

睡莲品种数量分类的初步研究

柏斌斌, 骆菁菁, 李虹, 俞红强, 游捷

(中国农业大学 园林植物与观赏园艺系, 北京 100193)

摘要:以《睡莲》中所描述的 245 个品种中的 101 个为运算单位, 利用 SAS 统计软件进行了主成分分析和 R 型聚类分析。结果表明: 品种描述所用的 16 个性状独立性强, 其中株型、花径、叶面积和株叶展幅对睡莲分类有重要影响。根据国际法规的规定, 系统聚类(离差平方和法)结合快速聚类的结果, 又考虑到睡莲品种分类的现状, 建议将睡莲分为热带大花睡莲品种群、热带中花睡莲品种群、热带小花睡莲品种群、耐寒小花睡莲品种群、耐寒中花睡莲品种群、耐寒大花睡莲品种群等 6 个品种群。该框架既符合现在睡莲品种分类方法的内涵, 又符合国际法规的形式。

关键词:睡莲; 品种; 数量分类; 聚类分析

中图分类号: S 682.32 文献标识码: A 文章编号: 1001-0009(2011)22-0075-04

睡莲是睡莲科睡莲属植物的统称。在我国原产的睡莲只有子午莲(*Nymphaea tetragona*)等数种, 而大量的睡莲种与品种都是从国外引进的。《睡莲》中提到从外国引进 31 个热带睡莲品种和 119 个耐寒睡莲品种。同时还自行培育了 24 个热带睡莲品种和 72 个耐寒睡莲品种。根据记载, 国际上白天开花热带睡莲有 133 个品种, 晚上开花热带睡莲 20 个品种, 还有耐寒睡莲 283 个品种^[1]。

关于睡莲品种分类研究, 迄今未见国内外的有关资料发表。最早关于睡莲属植物种的分类系统是 Dr. Henry Conard^[1]在 1905 年建立的, 采用传统的分类方法即形态分类。其一级分类标准是根据心皮是离生还是聚合, 把睡莲属植物划分为 2 组 5 个亚组, 即离生心皮组(Apocarpiae)包括缺柱亚组(Anepchya)和矮柱亚组(Brachyceras), 聚合心皮组(Syncarpiae)则包括南非睡莲亚组(Castalia)、棒柱睡莲亚组(Hydrocallis)和带柱睡莲亚组(Lotus)^[2]。在国外现有的睡莲著作和商品名录中是按生态类型来分的。可以分为热带睡莲类, 包括白天开花组和晚上开花组, 耐寒睡莲类包括具 Marliac 型块茎组、具 Odorata 型块茎组、具 Tuberosa 型块茎组、具指型块茎组和具地下走茎型组^[1]。但这些分类系统都未与国际接轨, 因此提出符合 ICNCP 规定的睡莲品种分类框架是迫切需要的。关于睡莲属植物分子系统发育研究, 只有刘艳玲等^[3]曾对睡莲属 11 个种的 ITS 序列进行分析, 探讨了其种间的系统发育关系。

第一作者简介: 柏斌斌(1987-), 女, 在读硕士, 研究方向为园林植物种质资源与遗传育种。E-mail: baibinbin121@163.com。

责任作者: 游捷(1963-), 男, 本科, 副教授, 研究方向为园林植物种质资源与遗传育种。E-mail: jamy10@sohu.com。

收稿日期: 2011-08-11

20 世纪 50 年代数量分类学的诞生把数学方法和计算机技术引入到植物分类研究中, 使其从定性描述走向了精确定量的分析水平。数量分类方法在观赏植物分类上的应用相对较晚, 但也已经成功运用于海棠、腊梅、玫瑰、蔷薇、百合、牡丹等^[4-9]多种观赏植物品种的分类研究中。同时在番茄、枇杷^[10-11]等蔬菜的品种分类中发挥重要作用。但尚未见到对睡莲品种进行数量分类的报道。

1 性状编码及分析方法

1.1 分类性状的选取和编码

根据《睡莲》^[1]对睡莲性状的描述, 同时借鉴国际睡莲品种注册登录时对相关性状描述的要求, 最终选取了详细记录的 16 个性状。性状分为 3 种类型: 数值性状(4 个)、二元性状(3 个)、有序多态性状(9 个), 并编码将其数量化(表 1)。

性状的编码采用等级数量编码的方法, 二元性状按 0 和 1 进行编码, 可以区分性状间进化关系的, 将原始性状编为 0, 进化性状编为 1; 不易确定的, 一般是为 1, 否为 0。有序多态性状, 尽可能按从原始到进化的顺序, 取连续排列的非负整数 1、2、3、... 进行编码。一些性状采用数据变换的方法以得到稳定的结果(如叶片面积=长×宽)。数值性状不编码, 直接以原始数据形式进入下一步运算^[12]。

1.2 数据处理

利用 101 个睡莲品种作为运算单位, 利用 Spss 11.5 软件对 16 个指标进行主成分分析、R 型聚类分析, 分析各性状之间的相关性, 并验证所选的性状是否合理。为了消除不同量纲对数据分析产生的影响, 首先对原始数值矩阵进行标准差标准化(STD)处理, 即正规化处理。利用软件对处理过的数据进行分类, 作出性状指标聚类图。

表 1 睡莲聚类分析所选取的性状及编码

序号	性状	类型	0	1	2	3	4	5	6
1	株型	多态		小型	中小型	中型	中大型	大中型	大型
2	花色	多态		白色	粉色	黄色	紫红色	紫蓝色	复色
3	花态	多态		杯状	放射状	碟状	星芒状	碗状	
4	花径	数值							
5	花瓣数	数值							
6	叶面积	数值							
7	株叶展幅	数值							
8	叶片近轴面颜色	多态		草绿色	黄绿色	绿色	橄榄绿色		
9	叶片远轴面颜色	多态		红色	紫色	淡灰色	褐色	绿色	
10	叶耳	多态		分离	部分分离	重合			
11	花药颜色	多态		黄色	紫红色	紫蓝色			
12	萼片 4 枚	二元	否	是					
13	花梗有毛	二元	否	是					
14	叶形	多态		圆形	近圆形	椭圆形	卵圆形	肾心形	
15	花梗颜色	多态		绿色	褐色	紫色	肉红色	古铜色	
16	叶柄和花梗颜色相同	二元	否	是					

样本聚类也要先对数据进行预处理,对样本进行快速聚类,并利用多种方法对样本进行 Q 型聚类。在聚类分析过程中,种质间距离为欧氏距离。

2 结果与分析

2.1 性状之间的相关性

对 101 个睡莲品种的 16 个性状进行主成分分析,得相关系数表。其中相关系数达到极显著水平的性状有株型与花径、花径与株叶展幅,它们都呈明显的正相关,反映了睡莲的株型越大,花径越大。花径与株叶展幅的相关,反映了睡莲的花径与叶性状相关。其它性状彼此之间也存在着或强或弱的依赖关系。在这些复杂的关系中,如果仅凭经验直观地确定某几个性状来进行分类,就可能造成偏差。

利用 Spss 程序,求得各主成分特征值、贡献率及累积贡献率,取累积贡献率大于 85% 的前几个主成分,结果如表 2。

表 2 重要主成分的特征值

	主成分					
	1	2	3	4	5	6
株型	.708	.312	.216	-.339	-.054	.051
花色	.547	-.347	-.221	-.034	-.073	-.134
花态	.042	.185	-.180	.106	.728	-.134
花径	.751	.317	.292	-.105	-.016	.118
花瓣数	-.231	.595	-.388	-.012	-.234	.064
叶面积	.844	-.022	.009	.049	-.042	-.122
株叶展幅	.787	.219	.213	-.029	.053	.034
叶片近轴面颜色	.059	-.362	.366	.181	-.015	-.547
叶片远轴面颜色	.229	-.323	-.464	-.129	.196	.347
叶耳	.644	-.022	-.055	.386	.121	-.016
花药颜色	.406	-.355	-.161	.333	-.355	.264
萼片 4 枚	-.102	-.247	.368	.232	.448	.566
花梗有毛	-.328	.094	.715	-.118	-.005	.104
叶形	.079	-.124	-.129	-.811	.188	-.007
花梗颜色	.067	.476	-.116	.230	.388	-.293
叶柄和花梗颜色相同	-.054	.639	-.012	.206	-.199	.228

由表 2 可知,第一主成分中,叶面积、株叶展幅、花径和株型的系数较大,是对第一主成分影响最大的特征向量,说明第一主成分是由它们几个组成的一个综合指标,主要反映了植物整个外部形态。第二主成分中,花瓣数和花梗是主要决定因子,可以综合为花因子。第三主成分中,花梗有毛所占比重最大,综合为花

梗因子。第四主成分中,叶形因子负荷量最大,主要反映其叶形状态。第五主成分中,花态占最大比重,将其命名为花型因子。第六主成分中,萼片数和叶面近轴面颜色占的比重较大,主要反映了萼片与叶色情况。综上所述,各个性状的贡献率比较均衡,而且叶面积、花径和花瓣数对睡莲分类的贡献较高。

为了进一步研究睡莲品种各性状之间的关系,对以上 16 个性状进行了 R 型聚类(图 1)。结果表明,所选性状相关系数均小于 0.68,没有明显的聚为几类。有少数性状在聚类图上表现为两两相关,如 1(株型)和 4(花径),6(叶面积)与 7(株叶展幅),但它们之间并没有必然的联系。

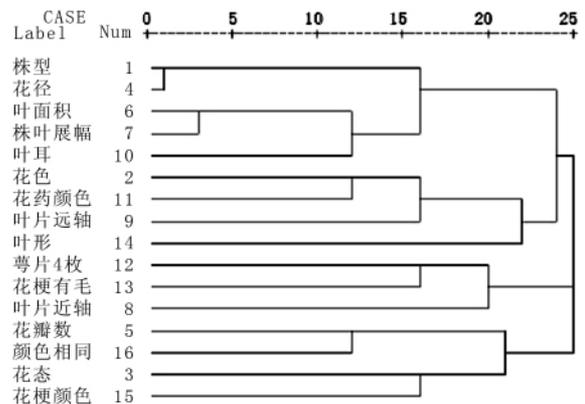


图 1 睡莲品种 R 型聚类树系聚类图

根据主成分分析和 R 型聚类的分析结果,选取了 16 个性状指标作为睡莲 101 个品种的聚类分析的相关性状。这 16 个性状间的相关性不强,分布分散,没有明显地分成几个大组,性状对进化均具有独立作用,对以后进行睡莲品种分类性状的选择起到一定的指导作用。

2.2 品种聚类分析

采用 Spss 通过标准化过程对睡莲 101 个品种进行预处理,再进行系统聚类分析。聚类分析采用了最短距离法、最长距离法、重心法、离差平方和法 4 种方法,对比发现离差平方和法最佳。为便于分析,做等级结合线 Lq1、Lq2 和 Lq3。

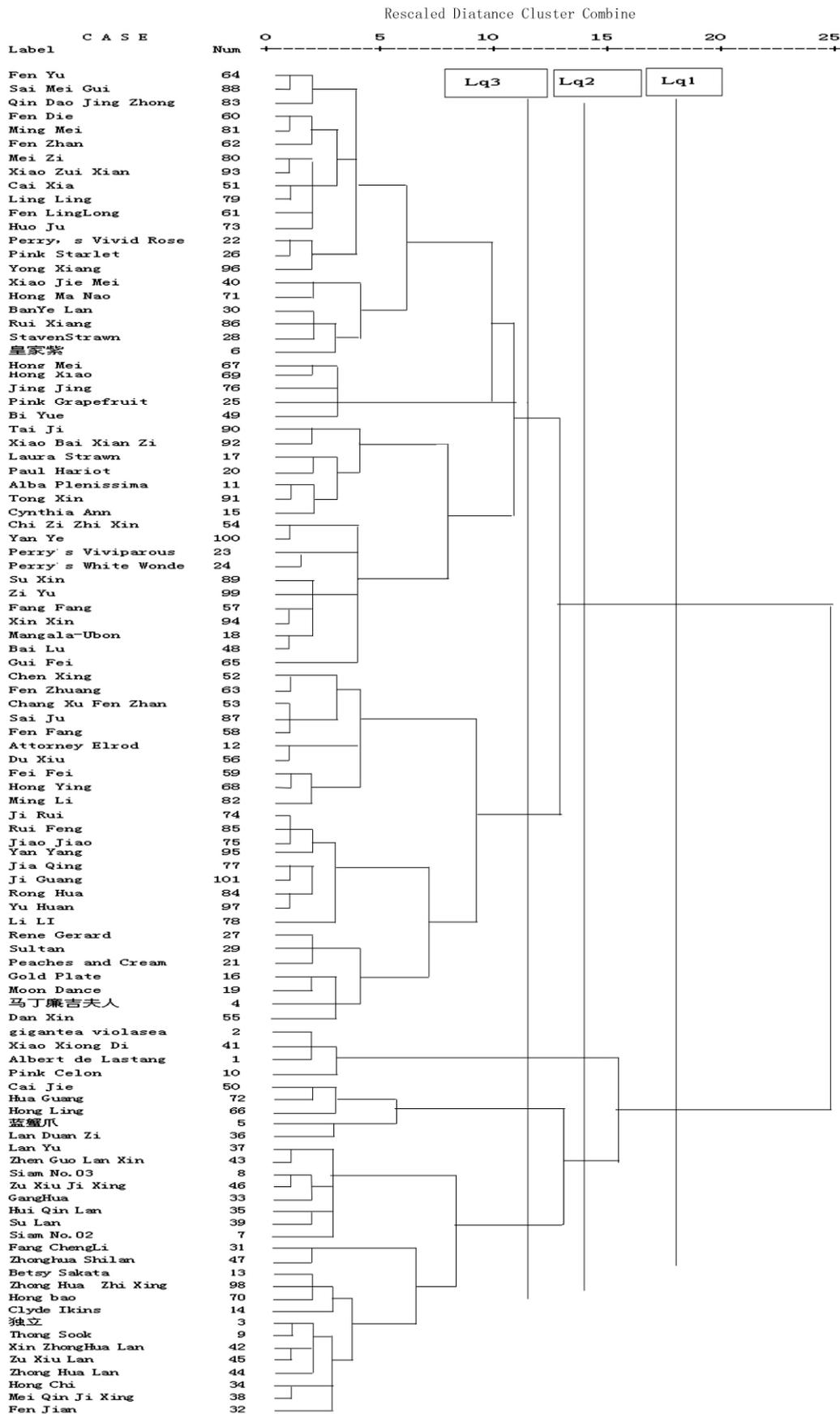


图2 睡莲品种 Q 型聚类树系图

从图 2 可看出,在结合线 Lq1 处将 101 个品种分为 2 个类群,在结合线 Lq2 处分为 3 个类群,在结合线 Lq3 处分为 5 个类群。

在 Lq1 处分为 2 个类群,主要依据其生态习性分成耐寒睡莲和热带睡莲,但其中也有个别品种划分依据不同。比如 40 ‘Xiaojiemei’、30 ‘Banyelan’、6 ‘RoyalPurple’本属于热带睡莲却在耐寒睡莲里。还有耐寒睡莲 50 ‘Caijie’、72 ‘Huaguang’、66 ‘Hongling’、13 ‘Betsy sakata’、98 ‘Zhonghuazhixing’、70 ‘Hong bao’和 14 ‘Clyde Ikins’却划分在热带睡莲里。从上述聚类分析结果可以看出,Q 型聚类分类结果基本上与传统睡莲品种的分类结果一致。

在 Lq2 处分为 3 个类群,主要依据是叶面积大小。第一类叶面积都偏小,小于 700 cm²。第三类叶面积偏大,大于 700 cm²。第二类根据花瓣数分出的,其花瓣数偏少,都小于 20 枚。其它 2 类都比较多。

在结合线 Lq3 处依据叶面积分为 5 个类群。同时调用 K-Means 聚类分析对睡莲 101 个品种进行快速聚类,共聚成 5 类。得出的最终结果是:第 1 类叶面积小(257 cm²),第 2 类叶面积较小(474 cm²),第 3 类叶面积较大(849 cm²),第 4 类叶面积较大(1 890 cm²),第 5 类叶面积大(1 340 cm²)。

3 讨论与结论

3.1 关于 Q 型聚类分析

根据 Q 型聚类分析所得的结果,定量地揭示了睡莲品种间的亲缘关系,根据数量分类的原理可知^[14],亲缘关系越近,就越早地聚合为一类,结合水平就越高;反之,亲缘关系越远,就越晚地聚合为一类,结合水平就越低。根据国际法规的规定,结合聚类的结果,又考虑到睡莲品种分类的现状,《国际栽培植物命名法规》(ICNCP7th)的规定,品种分类原则上只能采用品种群(Group, Gp)一个层次。该文将睡莲品种聚为 5 类的主要依据是叶面积,由于叶面积与花径的相关性比较大,因此可以以花径作为划分品种群的主要依据。考虑到传统分类对睡莲品种的深刻影响,在分类过程中考虑到生态学为基础的分类方法。最终将选取的睡莲 101 个品种分为 6 个品种群:热带大花睡莲品种群

(>21 cm)、热带中花睡莲品种群(14~21 cm)、热带小花睡莲品种群(<14 cm)、耐寒大花睡莲品种群(>15 cm)、耐寒小花睡莲品种群(<12 cm)、耐寒中花睡莲品种群(12~15 cm)。

3.2 关于性状选取

在选取性状过程中,根据《睡莲》中相关记载的 28 个性状和 246 个品种,选择其中 16 个性状记载齐全的 101 个品种。这说明现有睡莲品种中只有 57.14% 的性状记载比较齐全,41.06% 的品种 16 个性状记载齐全。这也就预示着睡莲性状描述过程中存在的问题:记录不规范。而且那些没有被用来做品种分类依据的 12 个性状中也是可以选取的,比如心皮附属物的形状和大小、颜色等也可以作为识别品种的标志。但由于这个性状记载不完全,所以不能采纳。这就要求在以后的研究中要注意性状记录的规范性。

参考文献

- [1] 黄国振,邓惠勤,李祖修,等.睡莲[M].北京:中国林业出版社,2009.
- [2] Swindells P. Waterlilies [M]. Oregon: Croom Timber Press, 1983.
- [3] 刘艳玲,徐立铭,倪学明,等.基于 ITS 序列探讨睡莲属植物的系统发育[J].武汉大学学报(理学版),2005,51(2):258-262.
- [4] 楚爱香,杨英军,汤庚国,等.河南垂丝海棠品种数量分类研究[J].园艺学报,2009,36(3):377-384.
- [5] 赵凯歌,虞江晋芳,陈龙清.蜡梅品种的数量分类和主成分分析[J].北京林业大学学报,2004,26:79-83.
- [6] 于守超,丰震,赵兰勇.平阴玫瑰品种数量分类研究的探讨[J].园艺学报,2005,32(2):327-330.
- [7] 张广进,赵兰勇,王芬,等.蔷薇品种的数量分类学研究[J].山东农业大学学报(自然科学版),2006,37(2):175-180.
- [8] 荣立苹,雷家军.东北地区野生百合数量分类研究[J].植物遗传资源学报,2010,11(1):99-102.
- [9] 龚双军,万卉敏,杨秋生.牡丹栽培品种数量分类的初步研究[J].江西农业学报,2009,21(12):66-69.
- [10] 张学时,李晴,韩玉珠,等.吉林省主栽番茄品种的数量分类研究[J].北方园艺,2010(4):38-40.
- [11] 胡文舜,骆键夫,黄爱萍,等.枇杷种质数量分类中系统聚类方法的探讨[J].福建果树,2009(2):15-22.
- [12] 陈俊愉.中国花卉品种分类学[M].北京:中国林业出版社,2001.
- [13] 张春英,戴思兰.桃花种资源的数量分类学研究[J].北京林业大学学报,1999,21(3):41-45.
- [14] 徐克学.数量分类学[M].北京:科学出版社,1994:94-110.

Preliminary Study on the Numerical Classification of *Nymphaea tetragona* Cultivars

BAI Bin-bin, LUO Jing-jing, LI Hong, YU Hong-qiang, YOU Jie

(Department of Ornamental Horticulture and Landscape Architecture, China Agricultural University, Beijing 100193)

Abstract: Using 101 *Nymphaea tetragona* cultivars from 245 cultivars described in *Nymphaea tetragona* as operational units, then carried through principle component analysis and R-cluster analysis by statistic software of Spss. The results showed that the 16 characters were independent strongly and the factors such as the plant types, flower diameter, leaf area and the spread of plant occupied the important status in classification. In accordance with international regulations, the results of fastclus analysis and Q-cluster analysis and the status of *Nymphaea tetragona* category, the authors proposed that *Nymphaea tetragona* was divided into 6 groups, consisting of large-sized-flower tropical water lilies group, medium-sized-flower tropical water lilies group, Small-sized-flower tropical water lilies group, large-sized-flower hardy water lilies group, Medium-sized-flower hardy water lilies group and small-sized-flower hardy water lilies group. The framework embodies the content of ecological classification, and accord with international regulations in the form.

Key words: *Nymphaea tetragona*; cultivar; establishment; cluster analysis