationship among varieties from different areas of China, and the result is just opposite in North America soybean varieties. Therefore the nature of these differences is worth while for further study. The speciality variation of two kinds of sulfur amino acids, and the problem of precision and replication of S-amino acid assay based on the chromatographic column has also been discussed.