

# 对藏族、畲族、朝鲜族人发中 微量元素的聚类分析研究

杨若明<sup>1</sup>, 金继红<sup>2</sup>

(1. 中央民族大学 生物化学系; 2. 信息与计算科学系, 北京 100081)

**摘要:** 用 ICP 发射光谱法对中国藏族、畲族和朝鲜族青年人体头发中的钙、镁、铁、锰、铜、锌、镉 7 种元素的含量进行了测定, 用聚类分析方法进行了比较研究。研究表明, 在 3 个民族的人发之间, 7 种元素含量的综合水平存在着较显著的差别。

**关键词:** 微量元素; ICP-AES; 头发; 聚类分析; 藏族; 畲族; 朝鲜族

**中图分类号:** Q983.62      **文献标识码:** A      **文章编号:** 1000-3193 (2001) 02-0157-05

人体内微量元素的代谢变化及其营养状态与头发中的微量元素含量有关<sup>[1]</sup>。因此, 头发中微量元素的测定已经被广泛应用于许多领域的研究工作中。我国是一个多民族的国家, 为了探讨各民族人体头发中微量元素的含量是否相同, 作者曾对我国土家族和瑶族青年人体头发中微量元素的含量进行过测定和聚类分析<sup>[2]</sup>, 两个民族之间在 7 种元素的综合水平上存在着明显差别。本文用等离子体原子发射光谱法 (ICP-AES) 对我国藏族、畲族、朝鲜族、维吾尔族等不同民族的青年头发中的钙、镁、铁、锰、铜、锌、铬等 7 种微量元素进行了测定, 并且对 3 个民族 (藏族、畲族、朝鲜族) 从 7 种元素含量的整体水平上用聚类分析方法进行了比较研究<sup>[3-5]</sup>, 发现藏族、畲族、朝鲜族样本基本上各自聚为 3 类; 又将一个维吾尔族样品的测定结果与上述 3 个民族的数据再次进行聚类分析, 结果维族数据不与上述 3 个民族中任何一个民族的数据聚成一类; 两次聚类结果都说明: 在微量元素含量的总体水平上, 民族之间区分比较明显。

## 1 对象与方法

**样品采集** 中央民族大学一年级新生中来自该民族聚居区的本民族男、女青年, 年龄在 17 至 21 岁, 身体健康。采样时用不锈钢剪刀剪枕部头发 2—3g。

**测定方法** 将头发用洗涤剂浸泡, 并用自来水和去离子水洗净、烘干<sup>[1]</sup>。精密称取一定量烘干后的头发, 用混合酸进行消化处理后, 用美国 Pseries 1000 型 ICP-AES 仪, 测定了发样中 7 种元素的含量, 数据用计算机处理。

收稿日期: 2000-10-26; 定稿日期: 2001-01-15

作者简介: 杨若明 (1947—), 女, 北京人, 中央民族大学民族地区环境资源研究所副所长, 生物化学系副教授, 理学硕士, 主要从事微量元素营养生态学和民族地区天然产物开发、利用的研究。

## 2 结果与讨论

### 2.1 3 个民族 7 种元素的平均值

藏族、畲族和朝鲜族 3 个民族人体头发中 7 种元素含量的平均值和置信区间见表 1。

表 1 3 个民族 7 种元素的平均值和标准偏差 (n=10, p=0.10,)

Average and standard deviation of seven elements from three nations

民族	测量统计项目	Ca (ppm)	Mg (ppm)	Fe (ppm)	Mn (ppm)	Cu (ppm)	Zn (ppm)	Cr (ppm)
藏族	平均值	1136.8	116.91	50.44	1.23	10.66	172.1	3.89
	标准偏差	508.31	39.55	34.47	0.53	1.65	26.90	0.47
畲族	平均值	812.8	87.64	10.83	1.20	8.26	137.16	7.36
	标准偏差	126.38	32.79	2.89	0.76	1.45	23.90	0.99
朝鲜族	平均值	783.1	122.41	48.4	2.45	9.88	134.23	4.20
	标准偏差	432.7	56.91	30.33	0.96	1.85	31.27	2.06

### 2.2 3 个民族 7 种元素的聚类分析

#### 2.2.1 数据的标准化

由于不同元素在含量的数量级上有差别,例如钙含量在 2000 ppm—300 ppm 范围,而铬含量仅在 1—10ppm 范围。为了分类,可以认为样品中钙含量 100ppm 数量级的差别与铬含量 0.1ppm 的数量级差别同等重要。可是如果直接用原始数据计算相似性量度,铬含量之间的差别与钙含量之间的差别相比,则显得微不足道。因此在进行聚类之前,应对原始数据进行规范化处理,以消除变量变化总幅度的影响。本文选用 Z 标准化,即对数据进行均数为 0,标准差为 1 的标准化。具体作法如下:

实验测得的数据构成原始数据阵 X (N × P), N 为样本数,每个样本有 P 个指标。为了使 P 个指标等权,对原始数据阵的每列进行标准化处理。对数据进行均数为 0,标准差为 1 的标准化。计算公式为:

$$Z_{ij} = \frac{X_{ij} - \bar{X}_{.j}}{S_j} \quad (i = 1, 2, \dots, N; j = 1, 2, \dots, p)$$

$$\text{式中: } \bar{X}_{.j} = \frac{1}{N} \sum_{i=1}^N X_{ij} \quad S_j = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_{ij} - \bar{X}_{.j})^2}$$

在表 2 矩阵中,各指标的代表值分别为: X1-Ca, X2-Zn, X3-Mg, X4-Cr, X5-Fe, X6-Mn, X7-Cu; 样本中 1—8 号为藏族, 9—17 号为畲族, 18—25 号为朝鲜族。

#### 2.2.2 主成分分析

由于变量比较多,所以我们先用主成分分析的方法,选用几个主成分进行聚类分析。计算出主成分的相关矩阵,并计算出相关矩阵特征值如表 3 和表 4 所示。

从表 4 中给出的特征值和各成分所占的方差比例可以看出:前 3 个特征值解释的方差累计为 75%。可以说前 3 个变量已经概括了大部分信息。后 4 个变量对方差的贡献率均小于 10%。因此,我们取前 3 个成分作为该问题的主成分进行聚类分析。

表 2 标准化数据矩阵  
Matrix of standardized data

NO	X1	X2	X3	X4	X5	X6	X7
1	0.68974	0.11528	0.55782	- 0.52075	2.52266	- 0.35821	0.95835
2	0.40394	0.90321	0.03410	- 0.63419	- 0.30322	- 0.41659	0.90281
3	0.92627	- 1.30905	0.30644	- 0.56612	- 0.14367	0.14774	2.34677
4	1.58329	0.78199	1.37481	- 0.63873	- 0.68369	- 0.60146	- 0.87993
5	0.67003	1.53961	0.41118	- 0.65688	- 0.54868	- 0.37767	- 0.14129
6	1.43875	1.75175	2.46414	- 0.45722	0.13248	- 0.88168	1.73587
7	0.19369	1.84266	1.60525	- 0.52982	2.06242	- 0.56254	1.06942
8	2.28302	0.29711	0.93489	- 0.73402	0.38714	- 0.90308	0.73620
9	0.38423	- 0.33929	- 0.10416	0.91771	- 0.81562	- 1.17260	- 1.14651
10	0.59447	0.35772	- 0.40163	1.02207	- 0.84017	0.97477	- 1.01322
11	- 0.41077	- 1.56361	- 0.96514	0.63637	- 0.82268	0.19639	- 0.99100
12	0.53534	0.47894	- 0.33669	1.46223	- 0.93805	- 0.36794	- 0.74109
13	- 0.35821	- 0.30899	0.83015	0.84964	- 0.71437	0.37152	- 1.47418
14	- 0.64073	- 0.30899	- 0.32202	0.78157	- 0.69903	- 1.37984	- 0.58003
15	- 0.26294	- 0.97570	- 1.15368	0.74527	- 0.92669	- 1.32244	- 0.67444
16	- 0.10526	1.17595	- 0.83107	1.85701	- 0.65300	- 0.22200	- 0.62446
17	- 0.05926	- 0.36960	- 0.80384	1.87970	- 0.80335	- 0.50416	0.01977
18	- 0.94953	- 0.36960	0.26454	0.66813	2.12379	1.41261	1.23603
19	- 1.41930	- 1.47573	- 1.43649	0.00563	- 0.21424	1.92829	1.06942
20	- 1.02837	- 0.09685	- 0.62787	- 1.85483	0.38714	1.12072	- 0.14129
21	0.34481	0.05467	1.41671	- 0.37554	- 0.38299	0.73153	- 0.21904
22	- 1.38316	- 1.70908	- 0.78498	- 0.89284	0.77681	- 0.14416	- 0.90214
23	- 1.47186	- 0.76356	- 1.18510	- 0.94729	0.11100	- 0.35821	- 0.44674
24	- 1.50800	- 0.33929	- 0.60483	- 0.80662	- 0.01787	0.04071	0.32522
25	- 0.45019	0.63046	- 0.64253	- 1.21048	1.00387	2.64830	- 0.42453

表 3 主成分的相关矩阵

Correlation Matrix of principal component

	X1	X2	X3	X4	X5	X6	X7
X1	1.0000	0.5325	0.6818	0.0280	- 0.0963	- 0.3927	0.2163
X2	0.5325	1.0000	0.6121	- 0.0715	0.1398	- 0.1659	0.1426
X3	0.6818	0.6121	1.0000	- 0.2283	0.2632	- 0.2412	0.3566
X4	0.0280	- 0.0715	- 0.2283	1.0000	- 0.4544	- 0.2253	- 0.3384
X5	- 0.0963	0.1398	0.2632	- 0.4544	1.0000	0.2598	0.5092
X6	- 0.3927	- 0.1659	- 0.2412	- 0.2253	0.2598	1.0000	0.0746
X7	0.2163	0.1426	0.3566	- 0.3384	0.5092	0.0746	1.0000

表 4 相关矩阵特征值和各成分贡献率

Cumulative proportion of correlation matrix eigenvalue and components

	Eigenvalue	Difference	Proportion	Cumulative
PRIN1	2.57457	0.62608	0.367796	0.36780
PRIN2	1.94849	1.19826	0.278356	0.64615
PRIN3	0.75023	0.11540	0.107175	0.75333
PRIN4	0.63482	0.10927	0.090689	0.84402
PRIN5	0.52555	0.18054	0.075079	0.91910
PRIN6	0.34502	0.12370	0.049288	0.96838
PRIN7	0.22132	.	0.031617	1.00000

### 2.2.3 谱系聚类分析

利用 SAS 统计软件，进行聚类分析。用平均距离法进行谱系聚类分析，得到聚类谱系图 1。

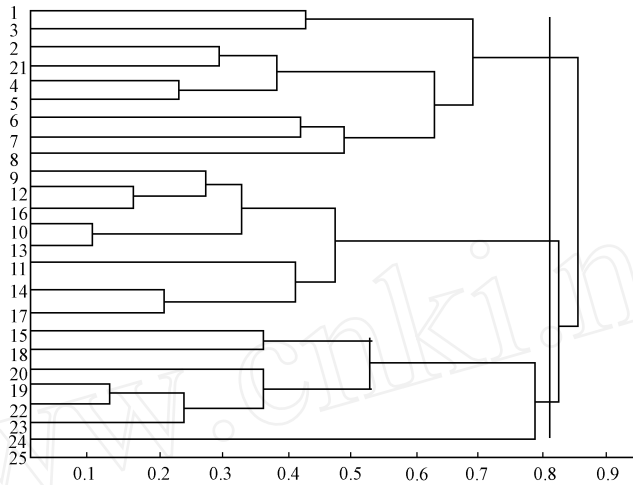


图 1 藏族、畚族和朝鲜族人体头发中微量元素聚类谱系图

Cluster analysis on contents of trace elements in human hair from Zang , She and Chaoxian nationalities

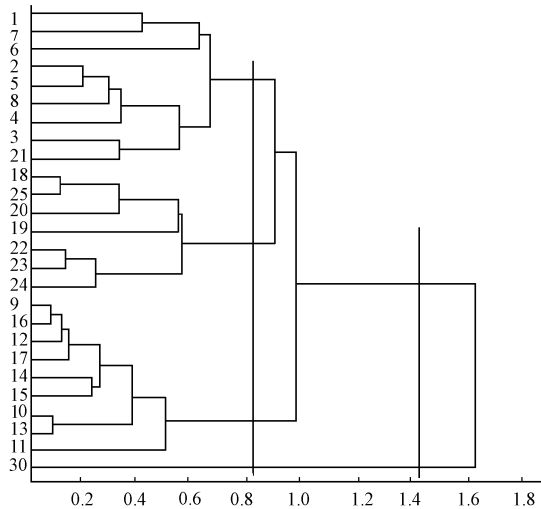


图 2 藏族、畚族、朝鲜族和维吾尔族人体头发中微量元素谱系聚类图

Cluster analysis on contents of trace elements in human hair from Zang , She , Chaoxian and Weiwuer nationalities

由图 1 可以看出：切断第二高的连线，我们看到样品聚成 (1, 3, 2, 21, 4, 5, 6, 8, 7), (9, 12, 16, 10, 13, 11, 14, 17, 15), (18, 20, 19, 22, 23, 24, 25) 3 大类。其中，样品 1—8 号来自藏族，9—17 来自畚族，18—25 来自朝鲜族。从上图我们看到以上 3 个民族基本上各自聚成一类，只有来自朝鲜族的 21 号被聚到了来自藏族的第一类，误判率为 4 %。

我们又将来自维吾尔族的一个样品加入到模式特征集中,与藏、畚、朝鲜族的样品一起聚类分析,结果如图 2 所示。

由图 2 可以看出:切断第一高线,来自维吾尔族的样品(30号)单独成为一类;切断第三高线,仍然有(1,7,6,2,5,8,4,3,21),(18,25,20,19,22,23,24)以及(9,16,12,17,14,15,10,13,11)3个大类,分别为来自藏族、朝鲜族和畚族的样品。表明,除了21号样品(朝鲜族)仍然聚到了第一类(藏族)中外,来自3个民族的样品之间的区分是明显的。

### 参考文献:

- [1] 冯宗榴,李增禧.现代微量元素研究[M].北京:中国环境科学出版社,1987,1—833.
- [2] 杨若明,王振英.用聚类分析法对土家族和瑶族人发中七种元素综合指标的比较研究[J].人类学学报,1999,18(2):126—132.
- [3] 郑连斌,郑明霞,陆舜华.亚洲21个人群体部特征的比较研究[J].人类学学报,2000,19(1):49—56.
- [4] 王碧泉,陈祖荫.模式识别理论、方法和应用[M].北京:地震出版社,1989,1—375.
- [5] 刘昆元译.聚类分析法解析分析化学数据[M].北京:化学工业出版社,1990,1—239.

## CLUSTER ANALYSIS OF ELEMENTAL CONCENTRATIONS IN THE HUMAN HAIR COVERING ZANG, SHE AND CHAOXIAN NATIONALITIES OF CHINA

YANG Ruo-ming, JIN Ji-hong

(Central University for Nationalities, Beijing 100081)

**Abstract:** In this paper, the concentration of Ca、Mg、Fe、Mn、Cu、Zn and Cr in the hair of young people from Zang, She and Chaoxian Nationalities were determined by inductively coupled plasma atomic emission spectrometry (ICP-AES). A data matrix of concentration of elements existing in the hair from three nationalities is evaluated comprehensively by using cluster analysis. The results show that there are considerable differences among these nationalities.

**Key words:** Trace element; ICP-AES; Hair; Cluster Analysis; Zang Nation; She Nation; Chaoxian Nation