Factor analysis and cluster analysis research on the relationship between the trace elements of flowers of Gentianaceae plants and the traditional Chinese medicine efficacy from Qinghai-Tibet Plateau

Libing Zhou
Guizhou Institute of Technology, Guiyang 550003, China

Corresponding author: zhoulibing7802@sina.com

Abstract. Flowers of Gentianaceae plants have historically been used in treating gastritis, dysuria, pruritus vulvae and detoxification. Gentianaceae plants have medicinal applications to treat a wide range of diseases such as rheumatic arthritis, pneumonia along with cough, chronic gastritis, et al. Therefore, the research on the relationship between human health and trace elements of flowers of Gentianaceae plants is of great significance. These studies provide the scientific base and theoretical foundation for the future large-scale rational relation development of Gentianaceae plants resources as well as the relationship between the race elements of them and traditional Chinese medicine efficacy.

1. Introduction
Bile (Gentiana) belongs to Kelon gentian plants around the world. There are about more than 400 kinds of plants around the world, distributed in Europe, Asia and northern Australia, New Zealand and the north, and along the Andes of Cape Horn and North America. China has 247 species distributed throughout the country, mainly in the mountains, rocky beach, alpine meadows and shrubs. There are abundant resources of Gentianaceae plants from Qinghai-Tibet Plateau. Flowers of Gentianaceae plants have historically been used in treating gastritis, dysuria, pruritus vulvae and detoxification. Gentianaceae plants have medicinal applications to treat a wide range of diseases such as rheumatic arthritis, pneumonia along with cough, chronic gastritis, et al. Therefore, the research on the relationship between human health and trace elements of flowers of Gentianaceae plants is of great significance [1-5].

Ning Wang pointed out that the chemical composition of gentian genus plant was divided into four categories: crack iridoid glycosides, xanthones glycosides, flavonoids and triterpenoids. In addition, it also contained alkaloids, volatile oil, sugar and a variety of ingredients [6]. Using high performance liquid chromatography (HPLC) method three kinds of active ingredients including oleanolic acid, mango glycosides and drugs flavin were isolated and analyzed [7] by Fengzu Hu, etc. in nine kinds of Gentianaceae plants from Qinghai-Tibet Plateau, the study has been of great significance to the research on the biodiversity of the Qinghai Tibet Plateau, the discovery and expansion of the new drug. The contents of six kinds of trace elements in flowers of Gentianaceae plants were measured by Xingwang Zhang [8], etc., the study has provided a certain basis for reasonable medicinal herbs of Gentianaceae plants flowers, further development and utilization. In summary, so far, the organic chemical composition and inorganic chemical component contents of
gentianaceae plants were measured that has been reported at home and abroad while the research on the relationship between race elements of them and traditional Chinese medicine efficacy would be much less using factor analysis and cluster analysis methods.

In this paper, using the world's most advanced MATLAB software, factor analysis and cluster analysis methods in the chemometric methods dig out hidden information between the trace elements of flowers of Gentianaceae resource and traditional Chinese medicine efficacy. These studies provide the scientific base and theoretical foundation for the future large-scale rational relation development of Gentianaceae plants resources as well as the relationship between the race elements and traditional Chinese medicine efficacy. Therefore, it is important practical significance to study the relationship between the trace elements of flowers of Gentianaceae plants and traditional Chinese medicine efficacy for the human.

2. Materials and methods

2.1 Materials

This paper selected 6 kinds of trace elements (including Ca, Mg, Cu, Fe, Mn and Zn) of flowers of Gentianaceae plants) including Gentian nubigena Edgew, tube flowers Gentiana flowers, linear Gentian flowers, yellow tube Gentiana flowers, manual twist Gentiana flowers and hosta blue Gentian flowers (in Qinghai-Tibet Plateau as an analysis sample (n = 6, said six parallel determination, unit: Mg/kg), the original data from the literature [8]. Raw data of trace elements flowers of Gentianaceae plants were shown in table 1.

| sample                        | Ca  | Mg  | Cu  | Fe   | Mn  | Zn  |
|-------------------------------|-----|-----|-----|------|-----|-----|
| Gentian nubigena Edgew        | 740.28 | 2305.26 | 5.53 | 939.61 | 43.28 | 35.15 |
| Tube flowers Gentiana flowers | 771.71 | 2007.38 | 2.98 | 150.69 | 14.66 | 33.85 |
| linear Gentian flowers        | 1756.06 | 3311.24 | 3.64 | 321.3 | 27.79 | 48.07 |
| yellow tube Gentiana flowers  | 1402.65 | 3187.14 | 1.49 | 493.06 | 21.92 | 23.74 |
| manual twist Gentiana flowers | 3008.61 | 4042.11 | 3.24 | 316.58 | 38.07 | 33.73 |
| hosta blue Gentian flowers    | 1048.77 | 2433.21 | 3.9  | 433.58 | 31.28 | 50.63 |

2.2 Factor analysis

Factor analysis is to seek a few factors to reflect comprehensively most of the information of all the factors, variable although more primitive variables are less, but the amount of information contains in the accounts for more than 85% of the original information. These new variables have high reliability about analysis essence of credibility, and they are not correlated with each other, eliminating the multicollinearity.

Factor analysis steps: Combining with the software MATLAB2010 data have been analyzed by factor analysis. The standardization of original data, namely the same variables minus the mean, and then are divided by the standard deviation in order to eliminate the dimension influence between the original data, so that the characteristics of the standardized data are comparable. In one problem there are n individuals, P indexes have been measured for each individual to avoid the influence of the index dimension, the original data for Standardization: \( Y_{ij} = (x_{ij} - \bar{x}_j) / s_j \), where \( Y_{ij} \) is standardized data after treatment , \( x_{ij} \) for the original data, \( \bar{x}_j \) for the mean value of the j-th index for n samples, \( s_j \) for the standard deviation of the sample; Determine the correlation coefficient matrix; Determine the characteristic root and variance contribution rate of the correlation coefficient matrix; The rotated factor loading matrix analysis; Calculate factor score and the comprehensive factor score.

2.3 Cluster analysis.
Cluster analysis is to classify the research object, put all the case classification in different classes, so that individuals of the same class have greater similarity, individuals of different categories have greater differences. MATLAB2010 software has been used to analyze the data. After extracting the main factors, the main factor data will be analyzed using hierarchical clustering method.

3. Results

3.1 Factor analysis process

The raw data have been standardized, standardized data show in table 2; Combining with MATLAB2010, the factor analysis methods have been used to analyze the data of trace elements and obtain the correlation coefficient matrix of trace elements in flowers of Gentianaceae from Qinghai-Tibet Plateau, the results are shown in table 3.

| sample                          | ZCa   | ZMg   | ZCu   | ZFe   | ZMn   | ZZn   |
|---------------------------------|-------|-------|-------|-------|-------|-------|
| Gentian nubigena Edgew          | -0.8362 | -0.7532 | 1.5694 | 1.8389 | 1.3173 | -0.2363 |
| Tube flowers Gentiana flowers   | -0.7994 | -1.1428 | -0.3670 | -1.0793 | -1.4187 | -0.3655 |
| linear Gentian flowers          | 0.3527 | 0.5627 | 0.1342 | -0.4482 | -0.1635 | 1.0474 |
| yellow tube Gentiana flowers    | -0.0609 | 0.4004 | -1.4985 | 0.1871 | -0.7246 | -1.3699 |
| manual twist Gentiana flowers   | 1.8188 | 1.5187 | -0.1696 | -0.4657 | 0.8193 | -0.3774 |
| hosta blue Gentian flowers      | -0.4751 | -0.5858 | 0.3316 | -0.0329 | 0.1702 | 1.3017 |

3.2 Characteristic root and variance contribution rate of correlation coefficient

The characteristic root and variance contribution rate of correlation coefficient were shown in table 4. Table 4 showed, factor analysis was adopted to extract 3 major factors, and the total contribution was 96.301%. The pre-election three main factor accumulative contribution rate reached 96.301%, the accumulative contribution rate of the 3 major factors to total variation accounted for over 85%, maintaining most of information of characters. So this paper chose the pre-election three main factors which represented 96.301% in information of 6 kinds of trace elements( including Ca, Mg, Cu, Fe, Mn and Zn) of flowers of Gentianaceae plants(including Gentian nubigena Edgew,tube flowers Gentiana flowers,linear Gentian flowers,yellow tube Gentiana flowers,manual twist Gentiana flowers and hosta blue Gentian flowers)in Qinghai-Tibet Plateau.

| Table 3. Correlation coefficient matrix
| Ca   | Mg   | Cu   | Fe   | Mn   | Zn   |
|------|------|------|------|------|------|
| Ca   | 1.000 | -0.269 | -0.335 | 0.286 | -0.072 |
| Mg   | 1.000 | -0.348 | -0.203 | 0.278 | -0.140 |
| Cu   | 1.000 | 0.602 | 0.714 | 0.490 |
| Fe   | 1.000 | 0.701 | -0.127 |
| Mn   | 1.000 | 0.188 |
| Zn   | 1.000 |

| Table 4. Characteristic root and variance contribution rate of correlation coefficient
| main factor | characteristic root | variance contribution rate % | the cumulative contribution rate % |
|-------------|--------------------|-----------------------------|----------------------------------|
| 1           | 2.576              | 42.939                      | 42.939                           |
| 2           | 2.039              | 33.977                      | 76.916                           |
| 3           | 1.163              | 19.385                      | 96.301                           |
| 4           | 0.190              | 3.161                       | 99.462                           |
| ...         | ...                | ...                         | ...                              |
3.3 Factor rotation
On the basis of principal component analysis, the factor load matrix has been 4 times to maximize orthogonal rotation, so that the load of each original variable in each common factor was rotated to be complete polarized. One of the purposes of factor analysis is to identify the factors that have practical significance, the factor loading matrix after rotation was shown in table 5. And rotated factor loadings Mn value reached the maximum, which indicated that the greatest influence trace element, Mn in flowers of Gentianaceae plants from the Qinghai Tibet Plateau.

Table 5. The factor loading matrix after rotation

| variable      | 1       | 2       | 3       |
|---------------|---------|---------|---------|
| Ca            | -0.062  | 0.990   | 0.019   |
| Mg            | -0.025  | 0.980   | -0.107  |
| Cu            | 0.801   | -0.258  | 0.469   |
| Fe            | 0.900   | -0.249  | -0.276  |
| Mn            | 0.925*  | 0.337   | 0.135   |
| Zn            | 0.089   | -0.062  | 0.972   |

3.4 Factor scores
The program automatically ran to get factor scores and comprehensive factor scores by professional software MATLAB2010 programming, operation results were shown in table 6.

Table 6. Factor and comprehensive factor scores

| sample                              | F1       | F2       | F3       | F         | ranking |
|-------------------------------------|----------|----------|----------|-----------|---------|
| Gentian nubigena Edgew              | 1.7899   | -0.7548  | -0.3801  | 0.4551    | 2       |
| Tube flowers Gentiana flowers       | -1.1768  | -1.0336  | -0.0562  | -0.9007   | 6       |
| linear Gentian flowers              | -0.2775  | 0.4365   | 1.0149   | 0.2346    | 3       |
| yellow tube Gentiana flowers        | -0.5488  | 0.0698   | -1.5969  | -0.5415   | 5       |
| manual twist Gentiana flowers       | 0.1793   | 1.7238   | -0.1047  | 0.6672    | 1       |
| hosta blue Gentian flowers          | 0.0339   | -0.4416  | 1.1230   | 0.0853    | 4       |

3.5 Cluster analysis
Cluster analysis is a method of mathematical statistics, the original data have been normalized, and then 6 kinds of trace elements in flowers of Gentianaceae plants (including Gentian nubigena Edgew, tube flowers Gentiana flowers, linear Gentian flowers, yellow tube Gentiana flowers, manual twist Gentiana flowers and hosta blue Gentian flowers) have been evaluated and classified by the methods of quantitative analysis of 3 main factors in clustering in Qinghai-Tibet Plateau. Based on the main factor clustering analysis, each two samples with Average linkage linked with the method of euclidean distance measurement, the sequence was plotted in figure 1. The tree graph of cluster analysis was shown in figure 1.

4. Discussion
Described in table 3, the correlation coefficient can be shown, the higher correlation coefficients between any two variables, each variable at least with one or more of the other variables have higher correlation coefficients. Therefore, it is suitable to study the relationship between variables with factor analysis. And it can be known, Ca and Mg was significant correlation, Mn and Cu, significantly correlated Fe.

The first commonality factor F1 had greater load on the indexes of Cu, Fe and Mn, as shown in table 5. Cu is one of the essential trace elements in human body, is involved in human life activities and is related to the metabolism of blood. Modern medicine has found that people and animals need Cu manufacture red blood cells and hemoglobin, and that Cu combined with certain drugs will have anti-rheumatism effect. The lack of Cu may cause bone mineralization, which can easily lead to osteoporosis in the elderly. Fe is one of the essential trace elements in human body, is the "building materials" for human development, is an important component of hemoglobin and a kind of important element to transport oxygen and oxygen exchange in the blood, is a component of many enzymes and is an activator of the redox enzyme. Mn is recognized as a kind of tumor suppressor element, is an integral component of the enzyme in the human body, and is involved in hematopoiesis process, redox, calcium and phosphorus metabolism. It can promote growth, development, the normal growth and the formation of bones and connective tissue, and maintain the blood pressure value, blood sugar and lipid levels in normal, and maintenance function of the central nervous system, and strengthen the function of the brain, improve memory, and also can alleviate the symptoms of fatigue and nervous, and promote physical and mental health. Table 5 showed, the second commonality factor F2 had greater load on the indexes of Ca and Mg. Mg, Ca have a variety of special physiological functions, can activate a variety of enzymes in the human body, inhibit the abnormal nerve excitability, and maintain the stability of the structure of nucleic acids, proteins involved in the synthesis, muscle contraction and temperature regulation in the human body. Table 5 showed, the third commonality factor F3 had greater load on the index of Zn. Zn is the main component of anti free radical in the human body and has the effect of inhibiting lipid peroxide damage, antioxidation on tissue and anti-aging. Zn has many pharmacological activities such as participating in the synthesis and activation of more than 200 kinds of enzymes, proteins, DNA and RNA synthesis in the human body, regulating body fluid acid-base degree, promoting collagen generation and making hair, skin and nails, etc. healthy growth and can help enhance memory and improve intelligence. Fe, Mn and Zn are the main components of anti free radicals in the human body and can inhibit the damage of lipid peroxidation and antioxidation on tissue, and they all have anti-aging effect.

Table 6 showed, six kinds of trace elements including Ca, Mg, Cu, Fe, Mn and Zn in flowers of Gentianaceae plants (including Gentian nubigena Edgew, tube flowers Gentiana flowers, linear Gentian flowers, yellow tube Gentiana flowers, manual twist Gentiana flowers and hosta blue Gentian flowers (in Qinghai-Tibet Plateau in the order were manual twist Gentiana flowers > Gentian nubigena Edgew > linear Gentian flowers > hosta blue Gentian flowers > yellow tube Gentiana flowers > tube flowers Gentiana flowers.

Figure 1 showed that the sample hierarchical cluster analysis clustered into 3 classes, linear Gentian flowers, hosta blue Gentian flowers and tube flowers Gentiana flowers were a class, as a class yellow tube Gentiana flowers and manual twist Gentiana flowers, Gentian nubigena Edgew as a class. Figure 1 showed trace elements spectrum in flowers there were similarity and provided the basis for further pharmacological study of flowers of Gentianaceae plants in Qinghai Tibet Plateau. On the basis of factor analysis it was feasible to clustering classification and identification for flowers’ samples of Gentianaceae plants in Qinghai Tibet Plateau. The conclusion is objective, credible and persuasive.

5. Conclusions
From the factor analysis and cluster analysis, the results showed that the contents of trace elements in manual twist Gentiana flowers were the highest, which indicated that Gentianaceae plants flowers as medicinal plants in Qinghai Tibet Plateau, in terms of content of trace elements, the best quality
manual twist Gentiana flowers, followed by Gentian nubigena Edgew, yellow tube Gentiana flowers, tube flowers Gentiana flowers worst.

Using chemometrics methods the innovation of this paper was to elucidate the contents of six kinds of trace elements including Ca, Mg, Cu, Fe, Mn and Zn in flowers of Gentianaceae plants in the high and low, one side reflected the quality of corresponding medicinal materials was good or bad, dig out the implied relationship between trace elements and traditional Chinese medicine of flowers of Gentianaceae plants to make up for gaps in previous studies of Gentianaceae plants in this field.

Combining the world's most advanced MATLAB software, factor analysis and cluster analysis methods in the chemometric methods were used to comprehensively analyze the data of trace elements and evaluate 6 kinds of trace elements and find out the relationship between the 6 kinds of trace elements including Ca, Mg, Cu, Fe, Mn and Zn in flowers of Gentianaceae plants in Qinghai-Tibet Plateau, which dig out hidden information between the trace elements of flowers of Gentianaceae plants resource and traditional Chinese medicine efficacy. Six kinds of trace elements in flowers of Gentianaceae plants in the order were manual twist Gentiana flowers >Gentian nubigena Edgew>linear Gentian flowers>hosta blue Gentian flowers>yellow tube Gentiana flowers>tube flowers Gentiana flowers.

These studies provide the scientific base and theoretical foundation for the future large-scale rational relation development of Gentianaceae plants resources as well as the relationship between the race elements and traditional Chinese medicine efficacy.

References
[1] Kumar, Vijay, Van Staden, Johannes 2016 A Review of Swertia chirayita (Gentianaceae) as a Traditional Medicinal Plant. Frontiers in Pharmacology 6 1-14
[2] Kumar V, Van Staden J 2016 A Review of Swertia chirayita (Gentianaceae) as a Traditional Medicinal Plant. Front Pharmacol 6 308
[3] Do, T., Popov, S., Marekov, N., et al. 1987 Iridoids from Gentianaceae Plants Growing in Bulgaria. Planta Medica 6 580
[4] Chun-Ying Xue, De-Zhu Li 2011 Use of DNA barcode sensu lato to identify traditional Tibetan medicinal plant Gentianopsis paludosa (Gentianaceae). Journal of Systematics and Evolution 49 267-270
[5] Zhang ZR, Yang J, Sun Y, et al. 2011 A set of novel microsatellite markers developed for the traditional Tibetan medicinal plant Halenia elliptica (Gentianaceae). American Journal of Botany 98,e173-5
[6] Ning Wang & Le Zhou 2001 Study in chemical components of genus Gentiana,Shanxi Forest Science and Technology, China 3 71-78
[7] Fengzu Hu, Yali Song, Mei Liu, et al. 2003 Analysis of Medicinal Bioactive Composition of Gentianaceae in Qinghai Tibet Plateau by High Performance Liquid Chromatography,Chinese Journal of Chromatography21 63-65
[8] Xingwang Zhang, Chen Chen, Ruitao Yu,et al. 2010, Determination of Trace Elements in Six Kinds of Flowers of Gentianaceae by Microwave Digestion-FAAS,Chinese Journal of Spectroscopy Laboratory2 1610-1613