Physichemical Properties of Different Corn Varieties by Principal Components Analysis and Cluster Analysis

[1] Jie Zeng *, [1] Haiyan Gao, [1] Guanglei Li and [1] Junliang Sun
[1]College of Food Science, Henan Institute of Science and Technology, Henan, Xinxiang 453003, P.R.China.
zengjie623@163.com*

(Received on 16th April 2013, accepted in revised form 8th August 2013)

Summary: Principal components analysis and cluster analysis were used to investigate the properties of different corn varieties. The chemical compositions and some properties of corn flour which processed by drying milling were determined. The results showed that the chemical compositions and physicochemical properties were significantly different among twenty six corn varieties. The quality of corn flour was concerned with five principal components from principal component analysis and the contribution rate of starch pasting properties was important, which could account for 48.90%. Twenty six corn varieties could be classified into four groups by cluster analysis. The consistency between principal components analysis and cluster analysis indicated that multivariate analyses were feasible in the study of corn variety properties.

Key words: maize; variety; physicochemical property; principal components analysis; cluster analysis.

Introduction

Corn starch is a valuable ingredient to the food industry, being widely used as a thickener, gelling agent, bulking agent and water retention agent [1-2]. Generally, both amylose content and chain-length distribution of amylopectin chains in starches exerts significant influence on the defectiveness of the crystalline lamellae through an alteration of their structural organization [3-4]. Some reports have characterized the pasting properties of starches from different corn varieties [5]. The viscosity parameters during pasting are controlled by the properties of both the swollen granules and the soluble materials leaking from the granules [6-8]. The distribution of starch molecular weights, structures and compositions, and subsequently their functional properties, depend on processing conditions as well as the variety [9].

The purpose of multivariate technique is to set linear relationships among the variables in order to create new variables which are not correlated among them. By means of this procedure, the dimension of the initial set of data can be reduced, and an improved understanding of the observed data is possible. The new variables called principal components are orthogonal and independent. [10].

Currently, the evaluation index system of maize quality has not been perfect, and the application of multivariate analysis techniques on maize quality research is still at the stage of feasibility research. This study intends to analyze physicochemical properties and chemical composition of 26 different varieties of corn flour and the statistically classification based on the differences. The corn varieties grown in Dandong city were compared on the basis of the composition and some properties by using principal components analysis and Q-type cluster analysis. This could be useful for selecting the appropriate variety for final use suitability and provide a new pathway for evaluating the quality of maize varieties.

Results and Discussion

Composition Analysis and Properties of Corn Flour

The composition analysis and properties of corn flours were listed in Table-1. The average, standard deviation (S.D.) and coefficient of variation (CV) of twenty six corn varieties were counted by the basic parameter estimation analysis. As shown in Table-1, greater difference in damaged starch, amylose, water solubility protein and lipid of the corn flours were observed and their coefficient of variation were 32.90%, 24.30%, 39.50% and 34.80%, The peak, trough, breakdown, final and setback in the pasting properties of twenty six corn varieties were also different and the coefficient of variations were 26.30%, 27.90%, 18.70%, 37.30%, 18.50% and 19.70%, respectively. This could be attributed to the different content of compositions in corn variety.

Principal Component Analysis (PCA)

The loadings of the principal components can define the direction of the greatest variability. The PCA was performed by means of statistical software (DPS 9.5) based on covariance matrix. The PCA provide an overview of the similarities and differences between the measured properties. For each principal component, the software provides the following attributes: Eigen value, proportion of total variance (percentage) and factorial matrix with the coefficients of polynomial expression among principal components and initial normalized variables. The normalization was made by subtracting the average value and dividing by standard deviation.

*To whom all correspondence should be addressed.

Table-1: Composition analysis and properties of corn flour.

Variety	*DS /mg.mL^{-1}*	*Amylose / %*	*WSP /mg.g^{-1}*	*Starch /%*	*protein /%*	*lipid /%*	*WBC /g.g^{-1}*	*Peak /cp*	*Trough /cp*	*Breakdown /cp*	*Final /cp*	*Setback /cp*	*Peak Time /min*	*Pasting Temp /℃*
Danyu48	**0.32**	**18.54**	**0.52**	**75.99**	**9.50**	**1.52**	**1.42**	**7299**	**3263**	**4036**	**5151**	**1888**	**4.27**	**74.15**
Shenyu22	**0.47**	**22.21**	**0.61**	**76.80**	**9.64**	**1.61**	**1.43**	**6796**	**2675**	**4121**	**4667**	**1992**	**4.07**	**73.60**
Shenyu21	**0.46**	**26.20**	**1.53**	**76.24**	**9.69**	**2.53**	**1.46**	**6663**	**2331**	**4332**	**4069**	**1738**	**4.07**	**74.2**
Danyu69	**0.33**	**20.37**	**1.52**	**76.38**	**10.05**	**1.52**	**1.33**	**6968**	**2758**	**4210**	**4821**	**2063**	**4.13**	**74.35**
Jingang12	**0.38**	**19.65**	**1.06**	**76.65**	**9.39**	**2.06**	**1.29**	**6599**	**2704**	**3895**	**4324**	**1620**	**4.00**	**73.50**
Tiedan22	**0.38**	**23.48**	**1.10**	**77.64**	**8.68**	**1.10**	**1.38**	**7760**	**3110**	**4650**	**5240**	**2130**	**4.07**	**71.75**
Danyu86	**0.40**	**10.23**	**0.55**	**76.24**	**10.20**	**3.00**	**1.44**	**3445**	**1609**	**1836**	**2754**	**1145**	**4.33**	**74.55**
Jiafeng10	**0.31**	**21.56**	**1.13**	**74.76**	**9.36**	**3.13**	**1.39**	**6840**	**2984**	**3856**	**4862**	**1878**	**4.13**	**72.70**
Danyu46	**0.33**	**23.45**	**1.01**	**76.90**	**9.45**	**2.81**	**1.41**	**6952**	**3126**	**3826**	**5123**	**1997**	**4.20**	**73.50**
Haihe12	**0.29**	**19.49**	**1.02**	**75.75**	**10.34**	**2.02**	**1.40**	**6971**	**2858**	**4113**	**4581**	**1723**	**4.13**	**75.15**
Danke2151	**0.29**	**12.66**	**1.08**	**76.79**	**9.46**	**3.08**	**1.46**	**6915**	**2964**	**3951**	**4834**	**1870**	**4.20**	**74.50**
Danyu57	**0.28**	**17.44**	**0.52**	**76.46**	**9.08**	**1.52**	**1.38**	**6661**	**2801**	**3860**	**4587**	**1786**	**4.13**	**73.45**
Tiedan15	**0.17**	**20.23**	**1.04**	**76.26**	**10.39**	**1.04**	**1.30**	**4408**	**2222**	**2186**	**3728**	**1506**	**4.53**	**75.05**
Liaodan565	**0.29**	**11.30**	**0.52**	**75.74**	**9.50**	**1.52**	**1.37**	**4258**	**2182**	**2076**	**3474**	**1292**	**4.47**	**74.35**
Dongdan60	**0.23**	**20.51**	**1.05**	**75.58**	**9.54**	**2.05**	**1.37**	**4795**	**2395**	**2400**	**3814**	**1419**	**4.40**	**76.00**
Mingyu2	**0.35**	**16.98**	**1.03**	**75.10**	**10.45**	**2.03**	**1.41**	**3796**	**2234**	**1562**	**3547**	**1313**	**4.67**	**73.55**
Danyu79	**0.39**	**19.47**	**0.57**	**75.54**	**9.30**	**1.57**	**1.48**	**4559**	**2277**	**2282**	**3816**	**1539**	**4.53**	**73.40**
Jinggang7	**0.34**	**10.32**	**1.10**	**74.32**	**10.84**	**1.10**	**1.36**	**3999**	**2033**	**1966**	**3222**	**1189**	**4.47**	**75.80**
Liao613	**0.29**	**20.89**	**1.01**	**74.87**	**9.42**	**1.01**	**1.31**	**4200**	**2338**	**1862**	**3620**	**1282**	**4.47**	**72.75**
Dongdan70	**0.37**	**17.12**	**1.04**	**74.80**	**9.69**	**2.04**	**1.29**	**4280**	**2021**	**2259**	**3251**	**1230**	**4.40**	**74.55**
Sanbei6	**0.30**	**21.19**	**1.52**	**73.12**	**9.99**	**2.52**	**1.19**	**3467**	**1880**	**1587**	**3059**	**1179**	**4.67**	**76.00**
Liaodan43	**0.30**	**12.64**	**0.50**	**73.89**	**9.28**	**1.50**	**1.42**	**4066**	**2225**	**1841**	**3532**	**1307**	**4.60**	**72.75**
Danke2165	**0.79**	**10.43**	**0.56**	**75.21**	**9.10**	**1.56**	**1.42**	**4352**	**2334**	**2018**	**3608**	**1274**	**4.67**	**75.90**
Zhengdan958	**0.56**	**18.68**	**0.50**	**78.67**	**8.60**	**1.50**	**1.44**	**8219**	**3051**	**5168**	**4929**	**1878**	**3.93**	**72.70**
Tiedan18	**0.31**	**17.34**	**0.50**	**75.98**	**10.02**	**1.50**	**1.29**	**6769**	**3323**	**3446**	**5135**	**1812**	**4.20**	**74.15**
Hu202	**0.37**	**17.25**	**0.40**	**78.80**	**9.70**	**1.40**	**1.39**	**8435**	**3359**	**5076**	**5265**	**1906**	**4.00**	**74.15**
Average	**0.36**	**18.06**	**0.88**	**75.94**	**9.65**	**1.86**	**1.38**	**5748**	**2579**	**3169**	**4192**	**1613**	**4.30**	**74.10**
S.D.	**0.12**	**4.39**	**0.35**	**1.30**	**0.56**	**0.64**	**0.07**	**1604**	**481**	**1181**	**774**	**317**	**0.23**	**1.12**
CV / %	**32.90**	**24.30**	**39.50**	**1.70**	**5.80**	**34.80**	**4.90**	**27.90**	**18.70**	**37.30**	**18.50**	**19.70**	**5.40**	**1.50**

DS, Damaged starch; WSP, Water solubility protein; WBC, Water bingding capacity; S.D., Standard deviation; CV, Coefficient of variation

Results of principal components analysis for different corn variety based on covariance matrix were shown in Table-2; It can be observed that five factors, that is, five principal components, derived from the orthogonal transformation of the matrix, explained 87.24 % of the total variation according to their eigen values (factor 1: 6.85%; factor 2: 2.15%; factor 3: 1.27%, factor 4: 1.11%, and factor 5: 0.83%).

Table-2: Eigen values of the correlation matrix and its contribution and cumulative contribution.

	Eigen value	Percentage of variance	Cumulated percentage %
Factor 1	**6.85**	**48.90**	**48.90**
Factor 2	**2.15**	**15.39**	**64.30**
Factor 3	**1.27**	**9.06**	**73.36**
Factor 4	**1.11**	**7.95**	**81.30**
Factor 5	**0.83**	**5.93**	**87.24**

The significant coefficients (p-value < 0.05) were in bold Figures in Table-3. In Table 3, it could be appreciate the composition analysis and properties of corn flour associated to each principal component. Factor 1 related to peak, breakdown, final and setback, was affected by the content of amylose content. Factor 2 negatively related to DS and WBC, and positively related to WSP and starch content, and the correlations were significant. Factor 3 related to lipid (r=0.7501)and WBC (r =0.4164). Factor 4 negatively related to WSP, and positively related to Protein content and pasting temperature. And Factor 5 related to DS, the correlation being significant among them (r=0.7037).

Table-3: Values of correlation coefficients of factorial matrix (p=0.05, r=0.3882; p=0.01, r = 0.4958).

		Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
***x* (1)**	**DS**	**0.0444**	**-0.4127**	**0.3326**	**-0.0914**	**0.7037**
***x* (2)**	**WSP**	**0.1825**	**0.4159**	**-0.0051**	**-0.4180**	**0.1782**
***x* (3)**	**Starch**	**-0.0567**	**0.5592**	**0.2613**	**-0.1814**	**0.2578**
***x* (4)**	**Amylose**	**0.3086**	**-0.1454**	**0.0027**	**0.2094**	**0.0829**
***x* (5)**	**Protein**	**-0.1971**	**0.2901**	**0.0087**	**0.5563**	**-0.1162**
***x* (6)**	**Lipid**	**-0.0060**	**0.1729**	**0.7501**	**-0.0369**	**-0.3458**
***x* (7)**	**WBC**	**0.1275**	**-0.3977**	**0.4164**	**0.0055**	**-0.3258**
***x* (8)**	**Peak**	**0.3734**	**0.0565**	**0.0158**	**0.1342**	**0.0624**
***x* (9)**	**Trough**	**0.3429**	**0.0323**	**-0.1564**	**0.1493**	**-0.0738**
***x* (10)**	**Breakdown**	**0.3673**	**0.0636**	**0.0852**	**0.1214**	**0.1149**
***x* (11)**	**Final Visc**	**0.3615**	**0.0692**	**-0.0954**	**0.1073**	**-0.0598**
***x* (12)**	**Setback**	**0.3613**	**0.1199**	**0.0045**	**0.0352**	**-0.0341**
***x* (13)**	**Peak Time**	**-0.3373**	**-0.0889**	**-0.1162**	**-0.1202**	**-0.0352**
***x* (14)**	**Pasting Temp**	**-0.2101**	**0.1234**	**0.1743**	**0.5891**	**0.3632**

Previous studies had only compared the physicochemical properties of different varieties of corn [11-12]. While the principal component analysis about the properties of corn flour have not been reported. Our study ultimately obtained five principal components, in which the starch quality is the most important principal component affecting the quality of corn flour.

Cluster Analysis

Among various clustering algorithms, frequently used hierarchical clustering algorithms were compared to determine the better method that is believed to more closely create naturally existing clusters. Ward’s minimum variance method tries to minimize the sum of squares within each cluster. Fig. 1 and Table-4 showed the presence of four clusters by ward method (squariance method) and Chi-square distance according to the average of pasting peak viscosity from high to low, which meaned that pasting properties was an important principal component. The first cluster with the least distance was formed by Danyu48, Danyu57, Tiedan18, Shenyu22, Tiedan22, Zhengdan958 and Hu202. The second cluster was formed by shenyu21, Danyu69, Haihe12, Jingang12, jiafeng10, Danyu46, Danke2151. The third one was formed by Danyu86, Liaodan565, Liaodan43, Danyu79 and Danke2165. The last cluster grouped Tiedan15, Liao613, Dongdan60, Dongdan70, Mingyu2, Jinggang7 and Sanbei6.

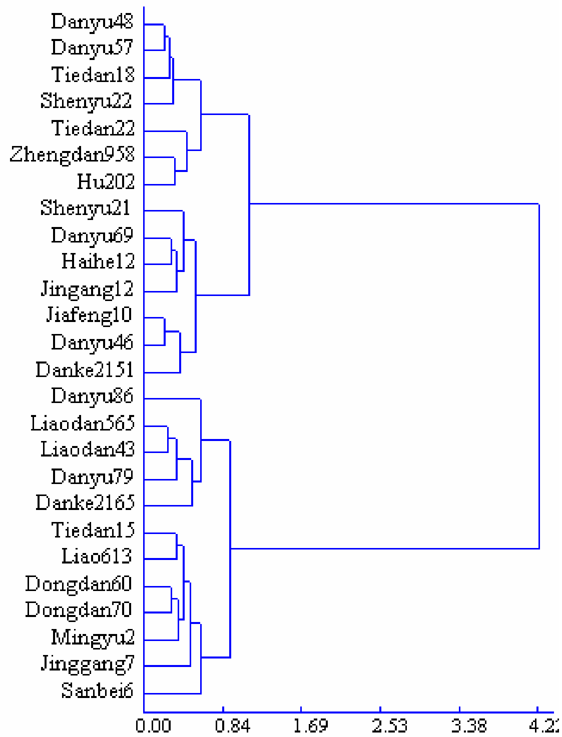

Fig. 1: Cluster Analysis of the twenty six “known” corn varieties

Currently, the cluster analysis about the processing qualities of corn grains also has not been reported. However, some reports discussed the effects of different regions, varieties, planting conditions, maturity and agronomic traits on the qualities of maize, and the cluster analysis were used to classify the corn varieties into several groups [13-14]. These results showed that the multivariate analysis was a viable research method in maize quality evaluation.

Experimental

Materials

Twenty six improved dent corn varieties were obtained from Maize Research Institute, Dandong P.R.China.

Corn was dehulled and milled in a mill with 100 mesh sieve and the obtained flour was stored at 4 °C.

Composition Analysis of Corn Flour

Crude protein content was determined according to the Kjeldahl method with a conversion factor of 6.25.

Crude lipid content was measured by the Soxhlet extraction method.

Total starch content was anlayzed by the enzymatic colorimetric method of Holmes et al [15] with a conversion factor of 0.9. The results included mono- and disaccharides which were disregarded since they were only present in small quantities in raw cereal grains.

Damaged starch (DS) was determined by a method based on the evaluation of susceptibility to amyloglucosidase hydrolysis [16].

Amylose content of the corn slurries was determined by using the method of Williams et al [17].

The concentration of water solubility protein (WSP) was determined by the method of Uhm et al [18], using bovine serum albumin as standard.

Table-4: Mean value of different category.

	Water retention capability(g/g)	Breakage starch (mg/mL)	Starcn (%)	Protein (%)	Peak (cP)	Trough (cP)	Breakdown (cP)	Final Visc (cP)	Setback (cP)	Peak Time (min)	Pasting Temp (℃)	SN (cP)
Ⅰ	1.40	0.38	77.91	9.21	7674.8	3116.8	4558	5034.4	1917.6	4.08	73.24	1993.8
Ⅱ	1.38	0.35	76.81	9.78	6830.33	2858.11	3972.22	4712.9	1854.8	4.126	73.961	1282.7
Ⅲ	1.38	0.38	75.09	9.40	4358.57	2253.14	2105.43	3587.86	1334.71	4.51	74.24	1409.9
Ⅳ	1.34	0.31	75.01	10.37	3823	1995.6	1827.4	3262	1266.4	4.534	74.99	1771

Water Binding Capacity of Corn Flour

Water binding capactity (WBC) of the corn flour or the fermented flour was determined according to the method described by Wang [19].

Pasting Properties of Corn Flour

Pasting properties of the corn flour or fermented flour were evaluated with the Rapid Visco Analyzer (RAV-4, Newport Scientific, Warriewood, Australia) according to the method of Sandhu et al [1]. The recorded parameters included pasting temperature, peak viscosity, trough viscosity (minimum viscosity at 95 °C), final viscosity (viscosity at 50 °C), breakdown viscosity (peak-trough viscosity) and setback viscosity (final-trough viscosity). All measurements were carried out in triplicate.

Statistical Analysis

The multivariate statistical technique called principal component analysis (PCA) is based on the calculation of linear combinations between the variables that explain the most variance of the data. As a result, data can be reduced to a set of new variables called principal components. The loadings of the principal component (PC) define the direction of the greatest variability. All compositions and properties were subjected to the principal component analysis to evaluate the relationships among them in order to identify the PC associated with the optimal separation of the grain components.

Cluster analysis was conducted to separate corn varieties into several subgroups with respect to composition analysis and properties of corn flour to produce properties uniform varieties. Different clustering algorithms often produce substantially different clusters when the same data set is used.

The principal component analysis and cluster analysis were performed using DPS 9.5 software package. All measurements were repeated three times, and the average values were used as data.

Conclusion

Different corn varieties behaved differently in terms of the composition contents and some of their physicochemical properties. Principal component analysis identified five components that explained 87.24% of the total variation among corn flours properties. Twenty six corn varieties were classified into four groups by cluster analysis. The results suggested that the application of multivariate statistics in corn variety characteristics was practical and effective.

Acknowledgements

The financial support provided by key teachers funding Scheme in Henan Province (project No. 2010GGJS-141) and Science and Technology Plan of the Henan Province in China (project No.132102110032) .

References

1. K. S. Sandhu and N. Singh, *Food Chemistry*, **101**, 1499 (2007).
2. N. Singh, J. Singh, L. Kaur, N. S. Sodhi and B. S. Gill, *Food Chemistry*, **81**, 219 (2003).
3. S. S. Kozlov, A. Blennow, A. Krivandin and V. P. Yuryev, *International Journal of Biological Macromolecules*, **40**, 449 (2007).
4. D. A. Koroteeva, V. I. Kiseleva, A. V. Krivandin, O. V. Shatalova, W. Błaszczak, E. Bertoftc, K. Piyachomkwan and V. P. Yuryev, *International Journal of Biological Macromolecules*, **41**, 534 (2007).
5. Y. Ji, K. Wong, J. Hasjim, L. M. Pollak, S. Duvick and J. Jane, *Carbohydrate Polymers*, **54**,305 (2003).
6. D. Kuakpetoona andY. J. Wang, *Carbohydrate Research*, **343**, 90 (2008).
7. K. S. Sandhu, N. Singh and M. Kaur, *Journal of Food Engineering*, **64**, 119 (2004).
8. K. S. Sandhu, N. Singh and N. S. Malhi, *Food Chemistry*, **89**, 541 (2005).
9. A. Denis, N. Brambati, B. Dessauvages, S. Guedj, C. Ridoux, N. Meffre and C. Autier, *Food Hydrocolloids*. **22**, 989 (2008).
10. A. N. C. Resio, M. P. Tolaba and C. Suárez, *Journal of Food Engineering*, **92**, 275 (2009).
11. K. S. Sandhu, N. Singh and N. S. Malhi. *Food Chemistry*, **101**, 938 (2007).
12. J. Singh, N. Singh, T. R. Sharma and S. K. Saxena. *Food Chemistry*, **83**, 387 (2003).
13. S. D.Jennings, D. J.Myers and L. A. Johnson, *Cereal Chemistry*, **79**, 697 (2002).
14. K. M.Krieger, L. M.Pollak and T. J Brumm. *Cereal Chemistry*, **75**, 656 (1998).
15. J. Holmes, L. Bjorck, A. Drews and N. Asp, *Starch*, **38**, 224 (1986).
16. K. Seetharaman, A. Tziotis, F. Borras, P. J. White, M. Ferrer and J. Robutti, *Cereal Chemistry*, **78**, 379 (2001).
17. P. C.Williams, F. D.Kuzina and I. Hlynka, *Cereal chemistry*, **47**, 411 (1970).
18. T. B Uhm, D. Y Jeon, S. M. Byun, J. S Hong and J. W. D. GrootWassink, *Biochimica et Biophysica Acta*, **926**, 119 (1987).
19. S. Wang, J. Yu, J. Yu, H. Chen, J. Pang and H. Liu. *Journal of Food Engineering*, **88**, 287(2008).