Research Article

Application of Principal Component Analysis Based on *In Vitro* Antioxidant Capacity

Hitesh Kumar*[1], Charanjit Kaur[1] and Sarika Jaiswal[2]

[1]Division of Food Science and Post Harvest Technology, Indian Agricultural Research Institute, New Delhi-110012
[2]Centre for Agricultural Bioinformatics, Indian Agricultural Statistics Research Institute, New Delhi-110012

Abstract

Interrelationships between different parameters (phenolics, flavonoids, and antioxidant activity) in different fruits and vegetables were investigated by principal component analysis (PCA) and correlation matrix. Generally there was high correlation between FRAP and phenols in fruits, vegetables and cereals. PCA revealed that the first two components represented 62% of the total variability in antioxidant activity and different antioxidant groups. In vegetables PC- I accounted for 71% and PC- II accounted for 14% variation whereas in fruits PC-I accounted for 67.9% and PC-II accounted for 21% variation. Cereals crops, however displayed scattered profile; PC-1 and PC-11 accounting for 45.79% and 28.2%. Overall high correlation between FRAP and phenolics in fruits, vegetables and cereals reveals that the phenolics are the major determinants for high antioxidant activity.

Keywords: Principal Component Analysis, Hierarchical cluster analysis, Correlation, Biplot, Antioxidant activity

***Correspondence**
Author: Hitesh Kumar
Email: hitesh.3971@gmail.com

Introduction

Total antioxidant capacity is the term used to describe the ability of antioxidants in different foods to clean harmful free radicals in the blood and cells. It takes into account the amount of water-based and fat-based antioxidants present in food. It is an analytic frequently used to assess the antioxidant status of biological samples and can evaluate the antioxidant response against the free radicals produced in a given disease [1]. The growing consumer awareness on health is increasing demand for health and functional foods. As a result there is increased consumption of fruits and vegetables. Many of the beneficial effects of fruits and vegetables have been found to be related to their high phenolic content. Phenolics are secondary metabolites; commonly found in both edible and non-edible parts of the plants and have been reported to have multiple biological effects, including antioxidant activity. The scavenging activity of phenolics is mainly due to their redox properties, which allows them to act as reducing agents, hydrogen donors, and singlet oxygen quenchers. In addition, many of the natural antioxidants exhibit a wide range of biological effects, including antibacterial, antiviral, anti-inflammatory, anti-allergic, antithrombotic, and vasodilatory actions [2]. Therefore, the potential of these phytochemical constituents of plant materials for the maintenance of health and protection from coronary heart disease and cancer is raising interest among researchers and food manufacturers as consumers have begun to move towards functional foods with special health effects [3]. Classification of fruits and vegetables is thus needed for dietary guidance materials to help people select appropriate types of these foods to meet their need for a healthy diet [4]. Many countries have food guides with graphic depictions of the food groups and subgroups, along with recommendations for consumption [5]. Multivariate mathematical approaches are powerful tools which often permit a relatively simple representation of similarities between samples on the basis of more-or-less complex analytical data. The present study aims to use chemometric tools to gain insights into variations in the complex antioxidant profiles between a selection of fruits and vegetables commonly consumed in India and to classify them based on antioxidant activity, levels of antioxidant groups.

Materials and Methods

I have selected sample 84 crops in which including (28 fruits), (38 vegetables) and (18cereals) Crops. They were taken from IARI, New Delhi, Directorate of sorghum Research, Hyderabad Andhra Pradesh and local Market. All analyses (extraction) were performed in triplicate; each replicate was quantified in duplicate. Data were expressed as means. Pattern recognition methods were applied to the data collection; these were principal component analysis (PCA) as an unsupervised classification method and hierarchical cluster analysis (HCA) as an unsupervised learning method. PCA was applied to the data set after standardization (the mean of the values for each variable is subtracted from each variable value and the result is divided by the standard deviation of the values for each variable.

Results

Correlation matrix between of fruits, vegetables and cereals

To evaluate the comparability between the five antioxidants under study (i.e. Total phenols, flavonoid, DPPH, FRAP and CUPRAC), a set of 38 vegetables, 28 fruits and 18 cereal crops rich in antioxidants were identified and analyzed for their antioxidant capacities. Correlation coefficients were calculated to study the relationship between antioxidant capacities for the data set under study. **Table 1** is the correlation matrix of all 84 crop data. **Figures 1-3** Shows the scatter plots of DPPH vs Total phenols, DPPH vs CUPRAC and FRAP vs CUPRAC. It can be seen from the table and respective figures that, CUPRAC showed a strong positive relationship with FRAP (correlation coefficient= 0.847, p<0.0001) followed by DPPH (correlation coefficient= 0.721, p<0.0001). Similarly DPPH was seen to have positive correlation with total phenols (correlation coefficient= 0.793, p=0.0001)

Table 1 Correlation matrix of total 84 crops data

	Phenol	Flavonoids	FRAP	DPPH	CUPRAC
Phenol	1	0.382(0.0003)	0.576**(<0.0001)	0.793**(0.0001)	0.506**(<0.0001)
Flavonoids		1	0.181(0.1)	0.443**(<0.0001)	0.081(0.465)
FRAP			1	0.689**(<0.0001)	0.847**(<0.0001)
DPPH				1	0.721**(<0.0001)
CUPRAC					1

**Correlation is significant at the 0.01 level (2-tailed)

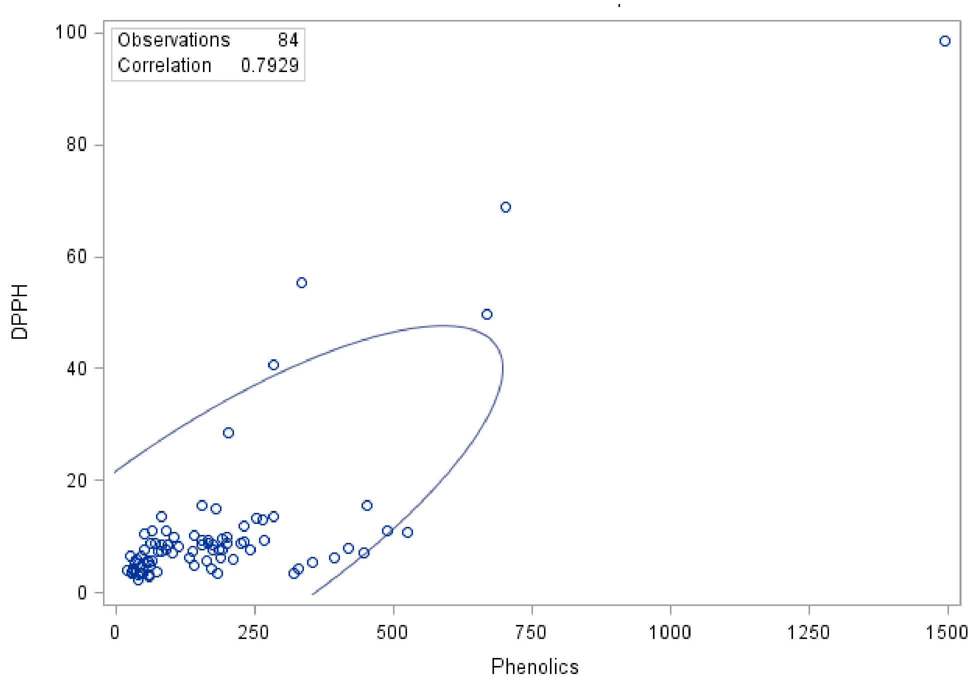

Figure1 Scatter plot between DPPH and total phenols of all vegetables, fruits and cereal crops

Correlation between vegetables and antioxidant activity

Table 2 represents the descriptive statistics of all vegetables. The comparison of antioxidant capacities of total 38 vegetables is summarized in **Table 3**. The correlation matrix shows that antioxidant capacity by DPPH is strongly correlated with Flavonoids (correlation coefficient=0.899, p< 0.000) followed by CUPRAC (correlation coefficient=0.777, p<.000). CUPRAC was strongly positively correlated with Flavonoids (correlation coefficient=0.776, p< .000).

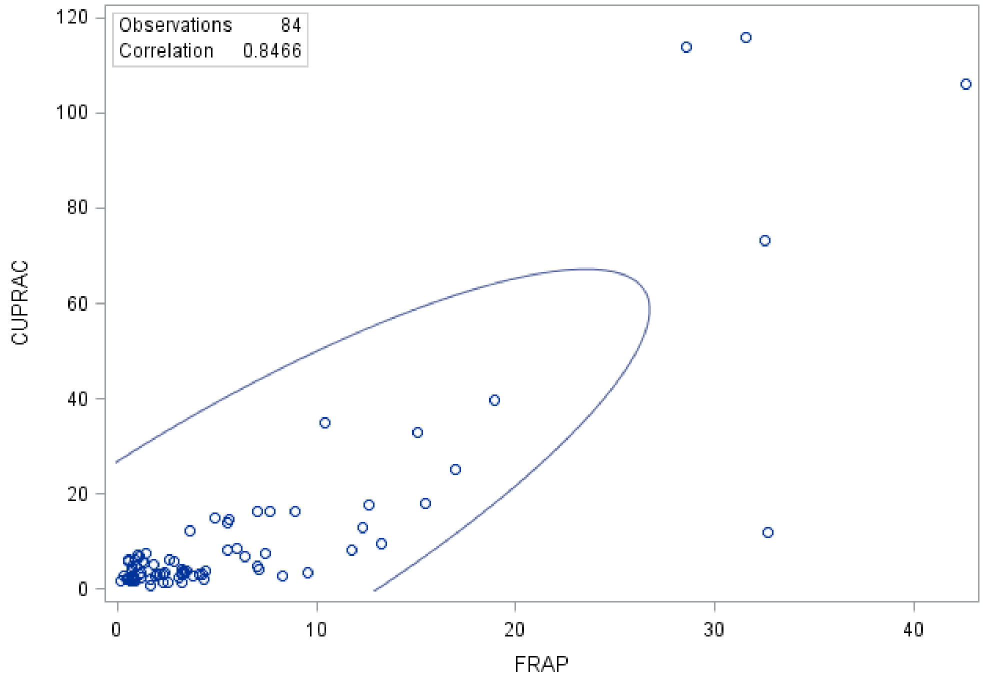

Figure 2 Scatter plot between DPPH and CUPRAC of all vegetables, fruits and cereal crops

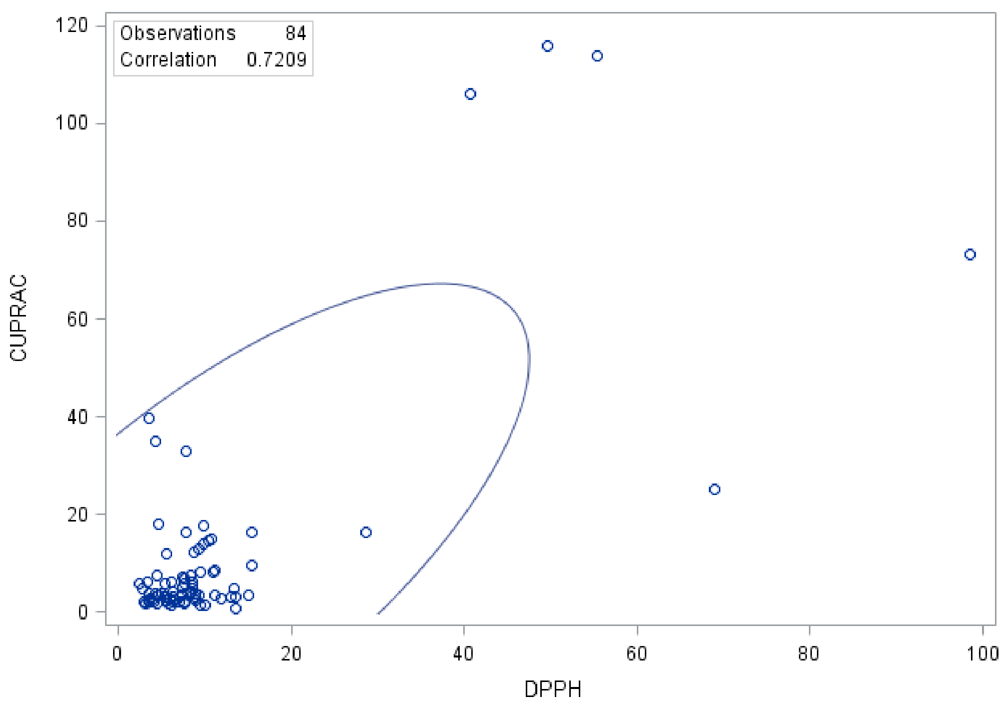

Figure 3 Scatter plot between FRAP and CUPRAC of all vegetables, fruits and cereal crops

Table 2 Descriptive statistics of all vegetables

	Minimum	**Maximum**	**Mean**		**Std. Deviation**
	Statistic	Statistic	Statistic	Std. Error	Statistic
Phenols	20.700	701.025	140.75263	25.048445	154.408986
Flavonoids	4.100	760.000	61.70829	21.406552	131.958852
FRAP	0.363	32.731	3.21176	0.943937	5.818821
DPPH	2.404	68.903	8.29058	1.779478	10.969442
CUPRAC	1.780	25.270	6.12842	.842315	5.192380

Table 3 Correlation matrix of total 38 vegetables

	Phenol	Flavonoids	FRAP	DPPH	CUPRAC
Phenol	1	0.729**(0.000)	0.402*(0.012)	0.639**(0.000)	0.666**(0.000)
Flavonoids		1	0.447**(0.005)	0.899**(0.000)	0.776**(0.000)
FRAP			1	0.446**(0.005)	0.625**(0.000)
DPPH				1	0.777**(0.000)
CUPRAC					1

** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

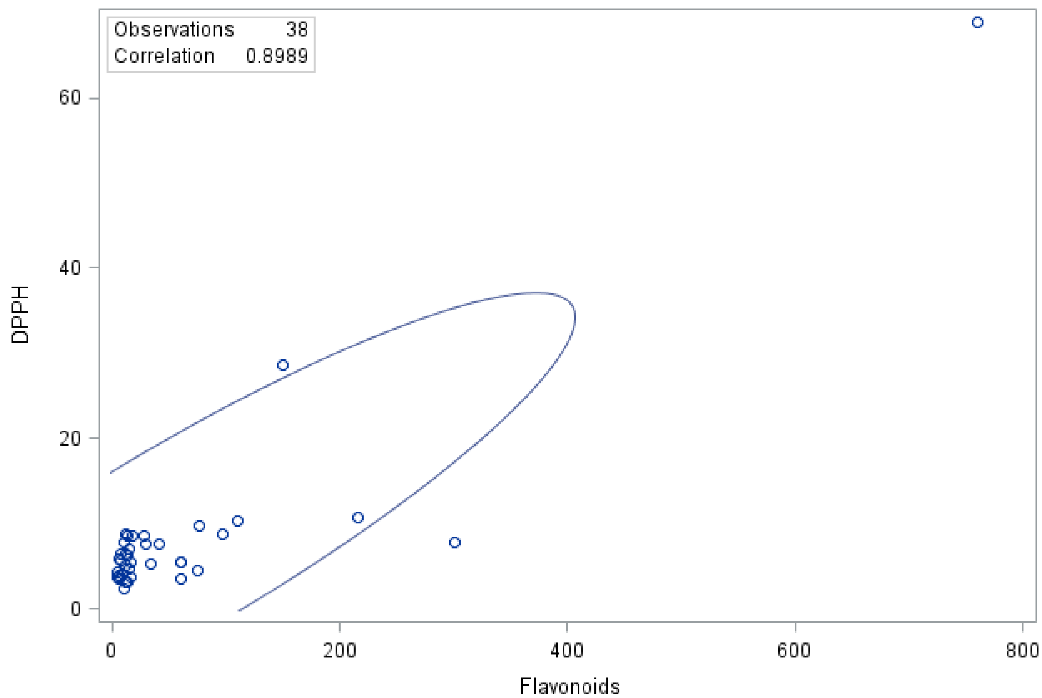

Figure 4 Scatter plot between Flavonoids and DPPH of vegetables

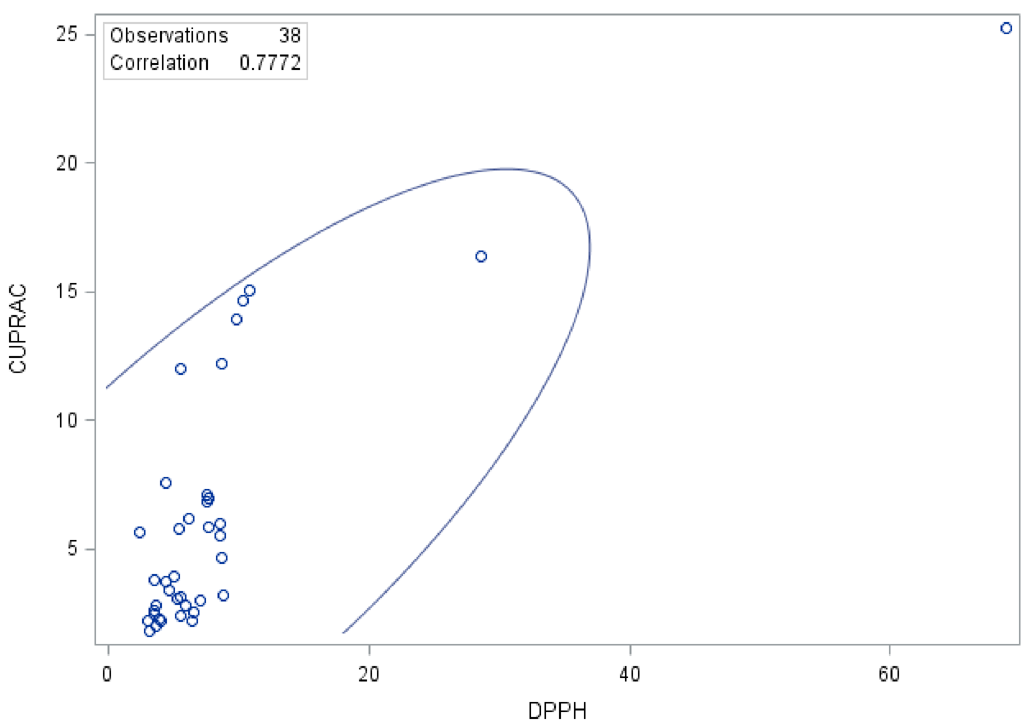

Figure 5 Scatter plot between DPPH and CUPRAC of vegetables

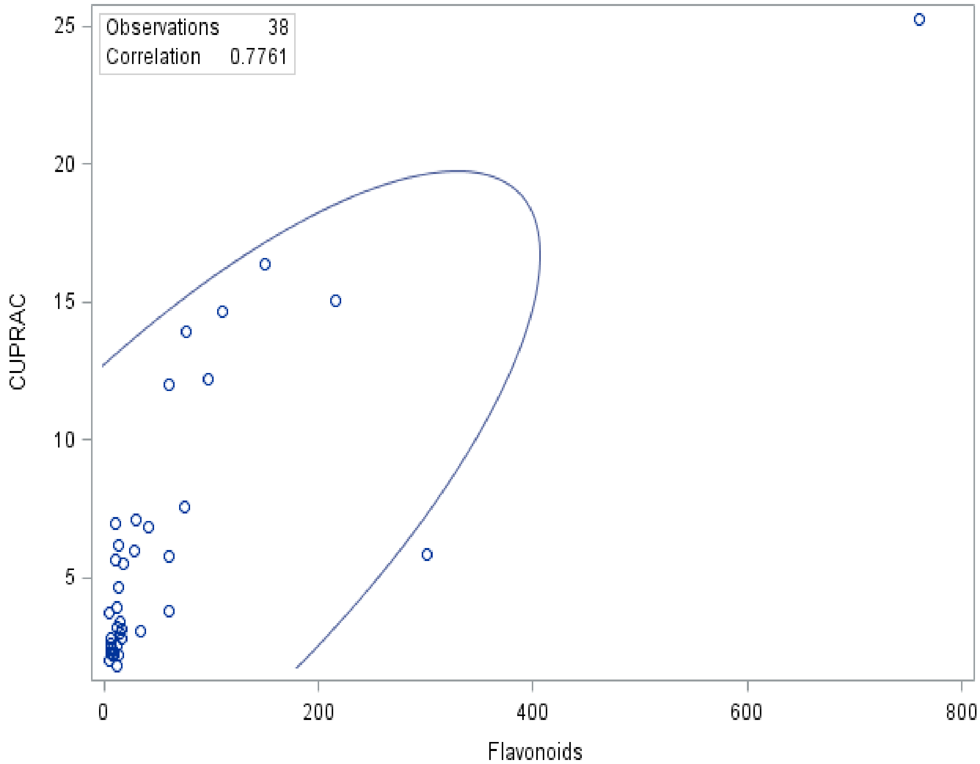

Figure 6 Scatter plot between Flavonoids and CUPRAC of vegetables

Correlation between fruits and antioxidant activity

Table 4 represents the descriptive statistics of fruits' antioxidant data. The antioxidant capacities of total 28 fruits taken under study delineated as correlation matrix in **Table 5**.

Table 4 Descriptive statistics of fruits

	Minimum	Maximum	Mean		Std. Deviation
	Statistic	Statistic	Statistic	Std. Error	Statistic
Phenols	40.68	1493.87	209.9825	55.12056	291.67060
Flavonoids	4.58	82.50	31.0325	4.35826	23.06172
FRAP	0.20	42.64	10.4550	2.13588	11.30203
DPPH	2.81	98.60	15.1521	4.00535	21.19434
CUPRAC	1.61	115.81	24.0829	6.54954	34.65693

CUPRAC is strongly correlated with FRAP in positive direction with significant correlation coefficient as high as 0.917 ($p< .0001$). DPPH is also found to be significantly positively correlated with Phenols with correlation coefficient=0.891, $p< 0.0001$. Surprising there was no correlation between CUPRAC and flavonoids. DPPH and CUPRAC show significant correlation coefficient of 0.774, $p< 0.0001$ between themselves.

Table 5 Correlation matrix of total 28 fruits

	Phenol	Flavonoids	FRAP	DPPH	CUPRAC
Phenol	1	0.473*(.011)	0.655**(.000)	0.891**(.000)	0.581**(.001)
Flavonoids		1	0.261(.18)	0.24(.218)	0(.999)
FRAP			1	0.770**	0.917**(.000)
DPPH				1	0.774** (.000)
CUPRAC					1

**Correlation is significant at the 0.01 level (2-tailed)
*Correlation is significant at the 0.05 level (2-tailed)

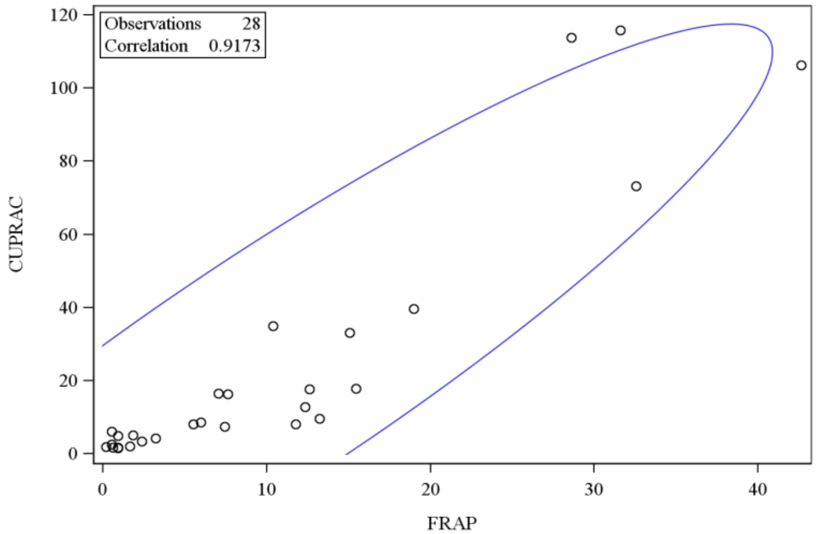

Figure 7 Scatter plot between FRAP and CUPRAC of fruits

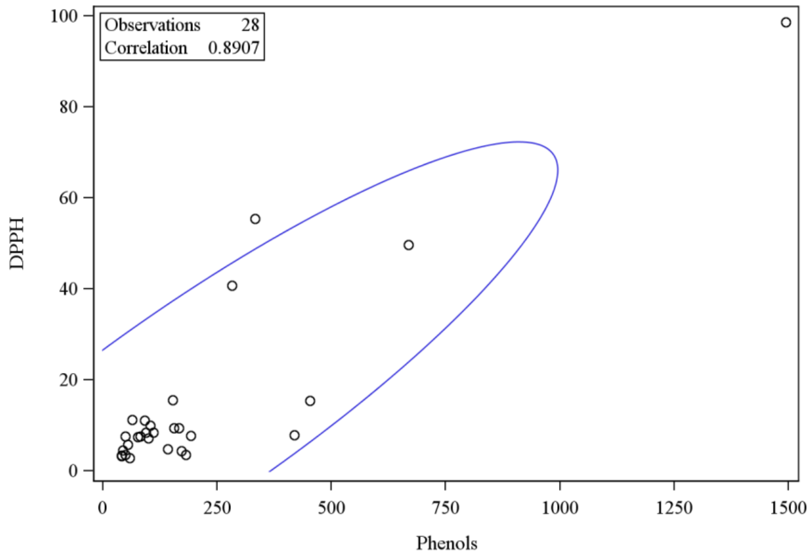

Figure 8 Scatter plot between Phenols and DPPH of fruits

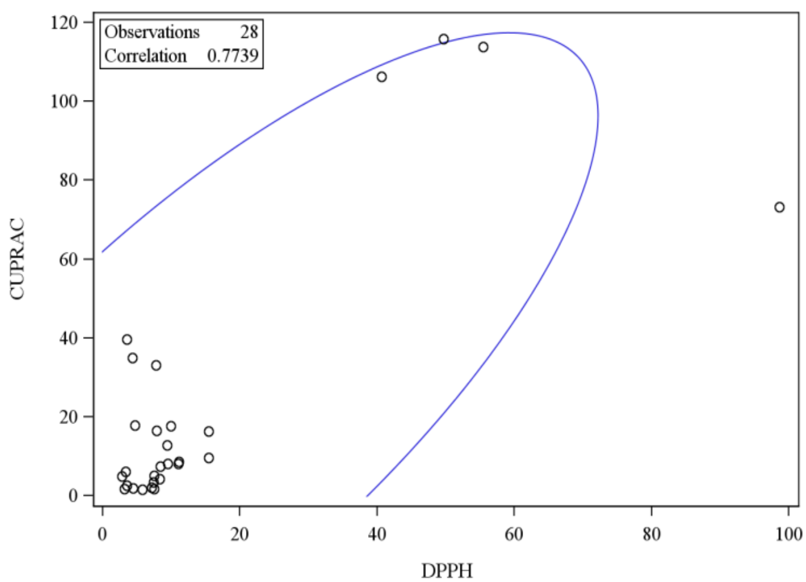

Figure 9 Scatter plot between DPPH and CUPRAC of fruits

Correlation between cereals and antioxidant activity

Total of 18 cereal crops rich in antioxidants were taken for the study. **Table 6** shows the descriptive statistics of all cereals' data. From **Table 7**, it can be seen that phenols are significantly and positively correlated with FRAP (correlation coefficient= 0.793, $p< .0001$) but phenol shows negative correlation with DPPH. Similarly flavonoids also share negative correlation with FRAP.

Table 6 Descriptive statistics of cereals

	Minimum	Maximum	Mean		Std. Deviation
	Statistic	Statistic	Statistic	Std. Error	Statistic
Phenol	81.38	488.41	235.0122	21.56057	91.47376
Flavonoids	2.62	116.89	31.3400	7.41925	31.47722
FRAP	1.66	9.60	4.3950	0.51183	2.17150
DPPH	6.10	15.03	10.1161	0.65526	2.78002
CUPRAC	0.65	4.68	2.8189	0.25296	1.07320

Table 7 Correlation matrix of cereals

	Phenol	Flavonoids	FRAP	DPPH	CUPRAC
Phenol	1	0.145 (0.565)	0.793** (0.000)	-0.094 (.710)	0.538* (0.021)
Flavonoids		1	-0.056 (0.825)	0.399 (0.101)	0.070 (0.783)
FRAP			1	-0.047 (0.854)	0.584* (0.011)
DPPH				1	0.087 (0.731)
CUPRAC					1

**Correlation is significant at the 0.01 level (2-tailed)
*Correlation is significant at the 0.05 level (2-tailed)

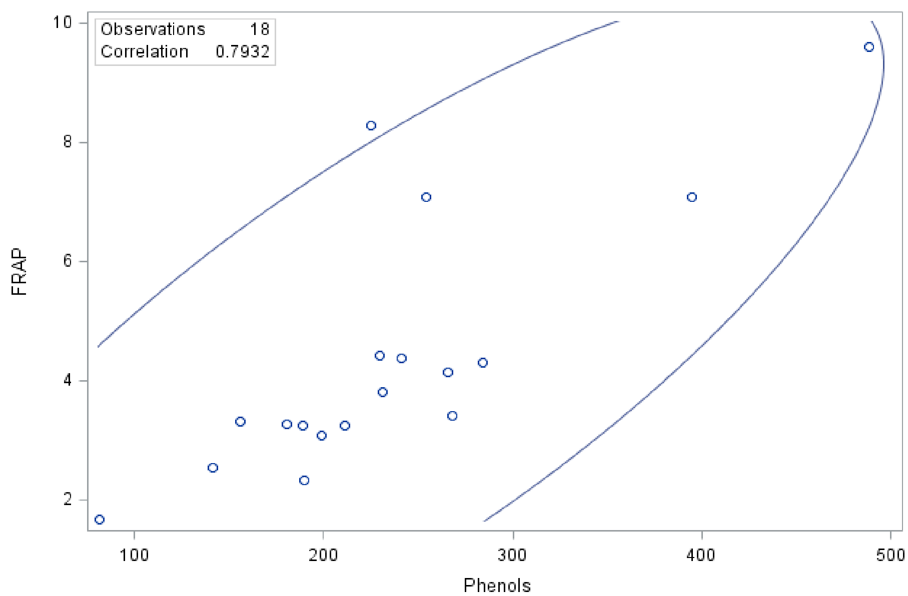

Figure 10 Scatter plot between Phenols and FRAP of cereal crops

Table 8 Eigen values and % variance

PC	Eigen value	%Variance	Cumulative %
1	3.604	72.086	72.086
2	0.705	14.108	86.194
3	0.396	7.927	94.122
4	0.204	4.076	98.198
5	0.09	1.802	100

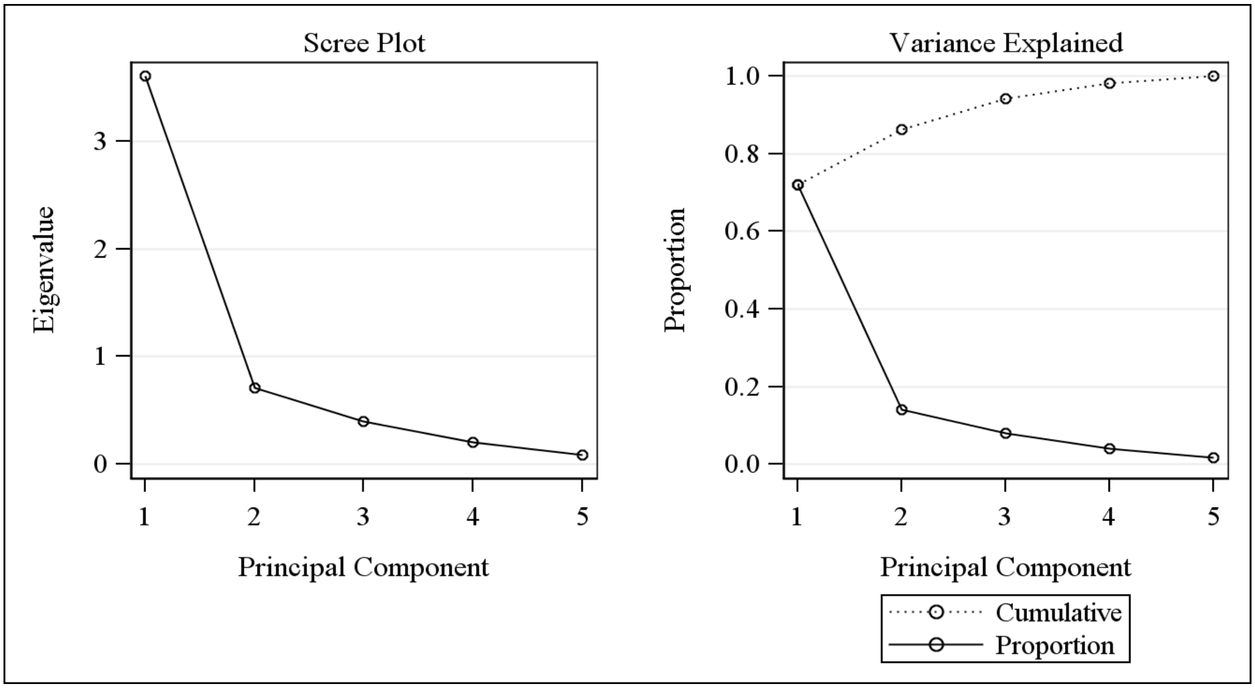

Figure 11 Scree plot and Variance plot of cumulative variance for all vegetables

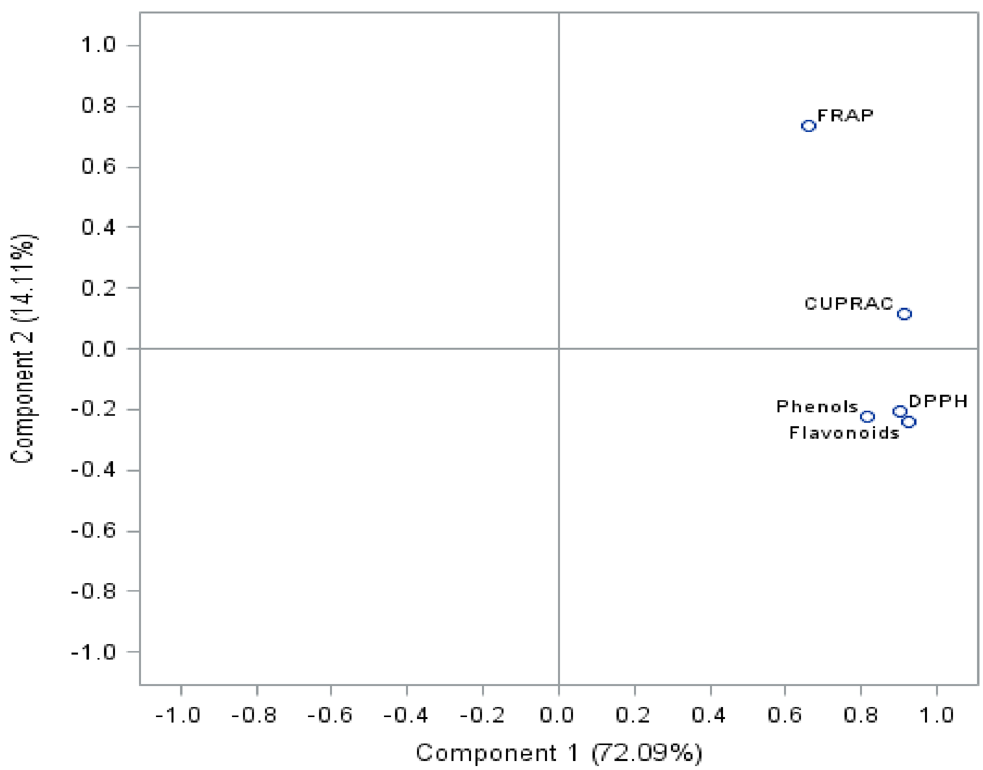

Figure 12 Loadings plot for five different antioxidant assays of vegetable data set (i.e. FRAP, CUPRAC, DPPH, Phenols, Flavonoids) (PC1 vs. PC2)

[1-Brinjal (Purple); 2-Brinjal (Green); 3-Brinjal (White); 4-Garlic; 5-Carrot (Red); 6-Carrot (Orange); 7-Carrot (Black); 8-Cauliflower (White); 9-Cauliflower (Purple); 10-Cauliflower (Orange); 11-Capsicum (Green); 12-Capsicum (Red); 13-Tomato (Red); 14-Tomato (Orange); 15-Tomato (Cherry); 16-Onion (White); 17-Onion (Red); 18-Onion (Orange); 19- Cabbage (Green); 20-Cabbage (Purple); 21-Broccoli; 22-Spinach leaf; 23-Fenugreek leaf; 24-Mustard leaf; 25-Beet root; 26-Drumstick flower; 27-Bitter gourd; 28-Radish (Pod); 29-Bathua leaf; 30-Gram grain (Green); 31-Kasuri methi leaf; 32-Potato; 33-Sungrow (Red carrot); 34-Rudhira (Red carrot); 35-Dumar; 36-Kachnarflower; 37-Dolichus bean (Purple); 38-Dolichus bean (Green)].

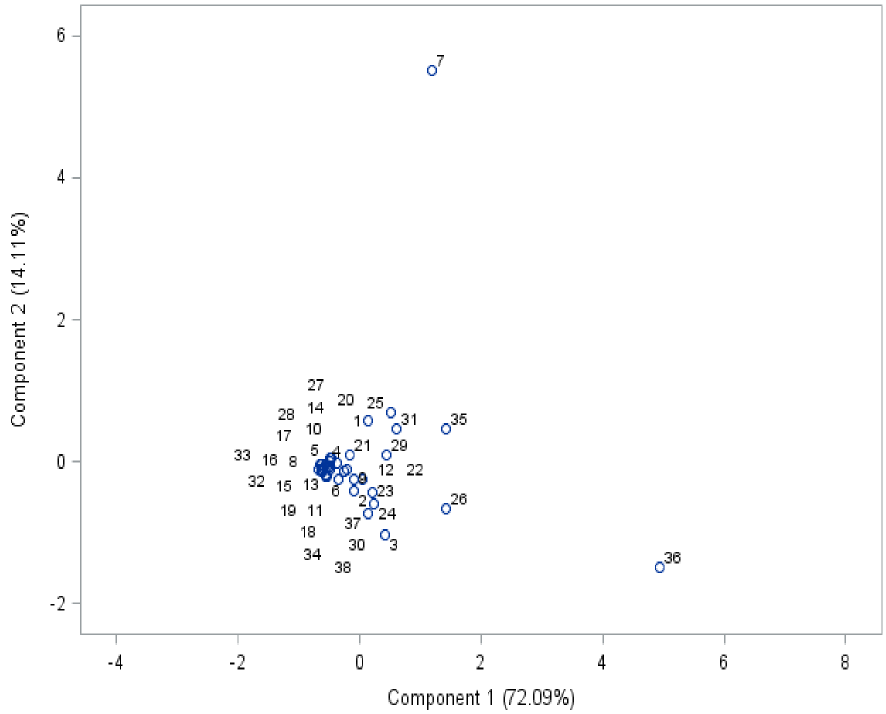

Figure 13 Scores plot for all 38 vegetables under study (PC1 and PC2)

As evident, in PC1, all the five assays have positive loadings, whereas in PC2, only FRAP and CUPRAC showed positive loading. The scores plot in **Figure 13** is used to get an overview of similarities or difference among various 38 vegetable categories. The vegetables like Carrot (represented by 7) and Kachnar flower (represented by 36) are very much different. The similar groups are formed by 4,5,6,7,8,9,10,11,13,14,15,16,17,18,19,20,21,22,27, 28,32,33,34,38.

Table 9 Eigen values and % variance

PC	Eigen value	%Variance	Cumulative %
1	3.396	67.92	67.92
2	1.082	21.64	89.56
3	0.420	8.39	97.95
4	0.060	1.2	99.15
5	0.042	0.85	100

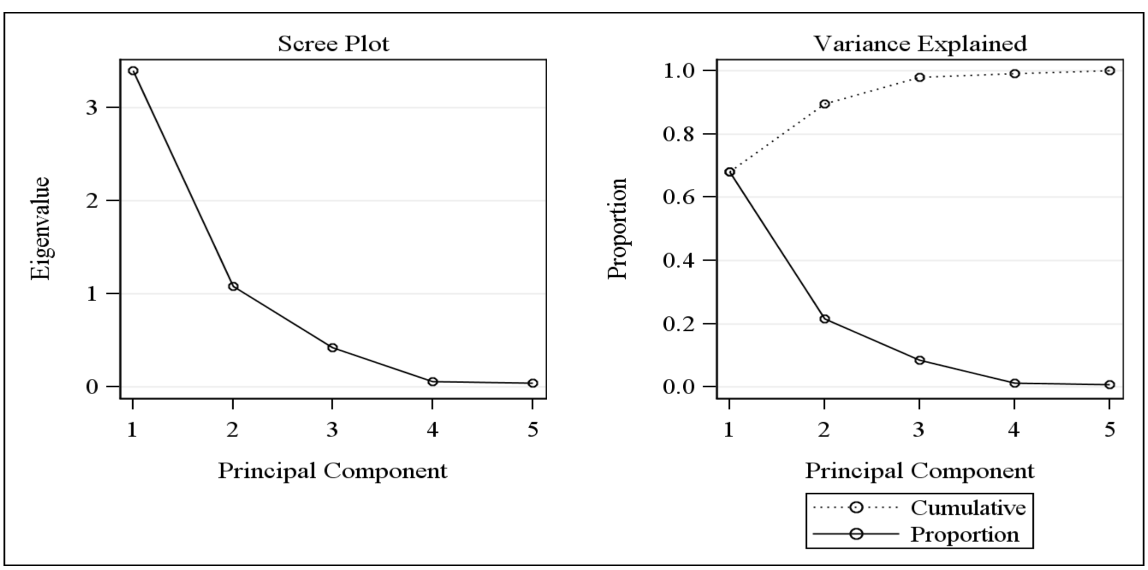

Figure 14 Scree plot and Variance plot of cumulative variance for all vegetables

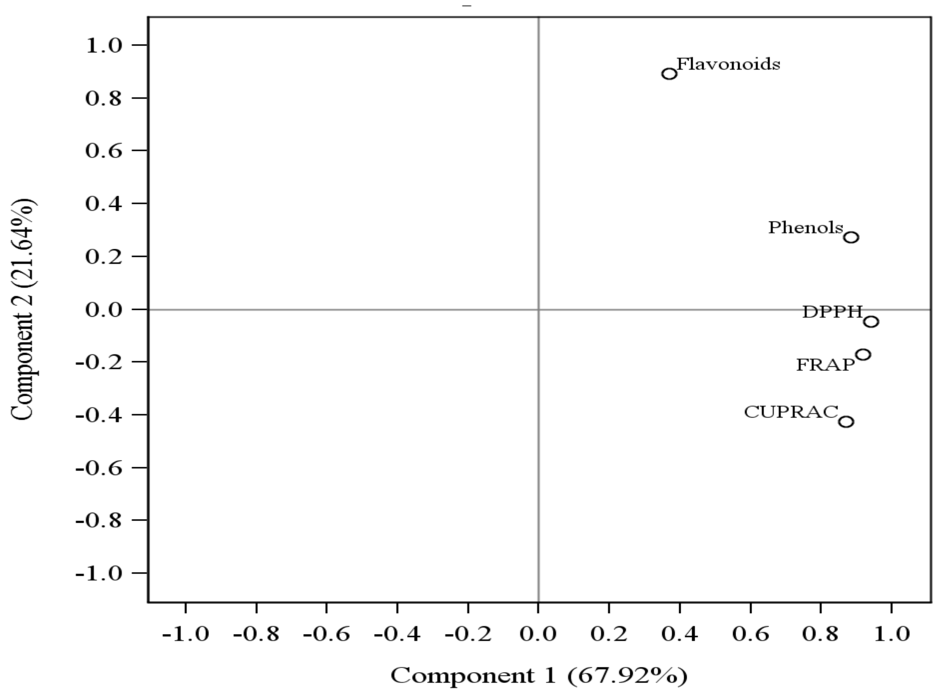

Figure 15 Loadings plot for five different antioxidant assays for fruit data set (i.e. FRAP, CUPRAC, DPPH, phenols, flavonoids) (PC1 vs. PC2)

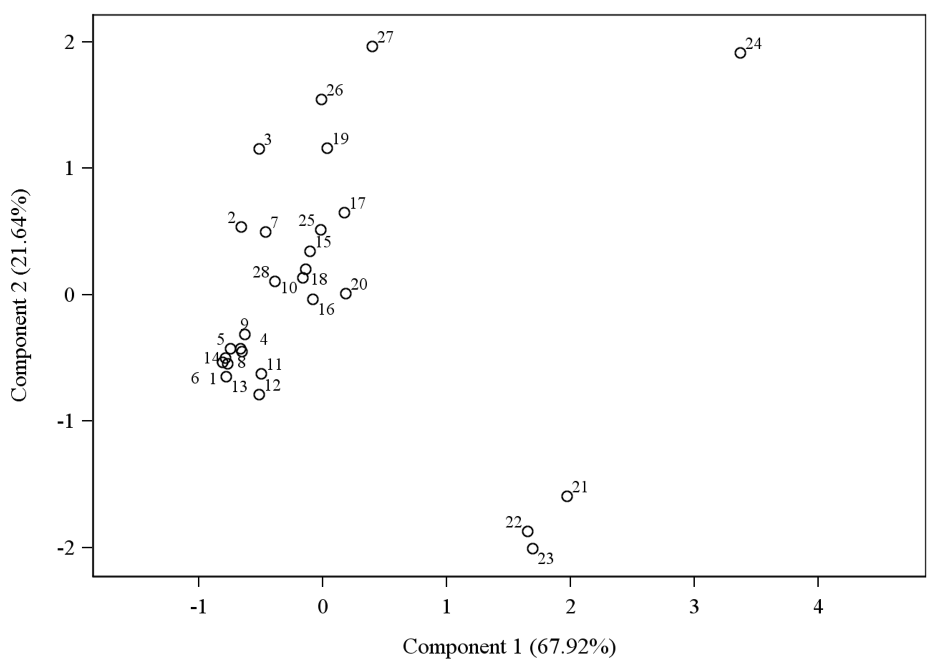

Figure 16 Scores plot for all 28 fruit data under study (PC1 and PC2)

[1-Grape fruit (White); 2-Grape fruit (Pink); 3-Pumelo (White); 4-Pumelo (Red); 5-Sweet orange (Mosambi); 6-Mandarin (Kinnow); 7-Guava; 8-Banana; 9-Pear; 10-Pomegranate (Mridula); 11-Pomegranate (Kadam); 12-Pomegranate (Ganesh); 13-Papaya; 14-Pineapple (Kew); 15- Apple (Royal Delicious); 16-Apple (Red Delicious); 17-Apple (Red spur); 18-Apple (Golden Delicious); 19-Apple (Organic spur); 20-Apple (Vanse Delicious); 21-Grape (Pusa Navrang); 22-Grape (Black muscat); 23-Grape (1612); 24-Aonla; 25-Plum (Santa rosa); 26-Plum (Beauty); 27-Plum (Frontier); 28-Plum (Green gage).

In PC1, all the five assays have positive loadings, whereas in PC2, only Phenols and Flavonoids shows positive loading. The scores plot in **Figure 16** is gives an overview of similarities or difference among various 28 fruit categories. The group of fruits falling under same category are those with subscript 8, 14, 5, 4, 9, 6, 13, 12. Fruit categories 24 are very distinct. 10, 18, 28, 16, 20, 15 again report similarity based on antioxidant contents.

Table 10 Eigen values and % variance

PC	Eigen value	%Variance	Cumulative %
1	2.290	45.790	45.790
2	1.415	28.291	74.081
3	0.668	13.362	87.443
4	0.463	9.259	96.702
5	0.165	3.298	100.000

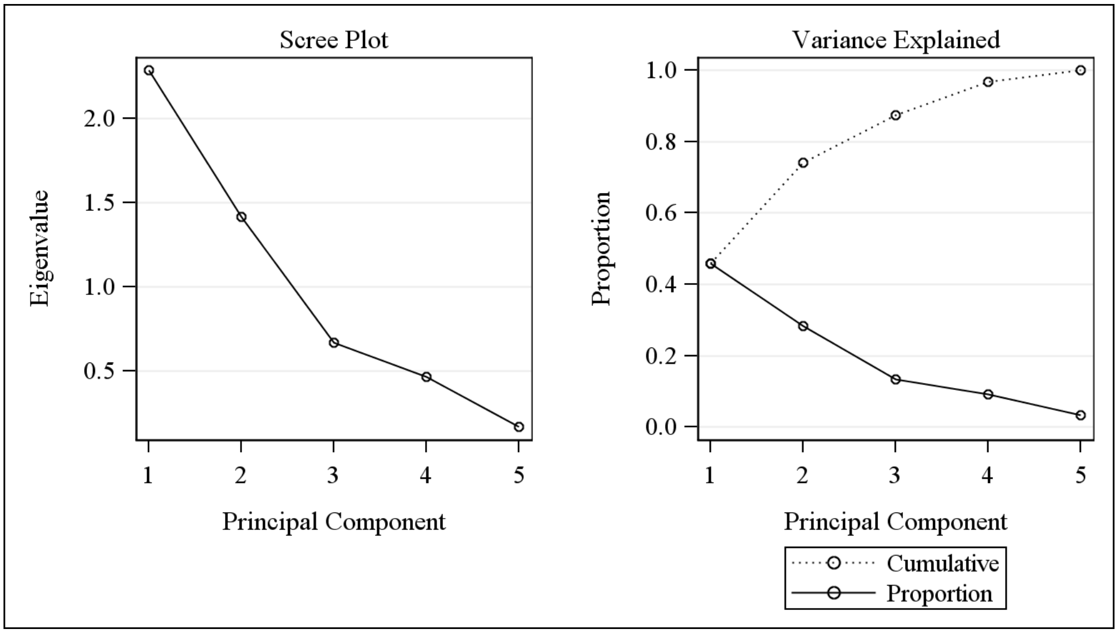

Figure 17 Screen plot and Variance plot of cumulative variance for all cereal crops

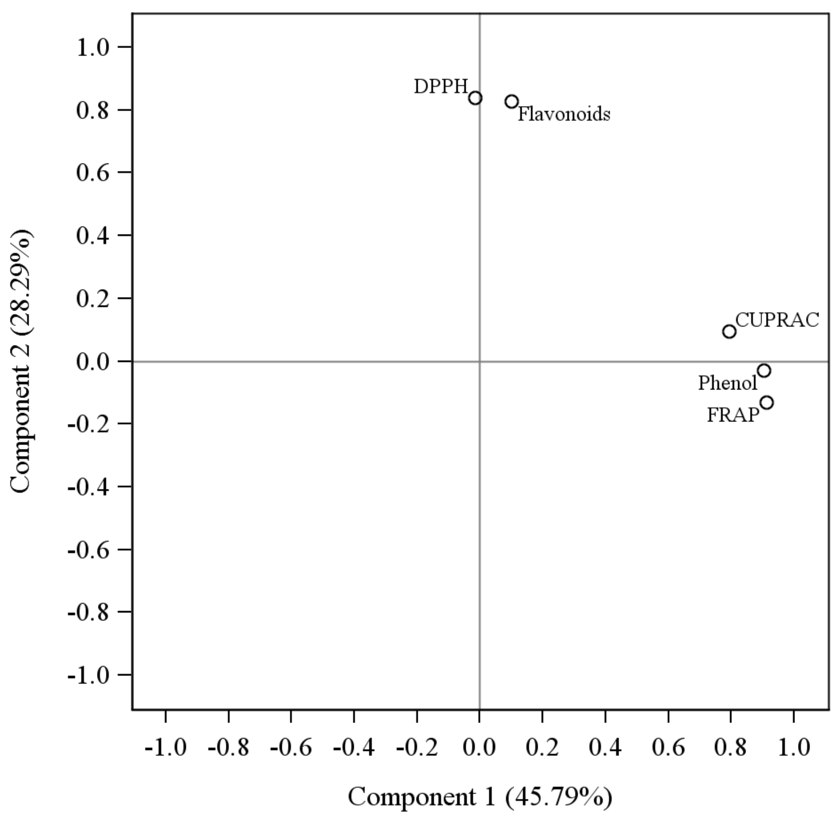

Figure 18 Loadings plot for five different antioxidant assays for cereal crops' data set (i.e. FRAP, CUPRAC, DPPH, Phenols, Flavonoids) (PC1 vs. PC2)

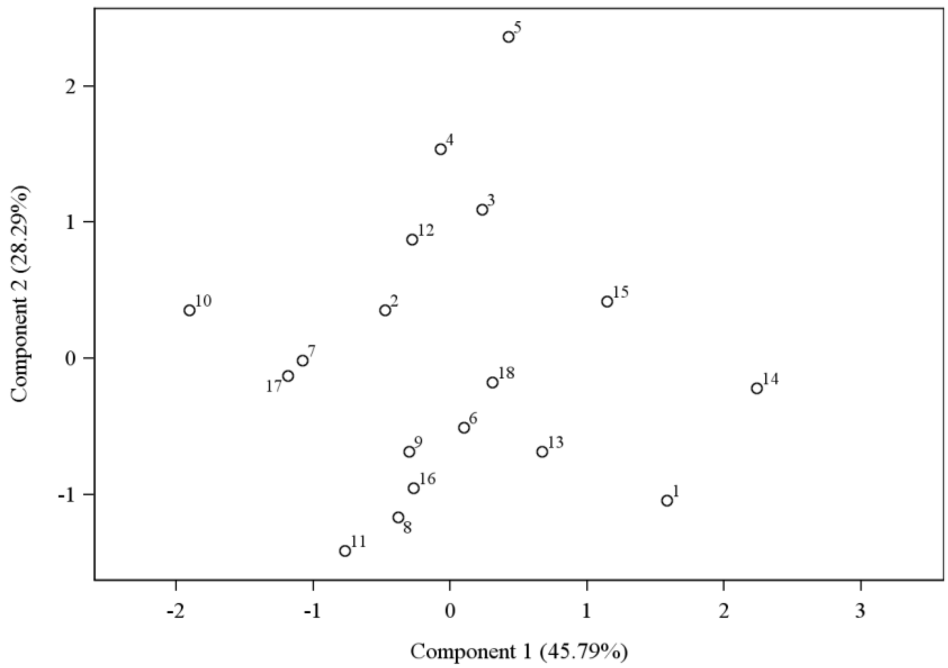

Figure 19 Scores plot for all 18 cereal crops data under study (PC1 and PC2)

[1-Finger millet (Local variety);2-Finger millet (GBU-67);3-Finger millet (GBU-48); 4-Finger millet (GBU-45);5-Finger millet (L-5); 6-Pearl millet;7-Sorghum (CSV18VR); 8-Sorghum (CSV14VR);9-Sorghum (M35-1); 10-Sorghum (Phule Yashoda);11-Barley;12-Oat;13-Red Rice;14-Black gram;15-Kidney bean;16-Soybean (Local variety);17-Soyabean (White);18-Soyabean (Black)

Discussion

Principal component analysis (PCA) is a multivariate technique that analyzes a data table in which observations are described by several inter-correlated quantitative dependent variables. PCA transforms the original, measured variables into new uncorrelated variables called principal components (Cam *et al*., 2009) of similarities and difference between different groups based on AOX and content of antioxidant groups. Its goal is to extract the important information from the table, to represent it as a set of new orthogonal variables called principal components [6]. In PC1, the four assays (FRAP, CUPRAC, phenols and flavonoids) have positive loadings, while DPPH has negative loading. In PC2, only Phenols and FRAP shows negative loading. The scores plot in **Figure 19** is gives an overview of similarities or difference among various 18 cereal categories. The cereal crops shows scattered plotting which shows that the cereal crops do not form good grouping/ clusters to differentiate the cereal categories based on antioxidant contents. Our results are in agreement with results of example, Cam *et al*. (2009).

Applied chemometrics to classify pomegranate juices on the basis of their antioxidant activity and reported the main determinant of this parameter to be cultivar. Wang et al. (2009) [7] carried out principal component analysis to gain an overview of the similarities and differences among 10 algal species and also investigated the relationships between total phenolic content and different antioxidant activity assays. Vallverdu-Queralt *et al* 2011 [8], evaluated phenolic profile and hydrophilic antioxidant capacity as chemotaxonomic markers of tomato varieties. They reported that phenolic compounds and hydrophilic antioxidant capacity accounted for major differences among tomato variety. Overall high correlation between FRAP and phenolics in fruits, vegetables, and cereals reveals that the phenolics are the major determinants for high antioxidant activity.

References

[1] Abdi, H. and Williams, L. J. (2010) Principal component analysis. Wares Computational statistics: 2(2), 433–459

[2] Cam, M., Yasar, H., Durmaz, G. (2009) Classification of eight pomegranate juices based on antioxidant activity measured by four methods. Food Chemistry 112,721–726.

[3] Cook N.C and Samman.S, (1996), Flavonoids-Chemestry, Metabolism, Cardioprotective effects and dietary sources. Nutritional biochemistry. 7, 66-76 5566-U555

[4] Javanmardi J, Stushnoff C, Locke E, Vivanco JM. (2003). Antioxidant activity and total phenolic content of Iranian Ocimum accessions. Food Chem. 83, 547-550.

[5] Painter, J., Rah, J.H., Lee, Y.K., 2002. Comparison of international food guide pictorial representations. Journal of the American Dietetic Association 102, 483–489.

[6] Pennington, J.A.T. and Fisher R.A. (2008), Classification of fruits and vegetables. Journal of Food Composition and Analysis. 22, S23–S31

[7] Rubio CP, Hernandez-Ruiz J, Martinez-Subiela S, Tvarijonaviciute A, and Ceron JJ.(2016) Spectrophotometric assays for total antioxidant capacity (TAC) in dog serum: an update. BMC Veterinary Research: 12: 166.

[8] Vallverdu-Queralt, A., Medina-Remon, A., Martinez-Huelamo, M., Auregui,O., Andres-Lacueva, C. and Lamuela-Raventos, R.M. (2011) Phenolic Profile and Hydrophilic Antioxidant Capacity as Chemotaxonomic Markers of Tomato Varieties. *Journal Agriculture Food Chemistry:* 59 (8), 3994–4001

[9] Wang, T., Jonsdottir,R. and Olafsdottir, G. (2009) Total phenolic compounds, radical scavenging and metal chelation of extracts from Icelandic seaweeds. Food Chemistry: 116, 240–248.

Publication History

Received	22nd June 2017
Revised	05th July 2017
Accepted	07th July 2017
Online	30th July 2017