数量分类方法在蜂蜜孢粉学中的探讨: 以山西中部地区为例

宋晓彦^①, 姚轶锋^{②*}

① 山西农业大学农学院, 太谷 030801;

② 中国科学院植物研究所系统与进化植物学国家重点实验室, 北京 100093

* 联系人, E-mail: yaoyf@ibcas.ac.cn

2014-08-12 收稿, 2014-09-17 接受, 2014-11-26 网络版发表

山西省青年科技研究基金(2010021032-2)和山西省农业科技攻关项目(20140311005-5)资助

摘要 蜂蜜孢粉学是研究蜂蜜中花粉的科学, 是孢粉学的一门重要分支学科. 本文选取山西中部地区 19 个蜂蜜样品作为分类运算单位, 61 种花粉类型作为分类性状, 采用数量分类方法对蜂蜜样品进行了系统聚类分析(Q 型聚类)和主成分分析. 系统聚类将 19 个样品分成 8 类, 而主成分分析将它们分为 5 类, 并论述了每类的显著特点. 结果表明, (i) 数量分类方法用在蜂蜜孢粉学中蜂蜜样品的分类是可行的, 而且分类结果直观易懂; (ii) 在系统聚类的实际分类过程中, 需要特别注意分类阈值 λ 的取值问题, 通常 λ 取值不同, 分类结果也不同, 因此 λ 取值要视具体情况而定, 既把相似性大的样品归为同类, 又把相似性小的样品尽量分开; (iii) 由于数量分类方法和蜂蜜孢粉定量分析方法都考虑了优势花粉类群的丰度, 因此分类结果有相似之处; 但由于前者同时还兼顾了非优势花粉类群的丰度, 所以分类结果又有别于蜂蜜花粉定量研究结果.

关键词

蜂蜜孢粉学
数量分类
聚类分析
主成分分析

数量分类学(numerical taxonomy)也称数值分类学, 是一门使用数学方法、统计学理论和电子计算机技术研究解决生物学中分类问题的科学^[1]. 数量分类方法通常是通过数学方法将所有的分类性状加以等权处理, 然后根据性状间的相似性来进行归类. 其分类效率和分类的客观性比传统的分类方法大大提高^[1].

近年来, 数量分类方法在植物分类和品种的亲缘关系分析等方面发挥着较好的作用. 例如, 彭焱松等人^[2]通过数量分类方法, 将中国栎属分为麻栎组、槲栎组、高山栎组、巴东栎组和檀子栎组等 5 个特征明显的表征群. 张广进等人^[3]、赵冰等人^[4]、顾翠花等人^[5]、王晓慧等人^[6]、张毓等人^[7]、张西丽等人^[8]、骆菁菁等人^[9]将数量分类方法分别成功用于蔷薇、蜡

梅、紫薇、大丽花、杓兰、百合、海棠等花卉的品种分类.

此外, 数量分类方法在蜜源植物花粉的分类方面也得到了初步尝试. 例如, 张峰等人^[10]采用系统聚类模型对山西玉米、向日葵等 13 种蜜源植物花粉中的天冬氨酸、苏氨酸等 23 个成分指标进行了数量分类研究, 将 13 种蜜源花粉分为 5 类. 曾志将等人^[11]利用聚类分析理论对我国主要的油菜、荞麦等 16 种蜜源花粉中的人体必需的 8 种游离氨基酸、磷脂等 20 种成分指标进行了数量分类研究, 将 16 种蜜源花粉分为四大类. 曾志将和王开发^[12]对山里红、紫云英等 26 种蜜源植物花粉中的钙、钠、磷等 23 种矿物质进行了数量分类研究, 将 26 种蜜源花粉分为 12 类. 这些研究表明, 数量分类方法用于蜜源植物花粉的分类是行之

引用格式: 宋晓彦, 姚轶锋. 数量分类方法在蜂蜜孢粉学中的探讨: 以山西中部地区为例. 科学通报, 2015, 60: 58–67

Song X Y, Yao Y F. Numerical taxonomy in melissopalynology: A case study in the central region of Shanxi, China (in Chinese). Chin Sci Bull, 2015, 60: 58–67, doi: 10.1360/N972014-00073

有效的。

蜂蜜孢粉学(Melissopalynology)是研究蜂蜜中花粉的科学,是孢粉学的一门重要分支学科^[13]。蜂蜜是一种纯天然、无污染的理想保健食品,具有抗菌消炎、润肺止咳、美容和增强细胞免疫等功效。花粉是天然蜂蜜的重要标志,通过蜂蜜中花粉的鉴别和分类,可以确定蜂蜜的种类和品质,以确保消费者的食用安全。迄今为止,数量分类方法在蜂蜜孢粉学中有过初步的探讨^[14-17]。例如,Devillers等人^[14]的研究表明,质量控制方法结合多变量统计分析有助于对不同地理来源的蜂蜜进行分类。Aronne和De Micco^[15]通过实验设计,结合对蜂蜜花粉数据进行聚类分析和主成分分析,探讨了蜂蜜在不同采集时间和地域条件下的花粉组成变化,从而有助于提高蜂蜜孢粉学在确定蜂蜜的植物和地理来源方面的应用。Herrero等人^[16]从植物和地理来源的角度出发,采用聚类分析和统计分析对西班牙西北部5个不同地貌单元的蜂蜜样品进行分类和特征描述。Upadhyay和Bera^[17]对印度东部34个蜂蜜的花粉数据进行主成分分析,证实了依据植物来源对蜂蜜进行分类的可靠性。本文以山西中部地区为例,在蜂蜜孢粉学研究基础上,尝试探讨和比较系统聚类方法、主成分分析方法以及蜂

蜜花粉定量分析方法所得结果的异同,并且对产生差异的原因进行了初步剖析。

1 材料与方法

(i) 蜂蜜样品采集方法。2010~2011年4~9月间,从山西中部柳林县、中阳县、寿阳县、祁县、文水县、汾阳县、临县、太谷县和清徐县等9个县采集到19个蜂蜜样品。现场采用摇蜜机通过离心力的作用将蜂巢中的蜂蜜分离出来,分别收集保存在150 mL塑料离心管中盖好,对蜂蜜样品进行编号(H1~H19),记录采集时间和地点。样品H1采自柳林县,H2采自中阳县,H3采自寿阳县,H4,H5和H14采自祁县,H6~H9采自文水县,H10~H12采自汾阳县,H13采自临县,H15和H16采自太谷县,H17~H19采自清徐县(图1)。

(ii) 分类单位和分类性状的选择。采用蜂蜜孢粉分析方法对19个蜂蜜样品进行了定性分析和定量研究,共鉴定61种花粉类型^[18],它们是:伞形科(Apiaceae)、五加科(Araliaceae)、菊科(Asteraceae)、十字花科(Brassicaceae)、大戟科(Euphorbiaceae)、豆科(Fabaceae)、龙胆科(Gentianaceae)、唇形科(Lamiaceae)、锦葵科(Malvaceae)、楝科(Meliaceae)、

图1 蜂蜜样品采集位置图(括号内为样品编号)

Figure 1 Location map of honey samples (Sample No. in bracket)

蔷薇科(Rosaceae)、玄参科(Scrophulariaceae)、茄科(Solanaceae)、马鞭草科(Verbenaceae)、藜科(Chenopodiaceae)、莎草科(Cyperaceae)、禾本科(Poaceae)、毛茛科(Ranunculaceae)、蒿属(*Artemisia* sp.)、黄栌属(*Cotinus* sp.)、盐肤木属(*Rhus* sp.)、猫尾树属(*Markhamia* sp.)、接骨木属(*Sambucus* sp.)、皂荚属(*Gleditsia* sp.)、甘草属(*Glycyrrhiza* sp.)、苜蓿属(*Medicago* sp.)、野豌豆属(*Vicia* sp.)、石栎属(*Lithocarpus* sp.)、葱属(*Allium* sp.)、百合属(*Lilium* sp.)、木兰属(*Magnolia* sp.)、葎草属(*Humulus* sp.)、杨梅属(*Myrica* sp.)、丁香属(*Syringa* sp.)、蓼属(*Polygonum* sp.)、大黄属(*Rheum* sp.)、苹果属(*Malus* sp.)、李属(*Prunus* sp.)、梨属(*Pyrus* sp.)、蔷薇属(*Rosa* sp.)、白鹃梅属(*Exochodra* sp.)、柳属(*Salix* sp.)、泡桐属(*Paulownia* sp.)、椴树属(*Tilia* sp.)、榆属(*Ulmus* sp.)、松属(*Pinus* sp.)、蒲公英(*Taraxacum mongolicum*)、梓树(*Catalpa ovata*)、金银忍冬(*Lonicera maackii*)、田旋花(*Convolvulus arvensis*)、大豆(*Glycine max*)、草木樨(*Melilotus suaveolens*)、刺槐(*Robinia pseudoacacia*)、国槐(*Sophora japonica*)、白刺花(*Sophora viciifolia*)、核桃(*Juglans regia*)、荞麦(*Fagopyrum esculentum*)、枣(*Ziziphus jujuba*)、栎树(*Koelreuteria paniculata*)、臭椿(*Ailanthus altissima*)、荆条(*Vitex negundo* var. *heterophylla*)。

本文选取19个蜂蜜样品作为分类运算单位(operational taxonomic unit, OTU), 上述61种花粉类型作为分类性状(unit character), 这样便可以得到一个61×19的原始数据矩阵(表1)。

(iii) 聚类分析方法. 鉴于表1中各分类性状的量纲相同, 所以不需要对分类性状进行标准化处理, 可以直接根据以下公式计算不同分类单位间的欧式

$$\text{距离: } d_{ij} = \sqrt{\sum_{k=1}^{61} (X_{ik} - X_{jk})^2} \quad (i=1, 2, 3, \dots, 19;$$

$j, k=1, 2, 3, \dots, 61)$, 这样便得到分类单位间的19×19欧式距离矩阵. 在此基础上, 采用类平均法对其进行Q型聚类分析. 整个聚类分析数值运算使用SPSS 17软件完成。

(iv) 主成分分析. 与聚类分析类似, 主成分分析时未对表1中的分类性状进行标准化处理, 但对各分类单元和分类性状进行了中心化处理. 在进行主成分分析前, 首先对分类性状进行降维对应分析(DCA分析), 以确定是选择单峰模型还是线性模型比较合适. 由于4个排序轴中梯度最大值为3.378(介于3~4之间), 所以选择单峰模型和线性模型均合适. 本文采用线性模型的主成分分析, 由CANOCO 4.5软件

完成。

2 研究结果

2.1 聚类分析结果

依据聚类分析结果, 绘制出19个蜂蜜样品的聚类树状图(图2). 树状图只显示了各个样品间的亲疏关系, 并没有达到分类的目的, 因此需要采取结合线法^[19], 通过取不同的分类阈值 λ 对树状图进行截取, 以便实现对样品的分类. 当 $\lambda=62.975$ 时(虚线Lq1所示为等级结合线), 19个样品被分为8类:

第Ⅰ类: 包括H1, H6, H8, H11, H13, H16等6个样品, 其相似特征为: 枣花粉含量较高, 尤其H1和H13相似程度更接近(枣花粉含量分别高达97.00%和92.44%), H6(50.44%)和H8(48.89%)枣花粉含量相当, H11(31.20%), H16(64.75%).

第Ⅱ类: 包括H4, H10, H12等3个样品, 其相似特征为: 十字花科花粉含量均在10%左右(H4: 9.25%, H10: 10.19%, H12: 10.25%). 此外, H4(8.36%)和H10(5.26%)样品中蒲公英花粉含量接近。

第Ⅲ类: 包括1个样品(H14), 其主要特征是以菊科(48.85%)和蔷薇科(25.48%)花粉为主。

第Ⅳ类: 包括1个样品(H7), 其主要特征是以豆科(60.63%)和柳属(20.00%)花粉为主。

第Ⅴ类: 包括1个样品(H5), 其主要特征是以臭椿(61.02%)和玄参科(17.13%)花粉为主。

第Ⅵ类: 包括H2, H18, H19等3个样品, 其相似特征为: 刺槐和国槐花粉含量接近. 刺槐花粉含量均超过60%(H2: 72.40%, H18: 61.01%, H19: 69.08%), 国槐花粉含量H2(12.00%)、H18(28.21%)、H19(13.74%).

第Ⅶ类: 包括H9, H17等2个样品, 其相似特征为: 以国槐花粉含量为主(H9: 40.88%, H17: 77.40%), 刺槐花粉含量接近(H9: 14.75%, H17: 15.20%).

第Ⅷ类: 包括H3, H15等2个样品, 其相似特征为: 荆条花粉含量接近, 在2个样品中分别为63.60%和62.37%。

2.2 主成分分析结果

主成分分析结果表明, 前2个主成分累积贡献率达到56.4%. 主成分分析将19个蜂蜜样品分为5类(图3):

第Ⅰ类(Cluster 1): 包括H1和H13等2个样品, 其相似特征是枣花粉含量非常高, 超过90%以上(H1:

表 1 19 个蜂蜜样品中的花粉类型和花粉率^{a)}
Table 1 Pollen types recovered from the 19 honey samples and their frequency^{a)}

分类 单元	COT	RHU	API	ARA	ART	AST	TAR	CAT	MAR	BRA	LON	SAM	CON	EUP	FAB	GLE	GLY	GLYC	MED	MEL	ROB
1	1	0	0	0	0	0	0.4	0	0	0	0	0	0	0	0	0	0	0	0.2	0	0.4
2	0	0	0.2	0	0.2	0	0	0	0	0	0	0	0.2	0	0	0	0	0	0.4	0	72.4
3	0	0	0	0.2	0	1.6	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.4	0
4	0	0	0	0	0	0	8.36	0	0	9.25	0.15	0	6.12	0	22.54	0	0	0	0	0	0
5	0	2.76	0	0	0.2	0	0.39	0	0	0.59	0.39	0	0	0	11.02	0	0	0	0	0	0
6	0	0	0	4.55	0	0	3.23	0	0	8.21	0	0	0	0	9.38	0	0	0	0	1.61	0
7	0	0	0.47	0	0	0	6.61	0	0	3.78	0.16	0	0	0	60.63	0	0	0	0	0	0
8	5.66	0	0	0	0.2	0	0	0	0	0	0	0	0	0	21.82	0	0	7.88	0	6.87	0
9	1.07	0	0	0	0.27	0.13	0	0	0.54	0.27	0.94	0	0	0	5.9	0	0	0	5.9	0.4	14.75
10	0	0	0	0	1.02	0	5.26	23.77	0	10.19	0	3.74	0	0	3.23	0	0	0.34	3.06	0.17	18.85
11	0	0	0	0	0	16.22	0	0.62	0	9.2	0.47	0	0	0	7.02	0	0	0	0.31	4.37	3.9
12	0	0	0	0	0.77	2.13	0	0	0	10.25	0	0	0	0	0	0	7.16	0	0.19	11.8	1.16
13	0	0	0	0	0.39	1.55	0	0	0	0.39	0	0	0	0	0	0	0	0	0	0	2.91
14	0	0	0	0	0.38	48.85	0	0	0	0.77	0	0	0.19	0	8.05	0	0.19	0	0	0	1.53
15	0	0	0	0	0	0	0	0	0.35	0	0	0	0	0.7	0	0	0	0	2.26	0.35	3.48
16	0	0	0.17	0	0	8.38	9.6	0	0.35	0.35	0.52	0	1.22	0	2.97	0.52	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0.2	0	2.6	0	0	0	0	0	0	0	0	0	15.2
18	0	0	0	0	0.16	0.48	0	0	0	0	0.63	0	0	0	0	0	0	0	0.16	0	61.01
19	0	0	0	0	1.34	0	0.95	0	0	2.48	1.53	0	0.38	0	0	0	0	0	0	0	69.08

a) COT, *Cotinus* sp.; RHU, *Rhus* sp.; API, Apiaceae; ARA, Araliaceae; ART, Artemisia sp.; AST, Asteraceae; TAR, *Taraxacum mongolicum*; CAT, *Catalpa ovata*; MAR, *Markhamia* sp.; BRA, Brassicaceae; LON, *Lonicera maackii*; SAM, *Sambucus* sp.; CON, *Convolvulus arvensis*; EUP, Euphorbiaceae; FAB, Fabaceae; GLE, *Gleditsia* sp.; GLY, *Glycine max*; GLYC, *Glycyrrhiza* sp.; MED, *Medicago* sp.; MEL, *Melilotus suareolens*; ROB, *Robinia pseudoacacia*

续表 1

分类 单元	SOPJ	SOPV	VIC	LIT	GEN	JUG	LAM	ALL	LIL	MAG	MAL	MEL	HUM	MYR	SYR	FAG	POL	RHE	ZIZ	EXO	MAL
1	0.6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	97	0	0.4
2	12	0.2	0	0	0	0	0	0	0	0	0	0	0.4	0	0.2	0	0.4	0	5	0	6.8
3	6.8	0.2	0.2	0	0	0.2	0.2	0	0	0	0	0	1.8	0	7.6	0.2	0.2	0	6.2	0	0.6
4	0	0	0	0	0	0	0.6	0.9	0	0	0	0.15	2.99	0	1.04	0	0	0	0	0	4.33
5	0	0.2	0	0	0.39	0	0.59	0	0	0	0	0	0.59	0	0	0	0	0	4.53	0	0
6	20.67	0	0	0	0	0	0	0	0	0.73	0	0	0	0.15	0	0.15	0	0	50.44	0	0
7	0	0	0	0	0.16	0	0	5.51	0	0.31	0	0	0.16	0	0	0	0	0	0	0	0
8	7.88	0	0	0	0	0	0	0	0	0	0.4	0	0.2	0	0	0	0	0	48.89	0	0
9	40.88	0	0	0	0	0	1.47	0	0	0	0	0	0	0	0	0	0	0	21.18	0	0
10	2.04	4.41	0	0	0	0	0	0	3.06	0	0	0	0.17	0	0	0	0	0	0.85	0	0
11	9.05	0.31	0	0	0	0	0	0	0.16	0	0	0	0.31	0	0	0	0	0.16	31.2	0	0
12	0.58	0	0	4.26	0	0	0.19	0	0	0	0	0.58	3.48	0	0	0	0	0	10.25	41.78	0
13	1.94	0	0	0	0	0	0.19	0	0	0	0	0	0	0	0	0.19	0	0	92.44	0	0
14	0	4.21	0	0	0	0	0.77	3.64	0.19	0	0	0	0.19	0	2.11	0	0	0	0.96	0	0
15	18.47	0	0	0	0.87	0	0.35	0	0	0.17	0	0.17	0.17	0	0	0	0	0	1.92	0	0
16	0.17	0	0	0	2.79	0	1.05	0.17	0.35	0	0	0	0	0	0.87	0.17	0	0	64.75	0	0
17	77.4	2.8	0	0	0	0	0.2	0	0	0	0	0	0.2	0	0	0	0	0	0	0	0
18	28.21	0	0	0	0	0	0	0	0	0	0	0	0	0	0.16	0	0	0	0	0	0
19	13.74	0	0	0	0	0	0	0	0.19	0	0	0	1.91	0	0.76	0	0	0	1.91	0	0

a) SOPJ, *Sophora japonica*; SOPV, *Sophora viciifolia*; VIC, *Vicia* sp.; LIT, *Lithocarpus* sp.; GEN, Gentianaceae; JUG, *Juglans regia*; LAM, Lamiaceae; ALL, *Allium* sp.; LIL, *Lilium* sp.; MAG, *Magnolia* sp.; MAL, Malvaceae; MEL, Meliaceae; HUM, *Humulus* sp.; MYR, *Myrica* sp.; SYR, *Syringa* sp.; FAG, *Fagopyrum esculentum*; POL, *Polygonum* sp.; RHE, *Rheum* sp.; ZIZ, *Ziziphus jujuba*; EXO, *Exochorda* sp.; MAL, *Malus* sp.

续表 1

分类 单元	PRU	PYR	ROS	ROSA	SAL	KOE	PAU	SCR	AIL	SOL	TIL	ULM	VER	VIT	CHE	CYP	PIN	POA	RAN
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0.2	0	0	0	0	0	0	0	0	0	0.4	0	0	0	0	0	0	0
3	0	0	0.2	0	0	0	0	0	0	0	0	0	3.2	63.6	2.6	0	0	3.6	0
4	0	1.64	0	5.22	3.13	0	28.21	0	0	1.34	0	0	0	0	0	0	0	0.9	0
5	0	0	0	0	0	0	0	17.13	61.02	0	0	0	0	0	0.2	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.29	0	0	0.59	0
7	1.73	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0.47	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0	0
9	0	0.67	0	0	0.13	0	0	0	0	0	0	0	0	0	0	0	0.67	0	0.27
10	0	0	0	0	0	0	0	0	0	0.51	0	0	0	0	0	0	0	0	19.35
11	0	0	0	0	0	2.34	0	0	3.74	0	0.16	0	0	9.2	0.16	0.16	0	0.16	0.78
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.19	0.58	0.19	0.58	1.16
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	25.48	0.38	0	0	0	0	0	0	0	0	0.57	0	0	0.38	0.57	0.19
15	7.14	0	0	0.7	0.17	0	0	0	0	0	0	0	0	62.37	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.54	0	0.17	0.52	0
17	1.4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	7.13	0	0	0	0	0	2.06	0	0	0	0	0	0	0	0	0	0	0	0
19	5.15	0	0	0	0.38	0	0	0	0	0	0	0	0	0	0.19	0	0	0	0

a) PRU, *Prunus* sp.; PYR, *Pyrus* sp.; ROS, *Rosa* sp.; ROSA, *Rosa* sp.; SAL, *Salix* sp.; KOE, *Koelreuteria paniculata*; PAU, *Paulownia* sp.; SCR, *Scrophulariaceae*; AIL, *Ailanthus altissima*; SOL, *Solanaceae*; TIL, *Tilia* sp.; ULM, *Ulmus* sp.; VER, *Verbenaceae*; VIT, *Vitex negundo* var. *heterophylla*; CHE, *Chenopodiaceae*; CYP, *Cyperaceae*; PIN, *Pinus* sp.; POA, *Poaceae*; RAN, *Ranunculaceae*

图2 蜂蜜样品的聚类树状图

Figure 2 The dendrogram of cluster analysis of honey samples

图3 蜂蜜样品的主成分分析结果

Figure 3 Result of principal component analysis of honey samples

97.00%, H13: 92.44%).

第II类(Cluster 2): 包括H2, H18和H19等3个样品, 其相似特征是刺槐和国槐花粉含量接近(刺槐 H2: 72.40%, H18: 61.01%, H19: 69.08%; 国槐 H2: 12.00%, H18: 28.21%, H19: 13.74%). 该结果与聚类分析结果第IV类一致。

第III类(Cluster 3): 包括H6, H8, H11和H16等4个样品, 其相似特征是枣花粉含量相对较高, 达到30%~65%(H6: 50.44%, H8: 48.89%, H11: 31.20%, H16: 64.75%).

第IV类(Cluster 4): 包括H9和H17等2个样品, 其相似特征是以国槐花粉含量为主(H9: 40.88%, H17: 77.40%), 刺槐花粉含量接近(H9: 14.75%, H17: 15.20%). 该结果与聚类分析结果第VII类一致。

第V类(Cluster 5): 包括H3, H4, H5, H7, H10,

H12, H14和H15等8个样品, 其相似特征是以枣、国槐、刺槐之外的其他花粉为主, 包括荆条、十字花科、蒲公英、臭椿、玄参科、豆科、柳属、菊科、蔷薇科等。

3 讨论

3.1 分类阈值 λ 取值问题初探

实践证明数量分类方法能够对大量生物学性状进行比较全面的综合分析, 摆脱了传统分类的主观性, 能够得出比较正确的分类结果^[1]。本文根据聚类树状图, 可以直观地判断蜂蜜样品间花粉组成和百分含量的相似性。当分类阈值 $\lambda=62.975$ 时, 可以明显地把山西中部地区19个蜂蜜样品分成8类。当然, λ 取值不同, 可以得到不同的分类结果。例如, 当 λ 取74.613时, 19个样品则可以分成5类。但是, 无论如何都不能改变蜂蜜样品间的相似性。因此, 在实际分类过程中, λ 取值要视具体情况而定, 既要考虑到把相似性大的样品放在同一类, 又要考虑把相似性小的样品尽量分开^[12]。

3.2 数量分类方法与蜂蜜花粉定量分析结果的比较

花粉是蜂蜜分析中的一个关键指标, 可以通过蜂蜜中花粉的鉴别和分类来确定蜂蜜的种类和品质。Song等人^[18]对山西中部地区19个蜂蜜样品进行了蜂蜜花粉的定量研究, 并按2种方法对样品进行了分类:

根据10克蜂蜜样品中花粉的绝对数目(absolute pollen count, APC)可以将19个样品分为3组, 即第I组($APC < 20000$), 包括H2, H3, H5, H7, H8, H10, H11, H12, H14, H15, H16, H18, H19等13个样品; 第II组($20000 < APC < 100000$), 包括H4, H6, H9, H17等4个样品; 第III组($100000 < APC < 500000$), 包括H1, H13等2个样品。

根据蜂蜜样品中是否有含量大于45%的花粉类型, 可以将19个样品分为2类: (i) 单花蜜, 包括H1, H2, H3, H5, H6, H7, H8, H13, H14, H15, H16, H17, H18, H19等14个样品; (ii) 多花蜜, 包括H4, H9, H10, H11, H12等5个样品。在14个单花蜜中, H1, H6, H8, H13, H16是以枣花粉为主; H2, H18, H19是以刺槐花粉为主; H3, H15是以荆条花粉为主; H5, H7, H14, H17分别是以臭椿、豆科、菊科、国槐花粉为主。

本文主要从花粉组成和含量的相似性特征出发,

采用数量分类方法对19个样品进行分类,可以将19个样品分为8类(图2,见聚类分析结果)或5类(图3,见主成分分析结果)。与上述2种分类方法相比,本文数量分类的最大特点是以树状聚类图或载荷图的形式直观地将分类结果表现出来,而且结果易于理解,各样品间的亲疏关系一目了然。然而,通常根据不同的分类方法产生的分类结果也不尽相同,本文的数量分类结果与上述的定量研究结果之间存在一定的相似性和差异性。例如,APC方法和聚类方法、主成分分析都将H1和H13归为同一类;根据APC方法,将H4, H6, H9, H17归入同一类,而聚类方法和主成分分析则将H4, H6, H9, H17归为不同的3类(其中H9和H17归为同一类);根据样品中某种花粉的含量是否大于45%的方法,将H1, H6, H8, H13, H16归为同一类(单花蜜),将H11归为另一类(多花蜜),主成分分析把它们分为两类,而聚类方法则将它们都归入同一类。鉴于每种分类方法都有各自的特点,因此不好评判各种方法的好坏。

3.3 与国外相关报道的对比

Herrero等人^[16]采用聚类分析和统计分析对西班牙西北部5个不同地貌单元的89个蜂蜜样品进行了分类。蜂蜜孢粉分析结果表明,89个样品包括43个多花蜜和46个单花蜜,而单花蜜可分为以金雀儿属(*Cytisus*)、矢车菊属(*Centaurea*)、欧石楠属(*Erica*)、悬钩子属(*Rubus*)、驴食草属(*Onobrychis*)、木犀草属(*Reseda*)、常春藤属(*Hedera*)、玄参属(*Scrophularia*)、栗属(*Castanea*)、百脉根属(*Lotus*)、向日葵属(*Helianthus*)和蓝蓟属(*Echium*)等花粉为主的12类。聚类分析将89个样品分成7类,即第1类以栗属花粉为主的单花蜜;第2类以木犀草属花粉为主的单花蜜;第3类以百脉根属花粉为主的单花蜜;第4类以金雀儿和树状欧石楠花粉为主的多花蜜;第5类以多花蜜和单花蜜(以驴食草、向日葵为主)混合为特征;第6类以欧石楠属花粉为主的单花蜜;第7类以常春藤属花粉为主的单花蜜。

Upadhyay和Bera^[17]通过对印度东部奥里萨邦沿

海地区34个天然蜂蜜的花粉数据进行主成分分析,将其中的26个单花蜜分成4类:第1类以水黄皮属(*Pongamia*)花粉为特征;第2类至少含有番石榴属(*Psidium*)、椰子属(*Cocos*)、桉属(*Eucalyptus*)、芸苔属(*Brassica*)、紫薇属(*Lagerstroemia*)和胡荽属(*Coriandrum*)等6种花粉中的一种;第3类以刺葵属(*Phoenix*)花粉为特征;第4类含蒲桃属(*Syzygium*)花粉。蜂蜜孢粉分析则将26个单花蜜分为17类,分别以胡荽属、桉属、芸苔属、刺葵属、蒲桃属、椰子属、番石榴属、水黄皮属、芒果属(*Mangifera*)、蕃茄属(*Lycopersicon*)、柑橘属(*Citrus*)、含羞草属(*Mimosa*)、山柑属(*Capparis*)、黄檀属(*Delbergia*)、辣木属(*Moringa*)、枣属(*Zizyphus*)和久树(*Schleichera oleosa*)等花粉为主。

本研究采用聚类分析将山西中部地区19个蜂蜜样品分为8类,而采用主成分分析将它们分为5类(见研究结果部分)。蜂蜜孢粉分析根据样品中优势花粉类群的丰度将它们分为5个多花蜜和14个单花蜜,其中单花蜜分别以枣、刺槐、荆条、国槐、臭椿、菊科和豆科等7种花粉类型为主^[18]。

通过比较发现,根据蜂蜜中花粉组成对样品进行分类时,数量分类方法和蜂蜜孢粉分析方法得到的结果之间往往存在一定的相似性和差异性,而造成这种差异的主要原因在于数量分类方法不仅依据蜂蜜中的优势花粉类群的丰度,而且也兼顾其他非优势花粉类群的丰度,然而蜂蜜孢粉分析方法通常只关注样品中的优势花粉类群的丰度。

4 结论

通过以山西中部地区为例,初步探讨了系统聚类分析和主成分分析在蜂蜜孢粉学中的应用。由于数量分类方法同时兼顾了蜂蜜样品中的优势和非优势花粉类群的丰度,所以该方法能够更加全面客观地对蜂蜜样品进行分类,而且分类结果直观,易于理解。但是在采用系统聚类分析实现对样品的分类时,通常需要综合考虑分类阈值 λ 的取值问题,以便得到更加合理的分类结果。

参考文献

- 1 Xu K X. Numerical Taxonomy (in Chinese). Beijing: Science Press, 1994 [徐克学. 数值分类学. 北京: 科学出版社, 1994]
- 2 Peng Y S, Chen L, Li J Q. Study on numerical taxonomy of *Quercus* L. (Fagaceae) in China (in Chinese). J Wuhan Bot Res, 2007, 25: 149-157 [彭焱松, 陈丽, 李建强. 中国栎属植物的数量分类研究. 武汉植物学研究, 2007, 25: 149-157]

- 3 Zhang G J, Zhao L Y, Wang F, et al. Study on quantitative taxonomy of cultivars in *Rosa multiflora* (in Chinese). J Shandong Agric Uni (Nat Sci), 2006, 37: 175–180 [张广进, 赵兰勇, 王芬, 等. 蔷薇品种的数量分类学研究. 山东农业大学学报(自然科学版), 2006, 37: 175–180]
- 4 Zhao B, Luo X Y, Zhang Q X. A study on numerical classification of the wintersweet cultivars (in Chinese). Acta Horti Sin, 2007, 34: 947–954 [赵冰, 雒新艳, 张启翔. 蜡梅品种的数量分类研究. 园艺学报, 2007, 34: 947–954]
- 5 Gu C H, Wang S X, Tian M. Numerical classification of the *Lagerstroemia indica* cultivars (in Chinese). J Zhejiang Forest Coll, 2010, 27: 903–907 [顾翠花, 王守先, 田苗. 紫薇品种数量分类研究. 浙江林学院学报, 2010, 27: 903–907]
- 6 Wang X H, Yuan Z H, Feng L J, et al. Study on numerical classification of *Dahlia* cultivars (in Chinese). Shandong Agric Sci, 2012, 44: 20–23 [王晓慧, 苑兆和, 冯立娟, 等. 大丽花品种数量分类研究. 山东农业科学, 2012, 44: 20–23]
- 7 Zhang Y, Wang M M, Zhang X, et al. The study on numerical taxonomy of *Cypripedium* in North and Northeast China (in Chinese). Acta Horti Sin, 2013, 40: 933–942 [张毓, 王苗苗, 张雪, 等. 中国华北和东北地区杓兰属植物的数量分类研究. 园艺学报, 2013, 40: 933–942]
- 8 Zhang X L, Zhou H G, Zhou Y, et al. A numerical taxonomic study on *Lilium* cultivars (in Chinese). Guihaia, 2002, 20: 325–328 [张西丽, 周厚高, 周焱, 等. 百合品种间的数量分类研究. 广西植物, 2002, 20: 325–328]
- 9 Luo J J, Bai B B, Li H, et al. Cultivars investigation and numerical classification of ornamental *Malus* spp. in Beijing (in Chinese). Chin Agric Sci Bull, 2012, 28: 290–296 [骆菁菁, 柏斌斌, 李虹, 等. 北京地区海棠品种调查与数量学分类. 中国农学通报, 2012, 28: 290–296]
- 10 Zhang F, Shangguan T L, Xie S L, et al. Study on numerical classification of the pollen of honey source plants from Shanxi (in Chinese). J Wuhan Bot Res, 1994, 12: 299–303 [张峰, 上官铁梁, 谢树莲, 等. 山西蜜源植物花粉的数量分类研究. 武汉植物学研究, 1994, 12: 299–303]
- 11 Zeng Z J, Wang K F, Yu X M. Study on numerical taxonomy of nectariferous plants pollen (in Chinese). Apic Sci Technol, 1999, 6: 7–9 [曾志将, 王开发, 俞小明. 蜜源植物花粉数量分类研究. 养蜂科技, 1999, 6: 7–9]
- 12 Zeng Z J, Wang K F. Study on numerical taxonomy of minerals in nectariferous plants pollen (in Chinese). Apic Sci Technol, 2001, 6: 2–4 [曾志将, 王开发. 蜜源花粉中矿物质数量分类研究. 养蜂科技, 2001, 6: 2–4]
- 13 Song X Y, Yao Y F. Advances in melissopalynology (in Chinese). Chin Agric Sci Bull, 2009, 25: 7–12 [宋晓彦, 姚铁锋. 蜂蜜孢粉学研究进展. 中国农学通报, 2009, 25: 7–12]
- 14 Devillers J M, Morlot M, Pham-Delegue H, et al. Classification of monofloral honeys based on their quality control data. Food Chem, 2004, 86: 305–312
- 15 Aronne G, De Micco V. Traditional melissopalynology integrated by multivariate analysis and sampling methods to improve botanical and geographical characterisation of honeys. Plant Biosyst, 2010, 144: 833–840
- 16 Herrero B, Valencia-Barrera R M, Martin R S, et al. Characterization of honeys by melissopalynology and statistical analysis. Can J Plant Sci, 2002, 82: 75–82
- 17 Upadhyay D, Bera S. Pollen spectra of natural honey samples from a coastal district of Orissa, India. J Apic Res, 2012, 51: 10–22
- 18 Song X Y, Yao Y F, Yang W D. Pollen analysis of natural honeys from the central region of Shanxi, North China. PLoS One, 2012, 7: e49545
- 19 Xu K X, Li D Z. A preliminary research on the numerical classification of the genus *Panax* from China (in Chinese). Acta Phytotax Sin, 1983, 21: 34–43 [徐克学, 李德中. 我国人参属数量分类研究初试. 植物分类学报, 1983, 21: 34–43]

Numerical taxonomy in melissopalynology: A case study in the central region of Shanxi, China

SONG XiaoYan¹ & YAO YiFeng²

¹ College of Agronomy, Shanxi Agricultural University, Taigu 030801, China;

² State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

Melissopalynology is an important branch of palynology that deals with the pollen in honey. In the present study, 19 honey samples from the central region of Shanxi and 61 pollen types from those honeys were selected as operational taxonomic units and unit characters, respectively. Hierarchical (Q-type) cluster analysis and principal component analysis were conducted on the samples to create a numerical taxonomy. The two analyses divided the 19 samples into 8 and 5 categories, respectively. The results showed that (i) numerical taxonomy is feasible for the melissopalynological classification of honey samples, and the taxonomical results are easy to understand. (ii) More attention should be paid to the threshold value (λ) when using cluster analysis. Different λ values may yield different taxonomical results. Therefore, the λ value should be based on the actual situation so that very similar samples will be categorized together and dissimilar samples will be distinguished from one another. (iii) Finally, both the numerical taxonomy and melissopalynological analyses emphasize the frequencies of dominant pollen types, so their taxonomical results are similar. Numerical taxonomy also considers the frequencies of non-dominant pollen types, so the taxonomical results may differ from those of melissopalynological analysis.

melissopalynology, numerical taxonomy, cluster analysis, principal component analysis

doi: 10.1360/N972014-00073