Investigation of interrelations in chemical substances content of different fruit varieties

A I Butenko¹, M Yu Akimov², E V Zhbanova², V A Koltsov² and N V Protasova¹

¹Michurinsk State Agrarian University, 101, International st., Michurinsk, 393760, Russia
²I.V. Michurin Federal Research Centre, 30, Michurin st., Michurinsk, 393774, Russia

E-mail: but_tolik@mail.ru

Abstract. The main task of the investigation was to study the interrelations and to single out the sustainable clusters which reflect the content of different chemical substances in fruit, but have a tendency of slight correlation between each other. All these should become a base for definition of combination of fruits and berries containing a balanced set of useful compounds. It will help to make a composition of special food production. More than 170 fruit varieties of 13 fruit and berry like crops (apple, pear, apricot, plum, cherry, strawberry, black currant, red currant, viburnum, sea-buckhorn, honeysuckle) were studied for chemical indices content in their fruits and berries. The composition of chemical substances was chosen taking into account the specific features of all the investigated crops (dietary fiber, ascorbic acid, malic acid, citric acid, glucose, fructose, sucrose, Na, K, Ca, Mg, Fe, Cu, Zn). In order to study the correlation of the substances in the content, the principal component analysis was applied. As a result of the research, five clusters of the interrelated characters were singled out, these ones were linked to the first 5 the most important principal components. These components account for 78.23% of total variability of traits. The first component which accounts for 25.46% of total variability of traits is closely correlated with dietary fiber and organic acids content. The second component which accounts for 18.20% of total variability is closely connected with Ca, Fe, Zn content. The third component is correlated with Na, K, fructose and glucose contents.

1. Introduction

The recently increased interest in the chemical composition and antioxidant value of fruits and berries leads to the need to select crops and individual varieties with a higher content of sugars, organic acids, vitamins, trace nutrients and other valuable compounds [1-7].

One of the main factors determining the taste qualities of fruits is the level of sugar content, which correlates with such organoleptic characteristics as the perception of sweetness, acidity, astringency and general tasteful sense. Most consumers prefer sweeter fruits, which is not only a consequence of the higher sugar concentration, but also the balance between acids, volatic aromatic components, and other components, such as polyphenols, which can affect the perception of sweetness [8]. The sugar content of fruits and berries is balanced by the presence of several predominant organic acids, such as citric and malic, as well as phenolic acids, which can give bitter or astringent flavors. The complex of organic acids in fruits and berries makes up the titratable acidity [8]. Currently, the attention of scientists and practitioners is attracted by fiber (food fiber), which was previously considered a ballast substance in food products and tried to clear the food ration of it as much as possible. Fiber is a
complex of complex carbohydrates: fiber (cellulose), hemicellulose, pectins, gum (gum), mucus, as well as non-carbohydrate lignin (a polymer of aromatic alcohols) [9, 10].

Many berry crops, in particular strawberries, black currants, are characterized by a high content of ascorbic acid (vitamin C) and serve as good sources of this vitamin [11-14]. The established level of physiological need for vitamin C in different countries is 45.0-110.0 mg / day [15]. Macro-and microelements are also necessary for human life and health: sodium, potassium, calcium, magnesium, iron, copper, zinc, iodine, etc. [11, 15].

The purpose of our investigation was to study the relationships and identify stable clusters in the content of various chemicals in fruits that are weakly correlated with each other, and on this basis to determine the principle of making combinations of fruits and berries containing a balanced set of useful compounds for composing compositions in a healthy diet and creating specialized food products.

2. Materials and methods

For the investigation, we used data on the accumulation of 16 biochemical substances in the fruits of 170 sorts of 13 fruit crops, including pomeferous (apple, pear), drupaceous (apricot, plum, cherry) and berrylike (strawberry, black currant, red currant, raspberry, gooseberry, viburnum, sea buckthorn, honeysuckle). The composition of the chemical substances characteristic of all the studied cultures was selected (dietary fiber, ascorbic acid, malic acid, citric acid, glucose, fructose, sucrose, Na, K, Ca, Mg, Fe, Cu, Zn). To study the relationship of the content of this group of substances, the analysis of the main components was used [16, 17]. All calculations for the analysis of the main components were carried out using Microsoft Excel spreadsheets and the statistical software package Statistica 6.0.

The main analysis of components is based on the assumption that the variability and interrelationships of the studied set of features can be described fairly accurately by a much smaller number of non-correlating variables—the principal components. In this case, each main component is a certain linear combination of the studied features. The coefficients are found so that the first principal component has the maximum dispersion among all linear combinations of the considered features. The dispersion of the second principal component would be maximal among all linear combinations uncorrelating with the first principal component. The dispersion of the third principal component would be maximal among all linear combinations uncorrelating with the first two principal components. The dispersion of the third principal component would be maximal among all linear combinations uncorrelating with the first and second components, and so on. With this choice of coefficients, all the main components are uncorrelated with each other and their dispersions represent a decreasing (non-increasing) sequence of numbers. The main part of the total dispersion falls on the first few main components, thus it is possible to describe the variability and interrelationships of a large number of features with a small number of main components.

In the "Statistica 6.0" package, the main components are calculated from a matrix of observations or features. Each column of this matrix represents itself values of one feature for all the objects under consideration, and each row of the matrix contains the values of all the features for some object. In our case, the objects are 170 sorts, and the characteristics are 16 chemicals. Thus, each column of the matrix contains parts of the day’s allowance of one of the 16 substances in 100 g of fruits of each variety [18, 19].

The matrix of the weights of the main components, their dispersions, and the matrix of the values of the main components are calculated.

The dispersions are their own values of the correlation or covariance matrices. We have selected a correlation matrix. In this case, the features are normalized in such a way that their average value is zero, and the dispersion is equal to one. The dispersions of the main components can easily be transferred as a percentage of the total dispersion of the features, which is equal to their number in the above-mentioned normalization [20].

The elements of the weight matrix are the correlation coefficients between the corresponding features and the main components. They help to distinguish "clusters" of closely interrelated features. To do this, you should combine in one cluster the features which have the highest absolute weight with one of the main components.
Using the matrix of the main component values, you can make diagrams for clear view of the distribution of sorts in the plane of the first main components.

3. Results and discussion
Using the "Statistica 6.0" package, the weights, dispersions, and values of the main components are obtained. Table 1 shows the weights of the first five main components. If, as mentioned earlier, we select the maximum absolute value of the weight of the main component in each row of the weight matrix, then all 16 features are grouped into 5 clusters, each of which is closely correlated with only one main component. In Table 1, the resulting clusters are grayed out. The weights of the other main components are significantly less than the ones given, so they are not specified. The above five main components account for 78.25% of the total variability of features.

Table 1. Weights of the first five main components

| Substances          | Main components | Daily requirement, mg |
|---------------------|-----------------|-----------------------|
|                     | 1       | 2       | 3       | 4       | 5       | 25.46 | 18.20 | 15.08 | 11.16 | 8.33 |
| Fiber (total)       | 0.950   | -0.041  | -0.141  | 0.026   | 0.084   | 22000 |
| Insoluble fiber     | 0.912   | 0.142   | 0.089   | -0.091  | -0.146  | 20000 |
| Soluble fiber       | 0.484   | -0.342  | -0.466  | 0.218   | 0.441   | 2000  |
| Na                  | 0.339   | -0.157  | 0.603   | -0.132  | 0.403   | 1300  |
| К                   | 0.397   | 0.162   | 0.681   | -0.168  | 0.284   | 2500  |
| Ca                  | 0.267   | 0.733   | -0.090  | -0.543  | -0.071  | 1250  |
| Mg                  | 0.226   | 0.574   | -0.266  | -0.635  | -0.199  | 400   |
| Fe                  | 0.078   | 0.723   | 0.064   | 0.567   | -0.133  | 10    |
| Cu                  | 0.094   | 0.530   | 0.075   | 0.607   | -0.109  | 1     |
| Zn                  | 0.072   | 0.709   | -0.038  | 0.321   | 0.196   | 12    |
| Citric acid         | 0.894   | -0.087  | 0.223   | 0.106   | -0.048  | 500   |
| Malic acid          | 0.553   | -0.133  | 0.550   | -0.040  | 0.094   | 500   |
| Ascorbic acid       | 0.671   | -0.245  | 0.054   | 0.414   | -0.085  | 70    |
| Fructose            | 0.022   | -0.526  | -0.641  | -0.023  | -0.071  | 35000 |
| Glucose             | 0.126   | -0.342  | 0.684   | -0.043  | -0.331  | 25000 |
| Sucrose             | 0.345   | 0.266   | -0.123  | -0.087  | 0.788   | 65000 |
| Dispersion of the main component, % | 25.46 | 18.20 | 15.08 | 11.16 | 8.33 | 400 |

On the plane of the first two main components, fruit crops are arranged in compact groups (figure 1). This indicates that the values of the main components of each culture vary within certain limits. The highest values of the first component were found in black currant, the lowest in plum and cherry. The highest values of the other main components are already represented in different cultures. The second component has the highest values in apricot varieties, the lowest-in pear varieties. By selecting one variety for each component, we get a combination of varieties in which all five components are shown to the maximum extent.

The Lel apricot variety has high values in the second and fifth components. Also, the strawberry variety Dukat can represent the second and fourth components. For the third components, we take a variety of gooseberry Mashek, and for the first – grade black currant Vigorous. Choosing 100 g of fruits of each of the 4 varieties, we get the total percentage of the daily requirement of each of the 16 chemicals (Table 2). As can be seen from Table 2, the percentage of substances such as sodium, zinc,
sucrose is very low, and iron, copper and acids are contained in large quantities. This is due to the low or high content of substances in the entire group of varieties.

Figure 1. Arrangement of varieties on the plane of the first two main components

Table 2. The content of biochemical substances (percentage of the daily requirement) in four varieties (per 100 g of fruit of each variety)

| Indicator I        | apricot Lel | strawberry Ducat | gooseberry Masheka | black currant Vigorous | Percentage of daily requirement |
|--------------------|-------------|-------------------|--------------------|------------------------|--------------------------------|
| Fiber (total)      | 10.5        | 9.5               | 13.5               | 24.0                   | 57.5                           |
| Insoluble fiber    | 4.0         | 4.5               | 8.5                | 14.0                   | 31.0                           |
| Soluble fiber      | 65.0        | 50.0              | 50.0               | 100.0                  | 265.0                          |
| Na                 | 2.1         | 0.1               | 2.5                | 2.4                    | 7.1                            |
| K                  | 26.1        | 3.0               | 21.0               | 17.4                   | 67.5                           |
| Ca                 | 32.2        | 2.8               | 3.2                | 2.7                    | 40.9                           |
| Mg                 | 43.0        | 1.5               | 6.6                | 6.0                    | 57.1                           |
| Fe                 | 35.0        | 73.0              | 6.1                | 14.7                   | 128.8                          |
| Cu                 | 29.0        | 230.0             | 3.6                | 7.2                    | 269.8                          |
| Zn                 | 9.7         | 6.5               | 1.8                | 1.6                    | 19.6                           |
| Citric acid        | 43.2        | 118.6             | 267.4              | 517.4                  | 946.6                          |
| Malic acid         | 82.2        | 52.4              | 280.0              | 32.0                   | 446.6                          |
| Ascorbic acid      | 1.9         | 64.6              | 48.9               | 214.9                  | 330.3                          |
| Fructose           | 2.9         | 6.6               | 2.9                | 10.9                   | 23.3                           |
| Glucose            | 7.6         | 9.6               | 18.4               | 6.4                    | 42.0                           |
| Sucrose            | 9.1         | 0.3               | 0.1                | 0.6                    | 10.1                           |
4. Conclusion
1. Using the analysis of the main components, we show that 16 characteristics reflecting the content of various chemicals in the fruits of the studied varieties of fruit crops are grouped into five clusters that closely correlate with the first five most significant main components.
2. On the plane of the first two components, fruit crops are arranged in compact groups, which indicates consistency in the manifestation of complexes of signs in different varieties of the same crop.
3. If you choose varieties with high values of different components, so that in a minimum set of 4-5 varieties all the components are represented, then this set will contain the best combination of chemical characteristics that can be obtained from the considered set of varieties.

References
[1] Nikitin A, Kuzicheva N and Karamnova N 2019 Establishing efficient conditions for agriculture development International Journal of Recent Technology and Engineering 8(2) 1-6
[2] Akimov M Yu, Makarov V N and Zhbanova E V 2019 The role of fruits and berries in providing humans with vital biologically active substances Achievements of science and technology of the agroindustrial Complex 33(2) 56-60
[3] Chupakhina G N, Maslennikov P V, Skrypnik L N, Chupakhina N Yu and Feduraev P V 2016 Antioxidant properties of cultivated plants of the Kaliningrad region: monograph (Kaliningrad: I. Kant BFU Publishing House) 145 p
[4] Kirina I B, Belosokhov F G, Titova L V, Suraykina I A and Pulpitow V F 2020 Biochemical assessment of berry crops as a source of production of functional food products IOP Conference Series: Earth and Environmental Science 548(8) 082068
[5] Skrovankova S, Sumczynski D, Mlcek J, Jurikova T and Sochor J 2015 Bioactive Compounds and Antioxidant Activity in Different Types of Berries Int. J. Mol. Sci. 16(10) 24673-24706
[6] Skinner M and Hunter D 2013 Bioactives in Fruits: Health Benefits and Functional Foods (John Wiley & Sons) 542 p
[7] Battino M, Beekwilder J, Denoyes-Rothan B, Laimer M, McDougall G J and Mezzetti B 2009 Bioactive compounds in berries relevant to human health Nutrition Reviews 67 145-150
[8] Blinnikova O M, Babushkin V A, Eliseeva L G and Usova G S 2020 Modeling a formulation and assessment of the consumer properties of the special purpose starch drink Sarhad Journal of Agriculture 36(3) 939-948
[9] Bronovets I V 2015 Fiber is an important component of a balanced healthy diet Medical news 07 46-48
[10] Baigarin E K 2012 The investigation of the content of fiber in domestic food products and their impact on the digestibility of macronutrients: abstract of candidate dissertation (Moscow) 22 p
[11] Pastushkova E V, Zavorokhina N V and Vyatkin A V 2014 Vegetable raw materials as a source of functional food ingredients Bulletin of SUSU. Series: Food and Biotechnologies 4(4) 105-113
[12] Akimov M Yu, Zhbanova E V, Lukyanchuk I V and Lyzhin A S 2019 Characteristics of the strawberry varietal fund by chemical composition and antioxidant value of fruits in the conditions of the Central Chernozem region Bulletin 01 56-60
[13] Zhbanova E V 2012 Evaluation of the biochemical composition of black currant varieties taking into account their genetic origin Fruit and berry growing in Russia 38(1) 139-144
[14] Szajdek A and Borowska E J 2009 Bioactive Compounds and Health-Promoting Properties of Berry Fruits: A Review Plant Foods for Human Nutrition 63 147-156
[15] 2009 Norms of physiological needs in energy and food substances for various groups of the population of the Russian Federation. Methodological recommendations MP 2.3.1.2432-08: (Moscow: Federal Center for Hygiene and Epidemiology of Rospotrebnadzor) 36 p
[16] Kendall M and Stewart A 1976 Multidimensional statistical analysis and time series (Moscow: Main edition of the physical and Mathematical literature of the Nauka Publishing House) 736 p
[17] Blinnikova O M, Babushkin V A, Akindinov V V, Perfilova O V and Novikova I M 2020 Production technology and mathematical method for modeling the formulation of fruit and jelly
candies enriched with collagen IOP Conf. Series: Materials Science and Engineering 919 052036. doi:10.1088/1757-899X/919/5/052036

[18]Perfilova O V, Akishin D V, Vinnitskaya V F, Danilin S I and Olikainen O V 2020 Use of vegetable and fruit powder in the production technology of functional food snacks IOP Conference Series: Earth and Environmental Science 548(8) 082071

[19]Kartechina N V, Bobrovich L V, Nikonorova L I, Pchelinceva N V and Abaluev R N 2020 Practical application of variance analysis of four-factor experience data as a technology of scientific research IOP Conference Series: Materials Science and Engineering 919(5) 052030. doi:10.1088/1757-899X/919/5/052030

[20]Suhareva T, Sergienko I, Kutsova A and Ratushny A 2019 Mathematical planning when choosing rational dosages of ingredients for adjusting the composition of bakery products International Journal of Engineering and Advanced Technology 8(6) 4562-4565