

Principal Component Clustering Analysis Apply to the Amino Acid Content in Antler Based on Matlab

Xin Zhao, Jing Wang, Yang Jiao, Haijiao Yu, Yao Qin, Quanming Li,
Guogang Zhao

► **To cite this version:**

Xin Zhao, Jing Wang, Yang Jiao, Haijiao Yu, Yao Qin, et al.. Principal Component Clustering Analysis Apply to the Amino Acid Content in Antler Based on Matlab. 10th International Conference on Computer and Computing Technologies in Agriculture (CCTA), Oct 2016, Dongying, China. pp.284-290, 10.1007/978-3-030-06155-5_28 . hal-02179981

HAL Id: hal-02179981

<https://hal.inria.fr/hal-02179981>

Submitted on 12 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Principal Component Clustering Analysis Apply to the Amino Acid Content in Antler Based on Matlab

Xin Zhao¹, Jing Wang², Yang Jiao³, Haijiao Yu⁴, Yao Qin⁴, Quanming Li⁴, Guogang Zhao⁴(✉)

¹School of Mathematics and Big Data, Huizhou University, Huizhou, China, 516007 China;
zhaoxin@hzu.edu.cn

²College of Information Technology, Jilin Agricultural University, Changchun, 130118, China;
374121414@qq.com

³Sixth Middle School in Changchun, Changchun, 130000, China;
61516131@qq.com

⁴The City College of JLU, Changchun, 130000, China;

haihai951kl@163.com, 470695861@qq.com, keeper007@126.com, zhaogg2010@foxmail.com,

Abstract. The antler is a very high nutritional value of supplements, but also a very important medicine, antler occupies a very important position in Chinese medicine. Antler is rich in amino acids, which contains more than seven kinds of essential amino acids. In the organic component of antler, amino acids are the topped content of nutrients, and in them, the highest is glycine; amino acids are the basic components of living organism tissue cells, and play a pivotal role for life events. If the body lacks any kind of essential amino acids, it can cause physiological dysfunction, affecting the normal antibody metabolism, leading to disease. In the paper, it accord to the data of amino acids which contain different specifications antler herbs, analyze and compare the relationship of amino acids between sika deer antler and red deer antler using the principal component cluster analysis. The results showed that sika deer antler with red deer antler have the similar medicine effect and different essential amino acids nutrients.

Keywords: amino · Amino acid · principal component cluster analysis · Cluster analysis method · Matlab

1 Introduction

Antler is not ossification of sika deer and red deer bucks dense hairs, and it is common health tonic medicine, it mainly has Warming impotence, Yijing blood, gluten bone health and other effects^[1-5]. Antler has been used as medicine two thousand years, and is recorded in Han Dynasty "Shen Nong's Herbal Classic"^[6-8]. So far, antler medicinal ingredient is still not very clear, in the works of process, quality assessment, identifying the authenticity, there is no scientific chemical composition as the standard indicators, and it restricts further development and utilization of the antler^[9-11]. Therefore, it is in urgent need to carry in-depth study, it should systematically study the antler active ingredients and pharmacological effects using modern analytical methods, this work not only has theoretical significance, but also important practical value, which can found the clear antler quality standars, and provide a scientific basis for the further development and utilization of the antler. This paper analyze and compare the relationship of amino acids between sika deer antler and red deer antler using the principal component cluster analysis.

2 The Data Extraction and Related Work

In this paper, the data is from the article by Zhang Song^[12] which name is Determination of amino acids in Cornu Cervi Pantotrichum of different specifications, and then extract the content of essential.

3 Mathematical Model Theory

3.1 Principal Component Cluster Analysis

In the principal component cluster analysis, when the variance contribution of first principal component F_1 is high (Generally more than 85%), it can be considered that the principal component can reflect the information provided by the original P variables, at this time can be based on the first principal component score to sort the assessment.

1. The erroneous zone of comprehensive score of principal component: In many cases, the variance contribution rate of first principal component is not high enough, it should to thought a number of main components which are named as F_1, F_2, \dots, F_r , in order to sort the sample, the common method is that we make the variance contribution rate a_i of principal components F_i as weight, and construct the composite score about main component:

$$H = a_1F_1 + a_2F_2 + a_3F_3 + \dots + a_rF_r, \quad (1)$$

Then, according to the formula, the main component of the sample is calculated and sorted.

The comprehensive score calculation looks very reasonable, it seems to increase the information content, that is, to increase the variance contribution rate, in fact, this is just an illusion, by calculating the variance of H we can found, it is possible to get completely opposite conclusions.

$$Var(H) = \sum_{i=1}^r a_i^2 Var(F_i) = \sum_{i=1}^r a_i^2 \lambda_i \leq \sum_{i=1}^p a_i^2 \lambda_i \leq \lambda_i, \quad (2)$$

From this derivation, we found that the variance of composite score is smaller than the variance of the first principal component, indicating that the information contained in comprehensive principal component scores than the first principal component contains even less information.

2. Principal component clustering: For the multi-index evaluation system sort problem, if the variance contribution rate of the first principal component is not enough, that is, the raw data that is expressed in the first principal component information is not enough, only by the first principal component score of the samples will have sidedness about sort. At this time, the principal component analysis and cluster analysis can be combined, that is principal component cluster analysis.

Cluster analysis is a method which classify the sample or variable according closeness degree of properties, although the cluster analysis can classify well the data which contain many indexes, but it can not get the evaluation result of the degree of excellence. The thought of principal component cluster analysis is to analysis principal components firstly, and then take a number of the main components of the sample to do cluster analysis, last combine with the first principal component score sort to classify the sample ranked. Thus we get a new comprehensive evaluation method, the specific work is following: selecting the R principal components according to the cumulative contribution rate calculate the principal component scores:

$$F_k = a_{1k}X_1 + a_{2k}X_2 + \dots + a_{pk}X_p, k = 1, 2, \dots, r \quad (3)$$

The selected new data matrix (F_1, F_2, \dots, F_r) can be analysed by system clustering. Then using the calculated mean score of the samples determine the order between different classes; Finally, according to the first principal component score of samples determine the order of samples, and give the comprehensive evaluation.

4 Results and Discussion

4.1 with the Principal Component Cluster Analysis Methods for Data Processing

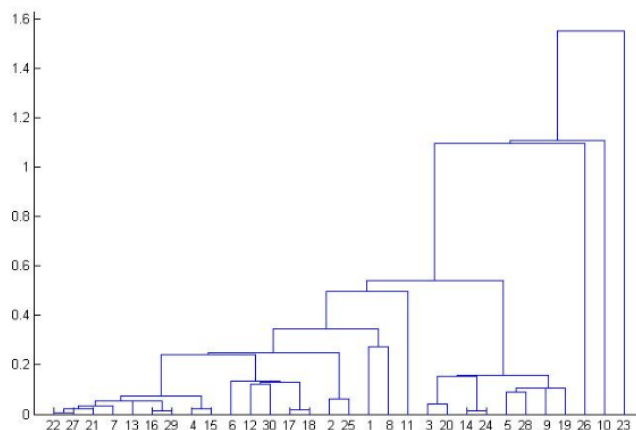


Fig. 1. Velvet antler human body essential amino acids principal components clustering

The result by Matlab of Seven kinds of essential amino acid content using principal component cluster analysis is shown in Figure 1, the waxes of (5, 9, 19, 28) are clustered into a large class, and the other three waxes are clustered into a class, the result of cluster analysis about sika deer antler and red deer antler is not clear. Sika Deer Velvet Deer Antler wax and Red Deer antler wax tablets can be regarded as the same level, and its amino acid content is similar, indicating that Sika Deer Velvet Deer Antler wax and Red Deer antler have similar medicine effect.

Each essential amino acid content reflects the quality of antler, we divide it into seven indicators, they are (Met) methionine X_1 , (Ile) isoleucine X_2 , (Leu) leucine X_3 , (Phe) phenylalanine X_4 , (Val) valine X_5 , (Lys) lysine X_6 , (Thr) threonine X_7 .

For the principal component analysis of essential amino acid about Red Deer, we calculate a linear combination of each main component represented by the variables,

$$Z_1 = 0.3554X_1 + 0.3158X_2 + 0.3796X_3 + 0.4001X_4 + 0.3935X_5 + 0.3942X_6 + 0.3994X_7, \quad (4)$$

$$Z_2 = 0.4808X_1 + 0.6675X_2 - 0.4002X_3 - 0.1874X_4 - 0.2935X_5 + 0.0877X_6 - 0.1849X_7, \quad (5)$$

The score for each factor in the amino acid content of Red deer:

Table 1. Red deer antler amino acids score

Red deer	The principal component score 1	The principal component score 2
1	4.9780496	-1.0985852
2	6.2697454	-1.2895272
3	6.991833	-1.70675
4	5.9332288	-1.0979714
5	7.353369	-0.451135
6	5.4390096	-1.1220619
7	5.7467428	-1.160236
8	4.400715	-0.7897087
9	7.166991	-0.598925
10	2.7852456	-0.4334046
11	3.898019	-0.4907397
12	5.0588716	-1.666578
13	5.854442	-0.959294
14	6.830866	-1.49329

The result of pedigree cluster analysis about principal component scores using matlab is following :

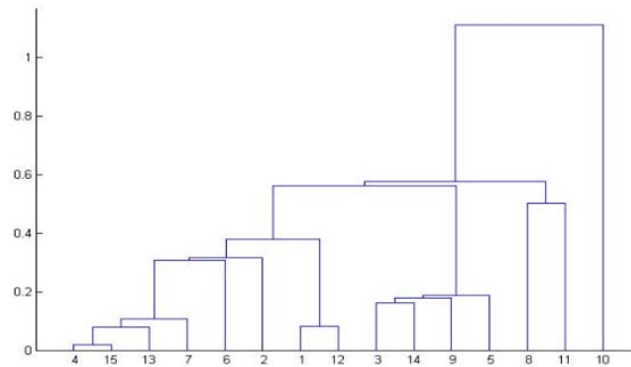


Figure 2. Ma pilose antler human body essential amino acids principal components clustering

For the principal component analysis of essential amino acid about Sika Deer, we calculate a linear combination of each main component represented by the variables,

$$Z_1 = 0.3651X_1 + 0.3646X_2 + 0.3732X_3 + 0.3844X_4 + 0.3835X_5 + 0.3868X_6 + 0.3875X_7, \quad (6)$$

$$Z_2 = 0.5572X_1 + 0.5434X_2 - 0.4752X_3 - 0.1764X_4 - 0.3402X_5 + 0.0649X_6 - 0.1316X_7, \quad (7)$$

The score for each factor in the amino acid content of Sika deer:

Table 2. Sika deer antler amino acids score

Sika deer	The principal component score 1	The principal component score 2
16	5.7469903	-1.5991154
17	5.2282053	-1.1448584
18	5.2473394	-1.1999092
19	7.423939	-1.014214
20	6.95486	-1.70147
21	5.6532929	-1.2921322
22	5.624774	-1.093257
23	10.072079	-1.679304
24	6.739466	-1.934784
25	6.144351	-1.527741
26	8.511762	-1.862528
27	5.6281678	-1.4326174
28	5.854442	-0.068999133
29	6.830866	-0.070970537
30	5.9528372	-0.064596331

The result of pedigree cluster analysis about principal component scores using matlab is following:

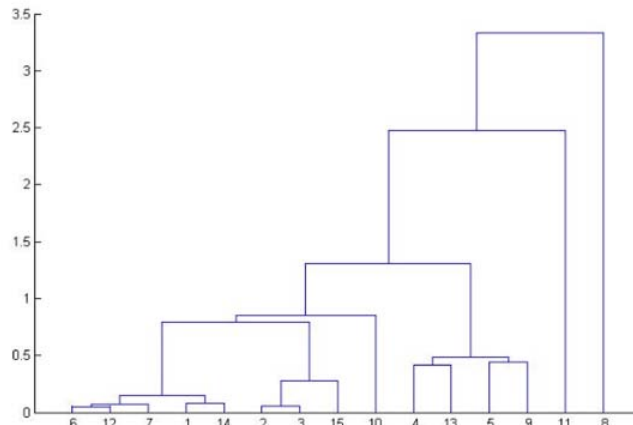


Figure 3. Plum velvet antler human body essential amino acids principal component cluster

In the figure 3, the number 4 and the number 13 get together. Table 1 shows that Figure 3 make the waxes get together.

Compared Sika Deer antler and Red Deer antler principal component analysis of essential amino acids, from the linear combination of the first principal component shows that when the quality of Red Deer antler and Sika Deer antler are the same, the methionine and isoleucine content of the Sika Deer antler are higher than Red Deer antler, and the rest of essential amino acids are lower than the Red Deer antler. Among them, the isoleucine content of sika deer antler is 0.05 times bigger than the red deer antler, the methionine content of sika deer antler is 0.01 times bigger than the red deer antler, the rest of the essential amino acids of sika deer antler are 1% less than red deer antler. This explains the differences about the amino acids content of sika deer antler and red deer antler.

5 Conclusions

In this paper, the amino acid content of sika deer antler and red deer antler was analyzed using principal component cluster analysis. The principal component cluster analysis explained the similar medicine effect and different active ingredient between the sika deer antler and red deer. From the charts of article, we can see that the different amino acid content can be clustered into a pile, this shows the amino acid content is very close, so red deer antler and sika deer antler have similar medicine effect. By the principal component analysis shows the difference of essential amino acids nutrients between the sika deer antler and red deer antler, and provides a theoretical basis for evaluation of the quality characteristics of the antler.

References

1. Chinese Pharmacopoeia Commission. Pharmacopoeia of the people's Republic of China[M]. Chemical Industry Press, 264-265(2000)
2. Ni jiakui. Summary of clinical application of deer products[J]. Jilin Journal of Traditional Chinese Medicine, 3, 33-34(1986)
3. Chen daixian, Guo yueqiu, Ren wei. Character identification of 12 kinds of pilose antler medicinal materials[J]. Journal of Chinese Medicinal Materials, 22(9), 441-444(1999)
4. Bernhard,K., Brubacher,G., Hediger,H., Bruhin,H. The chemical composition and structure of Deer antlers[J]. Experientia, 9, 138-140(1953)
5. Silave,A.B., Katrukha,G.S., Shampanova,O.M., Tevi,A.S.Amino acid and mineral compisition of antlers and pantocrine[J]. Sbornik nauchnykh rabot nauchno-Issledovatel' skoi laboratot -ii Pantovogo ollenevodstva, 2, 29-32(1969)
6. Zhang jinghua, Yang ruoming, Zhang linyuan, Cheng yanzheng. Determination of trace elements in elk antler of sika deer and red deer[J]. Studies of Trace Elements and Health, 17(4), 39-40(2000)
7. Ding yuhua, Xu anhong, Shen hua. Determination of the chemical constituents of deer antler[J]. special wild economic animal and plant research, 1, 36-37(1995)

8. Chen dan, Sun xiaoqiu. Sika deer and different parts of amino acids, total phospholipids, calcium and phosphorus content of research[J]. Journal of economic animal, 2(3), 31-34(1998)
9. Fan yulin. Research Progress on chemical composition of pilose antler[J]. Chinese Traditional Patent Medicine, (1), 23-25(1980)
10. Li zehong, Yao yuxia, Wang quankai. Comparison of amino acid content in two branched antler and antler three branched[J]. Amino Acids & Biotic Resources, 25(1), 10-11(2003)
11. Ma shengliang, Wang zhongwu, Li hai. Analysis of several chemical constituents in Xingkai Lake Sika Deer[J]. Special Wild Economic Animal and Plant Research, (2), 39-42(2003)
12. Zhang song, Li feng. The analysis of amino acid content of different specifications of goods in Antler[J]. China Journal of Chinese Materia Medica, 38(12), 1919-1923(2013)