

山西蜜源植物花粉的数量分类研究*

张峰 上官铁梁 谢树莲 凌元洁

(山西大学生物系 太原 030006)

提 要 应用系统聚类模型对山西 13 种蜜源植物花粉进行了数量分类研究,将它们分为 5 类,并论述了每类的显著特点。结果表明系统聚类方法应用于蜜源植物花粉的分类是有效的。此外,还讨论了系统聚类方法的适用性。

关键词 蜜源植物花粉,数量分类,山西

山西蜜源植物资源十分丰富,有 240 多种,其中重要的约 150 种,广布于全省各地的丘陵、山地和盆地,有着十分广阔的应用前景。蜜源植物花粉含有多种氨基酸、糖类、磷脂等营养物质,被誉为完全营养品,不仅在食品、日化工业方面,而且在医药工业领域都有广泛的用途。我们在进行山西蜜源植物花粉营养成分分析^[1]的基础上,采用系统聚类的方法对 13 种蜜源植物花粉进行了数量分类研究,以便为合理开发利用这些植物的花粉提供理论依据。

1 分类运算单位和性状的获得

本文选择了山西 13 种蜜源植物,它们是:(1) 沙棘(*Hippophae rhamnoides*)、(2) 向日葵(*Helianthus annuus*)、(3) 油菜(*Brassica campestris*)、(4) 玉米(*Zea mays*)、(5) 苹果(*Malus pumila*)、(6) 椴树(*Tilia mongolica*)、(7) 胡枝子(*Lespedeza bicolor*)、(8) 山楂(*Crataegus pinnatifida*)、(9) 荆条(*Vitex negundo* var. *heterophylla*)、(10) 香薷(*Elsholtzia ciliata*)、(11) 蒲公英(*Taraxacum mongolicum*)、(12) 芝麻(*Sesamum indicum*)、(13) 荞麦(*Fagopyrum esculentum*)等,用它们的花粉作为分类运算单位(OTU operational taxonomic unit)。用花粉的营养成分作为分类性状(unit character),这些成分有天冬氨酸、苏氨酸、丝氨酸、谷氨酸、甘氨酸、丙氨酸、缬氨酸、半胱氨酸、蛋氨酸、异亮氨酸、亮氨酸、酪氨酸、苯丙氨酸、赖氨酸、组氨酸、精氨酸、脯氨酸、磷脂、还原糖、蔗糖、V_A、V_E、V_C等。这样便可得到一个 23×13 的原始数据矩阵(表 1)。

2 聚类方法

2.1 原始数据的标准化

在表 1 中各种营养成分的量纲不尽相同,而且有的数值差异较大,为此要实行标准化。此处用离差标准化的方法^[2,3],对每一性状实行标准化。公式如下:

$$X_{ij} = \frac{X'_{ij} - \bar{X}_j'}{S_j} \quad (i = 1, 2, \dots, 13; j = 1, 2, \dots, 23) \quad (1)$$

本文于 1993 年 3 月 11 日收到,同年 5 月 25 日收到修改稿

* 山西省自然科学基金资助项目,本研究曾得到姜树德同志的帮助,谨致谢意。

表1 山西蜜源植物花粉营养成分

单位:mg/100g(注明除外)

Table 1 The nutritive compositions of pollen of 13 nectariferous plants in Shanxi

种名 Name of species	Unit:mg/100g(excent marked)															
	天冬氨酸 Asp	苏氨酸 Thr	丝氨酸 Ser	谷氨酸 Glu	丙氨酸 Ala	缬氨酸 Val	亮氨酸 Cys	蛋氨酸 Met	异亮氨酸 Ile	亮氨酸 Leu	酪氨酸 Tyr	苯丙氨酸 Phe	赖氨酸 Lys	组氨酸 His	精氨酸 Arg	脯氨酸 Pro
	甘氨酸 Gly	还原糖 (g/100g)	蔗糖 (g/100g)	脂肪 (g/100g)	V _A (IV/100g)	V _E	V _C									
沙棘 <i>Hippophae rhamnoides</i> L.	2.455	0.716	0.850	2.893	1.349	1.267	0.016	0.537	1.116	1.665	0.796	1.087	1.583	0.546	1.158	1.312
向日葵 <i>Helianthus annuus</i> L.	3.183	0.976	0.928	2.344	1.446	1.399	0.016	0.392	1.245	1.832	0.822	1.151	1.680	0.674	1.328	1.312
油菜 <i>Brassica campestris</i> L.	2.184	0.966	0.911	2.648	1.737	1.289	0.016	0.560	1.202	1.696	0.882	1.023	1.776	0.546	1.130	1.934
玉米 <i>Zea mays</i> L.	1.550	0.755	0.762	1.716	1.118	0.945	0.016	0.369	0.759	1.135	0.583	0.671	1.080	0.369	0.791	2.831
苹果 <i>Malus pumila</i> Mill.	2.466	1.098	1.398	3.062	1.588	1.209	0.016	0.515	1.081	1.785	0.611	1.135	1.875	0.573	1.158	2.373
椴树 <i>Tilia mongolica</i> Maxim	1.175	0.664	0.744	1.847	0.887	0.784	0.015	0.078	0.715	1.105	0.477	0.703	0.984	0.385	0.847	2.141
胡枝子 <i>Lespedeza bicolor</i> Turcz	2.750	1.097	1.077	3.105	1.543	1.524	0.016	0.492	1.288	1.983	0.902	1.215	1.476	0.642	1.328	4.627
山楂 <i>Crataegus pinnatifida</i> Bunge	1.304	0.513	0.543	1.569	0.851	0.718	0.016	0.325	0.477	0.984	0.477	0.608	1.123	0.706	0.621	2.003
荆条 <i>Vitex negundo</i> var. <i>heterophylla</i> Rehd.	1.337	0.510	0.431	1.479	0.781	0.792	0.049	0.050	0.714	1.021	0.398	0.519	1.009	0.652	0.505	0.937
香薷 <i>Elaeagnus cilata</i> Hyland	2.067	0.795	0.771	2.419	1.179	1.135	0.016	0.425	1.002	1.499	0.689	0.895	0.899	0.706	0.847	0.829
蒲公英 <i>Taraxacum mongolicum</i> Hand. —Mazz.	1.503	0.631	0.689	1.634	1.134	0.879	0.021	0.075	0.821	1.251	0.495	0.682	1.697	0.749	0.640	1.934
芝麻 <i>Sesamum indicum</i> DC	2.173	0.916	1.007	2.468	1.337	1.069	0.016	0.470	0.916	1.438	0.716	0.863	1.401	0.514	1.215	2.831
荞麦 <i>Fagopyrum esculentum</i> Moench	1.703	0.644	0.587	3.351	0.996	0.930	0.016	0.090	0.802	1.181	0.477	0.703	0.481	0.369	0.650	0.691

其中 $\bar{X}_j = \sum_{i=1}^{13} \frac{X_{ij}}{13}$, $S_j = \sqrt{\sum_{i=1}^{13} (X_{ij} - \bar{X}_j)^2}$.

2.2 求 OTU 间的距离系数(d_{jk})

$$d_{jk} = \sqrt{\sum_{i=1}^{23} (X_{ji} - X_{ki})^2} \quad (j, k = 1, 2, \dots, 13) \quad (2)$$

这样便可得到 OTU 间的 13×13 距离系数矩阵。

2.3 聚类方法

以距离系数矩阵为基础,采用系统聚类模型进行聚类分析。其模型如下:

$$D_{ij} = \alpha_1 D_{iu} + \alpha_2 D_{ju} + \beta D_{ik} + \gamma |D_{iu} - D_{ju}| \quad (3)$$

在(3)中,随着参数 $\alpha_1, \alpha_2, \beta, \gamma$ 的不同,它们分别是最短距离法、最长距离法、重心法、中线法、类平均法、离差平方和法和可变法^[4,5]。

3 分类结果与讨论

3.1 分类结果

依系统聚类模型的方法,对 13 种蜜源植物花粉进行聚类。按照聚类结果可作出相应的树状图来,此处仅作出类平均法的树状图(图 1)。聚类树状图并没有达到对 13 种蜜源植物花粉的分类目的,为此依结合线法^[6],取不同的阈值 λ 对树状图进行截取,便可实现对 13 种蜜源植物花粉的分类,其中 $\lambda = 70.07$ 是最好的(即第 8 次聚合与第 9 次聚合的结合系数之差,是所有相邻两次聚合中最大的)。此时,13 种蜜源植物花粉被分为 5 类,结果如下:

I (1, 2, 3, 12) 包括沙棘、向日葵、油菜、芝麻,其相似特征为:丝氨酸(0.85—1.007),甘氨酸(1.222—1.446),半胱氨酸(0.16),蛋氨酸(0.396—0.56),酪氨酸(0.766—0.882),组氨酸(0.514—0.674),精氨酸(1.13—1.328)。天冬氨酸(2.173—3.183)和 V_C (37—83.5)含量最高,赖氨酸含量较高(1.401—1.776), V_E 差异较大。

II (4, 6, 13) 包括玉米、椴树、荞麦,其相似特征为:甘氨酸(0.741—0.855),缬氨酸(0.784—0.945),半胱氨酸(0.15—0.16),异亮氨酸(0.715—0.802)。酪氨酸(0.477—0.583),苯丙氨酸(0.671—0.703),组氨酸(0.369—0.385),精氨酸(0.65—0.847),还原糖(30.94—32.08)。天冬氨酸较低(1.175—1.703), V_C, V_E 中等且差异不大。

III (10) 仅含香薷一种。与最相似的 I 组距离系数达 54.21,表明它与 I 组差异性较大。事实上,除组氨酸、蔗糖皆比 I 组中的所有种高,亮氨酸、异亮氨酸、苯丙氨酸、缬氨酸比芝麻高外,其它营养成分皆低于 I 组中的所有种。因此,香薷被独立分为一组就不奇怪

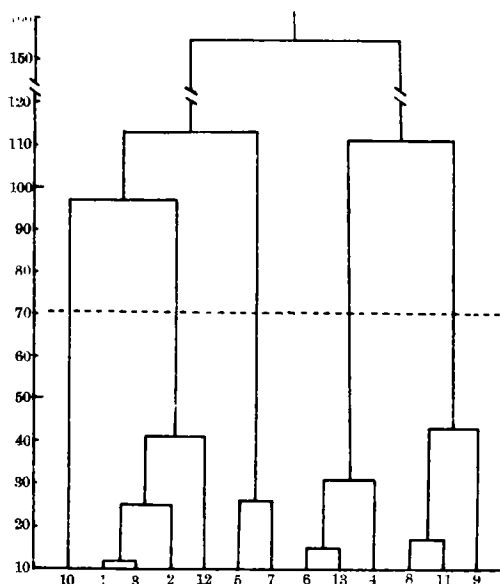


图 1 类平均法的树状图

Fig. 1 The dendrogram of group-average method

了。

IV(5,7) 包含苹果、胡枝子, 它们的苏氨酸(1.090—1.097), 丙氨酸(1.543—1.588), 谷氨酸(3.062—3.105), 蛋氨酸(0.492—0.515), 异亮氨酸(1.081—1.288), 亮氨酸(1.785—1.983), 苯丙氨酸(1.135—1.215), 组氨酸(0.573—0.642), 精氨酸(1.158—1.328), 磷脂(2.18—2.36), 蔗糖(2.98—3.19)极为相近, 半胱氨酸相同(0.016), V_E 、赖氨酸较高, V_A 差异极大(39230—92310)。

V(8,9,11) 包括山楂、荆条、蒲公英, 其苏氨酸(0.510—0.631), 缬氨酸(0.718—0.879), 半胱氨酸(0.016—0.049), 酪氨酸(0.398—0.495), 苯丙氨酸(0.519—0.682), 组氨酸(0.706—0.749), 精氨酸(0.505—0.640), 磷脂(0.67—0.76)较为相似, V_A 差异较大。

3.2 关于聚类方法的选择

数量分类方法自诞生以来, 在生态学、系统分类学等方面已有了广泛的应用, 显示了强大的生命力, 但是有关花粉方面的应用研究尚未见报道。因此采用哪种方法对花粉进行聚类效果更好, 目前尚无客观标准可供参考。本文应用的7种聚类方法中, 效果较好且应用广泛的首推类平均法, 因为其聚合过程是单调的, 亦即在聚类过程中, 始终保持上一级聚合的相似系数小于(或等于)下一级^[3]; 同时又是空间保持的, 也就是说随着聚合过程的进行, 不会使某一组与另一组更加接近, 从而使它们有进一步与别的组合并的趋势; 也不会使某一组与另一组有更加疏远的趋向, 因而起着阻碍进一步与别的组合并的作用。这样就使得聚类结果符合客观情况, 这是其余6种聚类方法所不具备的优点。更进一步地分析分类结果, 除类平均法外, 其余6种方法所得结果都有一些不合理类群, 这表明类平均法用于蜜源植物花粉的分类是比较理想的。

3.3 关于蜜源植物的合理利用问题

从分类结果和表1可以看到, 第I组蜜源植物花粉富含赖氨酸和 V_C , 而这些植物分布范围广, 花粉产量极大, 具有广阔的开发应用前景。它们的花粉可作为食品添加剂, 弥补食品中赖氨酸和 V_C 的不足, 起到强化赖氨酸和 V_C 的作用, 有利于提高人民的健康水平, 特别有利于促进婴幼儿的健康成长。第IV组蜜源植物花粉最明显的特征是 V_E 含量比其它组高, 因此它们可作为食品饮料和化妆品的添加剂, 起到抗衰防衰作用, 应引起足够重视, 有步骤有计划地加以开发利用。

一般来说, 花粉和蜂蜜的营养成分呈正相关性。通过花粉营养成分的聚类分析, 可以推测其相应的蜂蜜也有类似的营养成分, 因此在开发利用花粉资源的同时, 应积极开发利用营养价值较高且便于利用的蜂蜜, 如向日葵、油菜、苹果等蜂蜜, 取得较好的经济效益。

参 考 文 献

- 1 王开发、凌元洁等. 山西省蜜源植物花粉营养成分的分析. 山西大学学报, 1992, 15(3): 299—308
- 2 阳含熙等. 植物生态学的数量分类方法. 北京: 科学出版社, 1981, 29—30, 120
- 3 张峰等. 模糊图论在山西植被区划中的应用. 植物生态学与地植物学学报, 1991, 15(1): 94—100
- 4 朗奎健等. IBM PC 系列程序集——数理统计调查规划 经营管理. 北京: 中国林业出版社, 1989, 131—134, 439—440

- 5 钟扬等. 数量分类方法与程序. 武汉: 武汉大学出版社, 1990. 27
- 6 徐克学等. 我国人参属数量分类初探. 植物分类学报, 1983, 21(1): 34—43

ON NUMERICAL CLASSIFICATION OF THE POLLEN OF HONEY SOURCE PLANTS FROM SHANXI

Zhang Feng Shangguan Tieliang Xie Shulian Ling Yuanjie

(Department of Biology, Shanxi University Taiyuan 030006)

Abstract In present paper, system clustering model is used to study numerical classification of the pollen of 13 species of honey source plants from Shanxi Province, the 13 species are divided into 5 groups and the remarkable features of the each group are discribed. We believe that it is effective to apply system clustering method to the taxonomy of the pollen. Besides, the suitability application system clustering method to the classification of the pollen is discussed.

Key words The pollen of honey source plant, Numerical classification, Shanxi

《武汉植物学研究》编辑委员会 关于调整本刊收费标准的决定

根据中国科学院(92)出字 33 号文件精神,并结合本刊实际情况,编委会决定对本刊收费标准进行调整。具体内容如下:

1. 从 1994 年 3 月 1 日起,每篇来稿一律收取稿件审理费 50 元。
2. 从 1994 年(第 12 卷)第 3 期开始,录用稿件,每版(含图表)收发表费(版面费)50 元,论文、综述等超过 6 版者,简报超过 2 版者,超版部分每版收费 100 元。凡专页铜版纸图片,不论超版与否,每版皆收 100 元。

以上费用,请信汇或邮寄本刊编辑部。地址:武汉市 74006 信箱,邮编:430074。信汇帐号:中国科学院武汉植物研究所 2612—88—037053236,开户银行:工商银行武汉关山办事处²⁶⁵³¹₂₀₂₀₂。汇单附言栏内注明稿件审理费或版面费。本刊可以开具报销单据。

自本决定公布施行之日起,本刊原有收费标准即告取消。

《武汉植物学研究》编辑委员会

1994 年 1 月 20 日