

·技术与方法·

基于 MATLAB 的花粉形态量化分析

贾红娟¹, 张军文^{2*}

¹ 河北师范大学资源与环境科学学院, 石家庄 050016; ² 军械工程学院, 石家庄 050003

摘要 利用数字照相技术和MATLAB软件测量了5种杨树、3种柏树各50粒花粉的面积、扁圆率、极轴长度、赤道轴长度, 计算了面积标准差和极轴长度/赤道轴长度。对8种花粉的面积作了方差分析。结果显示, 在99%置信水平上用面积能区分8种花粉。用5种杨树花粉的扁圆率平均值、极轴平均值、赤道轴平均值和P/E值作聚类分析, 聚类结果显示, 5种杨树可分成3组, 加拿大杨和小叶杨是一组, 毛白杨和山杨是一组, 青杨单独为一组。电镜观察纹饰特点与以上聚类结果一致。本研究为定量研究花粉形态提供了一种快速有效的方法及更多的有效数据, 比目测精确度提高了10倍。

关键词 聚类分析, 数字照片, MATLAB, 花粉形态, 方差分析

贾红娟, 张军文 (2007). 基于MATLAB 的花粉形态量化分析. 植物学通报 24, 511-515.

MATLAB软件是美国Math Works 公司于1984年推出的, 现已成为最优秀的科学与工程计算软件工具之一。MATLAB 是解释型语言, 其语法规则非常类似BASIC 语言, 易学易用。它提供了 30 多个工具箱, 包括功能性工具箱和学科性工具箱, 如控制工具箱、图像工具箱、小波工具箱等。借助这些工具箱, 研究人员可以专注于个人领域的问题, 不必过多考虑计算机的实现方法。MATLAB具有强大的数值计算功能(陈桂明等, 2002), 许多复杂的计算问题在MATLAB中只需几行代码就可以解决。

随着数码照相机的广泛使用, 对数字图像的处理就变得非常重要。MATLAB 是基于矩阵的语言, 从本质上提供了对图像的支持, 它可以对由数字图像形成的离散数据矩阵进行各种处理。利用MATLAB可以对显微图片结构进行准确而又细致的分析, 从而能更好地对显微物体进行外形分析(杨云和黄亮, 2006)。成芳和应义斌(2004)利用MATLAB成功地进行了稻种的形状分析和识别。MATLAB图像处理工具箱中提供了数字图像处理的大多数经典算法, 有100多个工具函数, 它可

以让使用者把精力集中在算法而不是编程上。MATLAB 具有很强的开放性和适应性, 可以根据需要对函数进行修改和扩展, 方便用户使用(董长虹等, 2004)。

对花粉形态的研究, 以往人们多是在光学显微镜下用测微尺目测花粉的大小(Erdtman, 1962; 王开发和王宪曾, 1983; 王伏雄等, 1995; 王萍莉, 1997; 周忠泽等, 2004), 这种方法易受人为因素影响, 且较繁琐。随着计算机技术、数字图像技术的发展, 摆脱人工方法, 找到更简单快捷的方法测量花粉成为可能。曾志将等(2000)利用计算机对花粉电镜照片进行了研究。我们尝试着用日本OlympusBX51光学生物显微镜40倍物镜及700万像素的Olympus数码相机3倍光学变焦下对几种杨树和柏树花粉照相, 利用MATLAB程序对数字照片进行计算, 得到各种花粉的扁圆率、面积、极轴、赤道轴等各种参数, 为测量花粉形态参数提供一种快速有效的方法。利用获得的原始数据进行了方差分析和聚类分析, 把对花粉分析从定性分析上升到定量分析, 为植物系统演化提供证据。

收稿日期: 2006-07-31; 接受日期: 2006-10-30

基金项目: 国家自然科学基金(No.40571166)和国家教委博士点基金(No. 20050094004)

* 通讯作者。E-mail: jhjzjw@sina.com

1 材料和方法

1.1 材料来源

观察材料均为现采花粉(表 1)。

1.2 采样和照相

采样时统一选含苞待放的花蕾, 以确保花粉成熟度一致。光学显微镜观察材料均经过 Erdtman 醋酸酐分解法处理(Erdtman, 1962)。在日本 OlympusBX51 光学生物显微镜 40 倍物镜及 Olympus 数码相机 3 倍光学变焦条件下对花粉照相。在显微镜下选取较清晰、粘连少、视野中品相好的花粉区域照相, 相同倍数照下测微尺, 作为最后换算成公制单位的依据。

1.3 照片的预处理

把选好的照片用 photoshop 进行处理, 去掉背景, 选择合适的花粉复制到一个新文件中。在新文件中用橡皮擦工具擦去杂质和有缺陷的花粉(不必擦得很仔细, 在以后的计算时会自动除去面积小的杂质), 保存成 JPG 格式文件。

1.4 在 MATLAB 中计算(主要代码如下所示)

首先对彩色照片进行二值化处理, 填充花粉图像中的空洞, 移除杂质(面积小的对象), 计算整个照片中的花粉个数, 标记每个花粉。

```
photo=imread('simonii.jpg');
i=graythresh(photo);
a1=im2bw(photo,i);
```

```
a2=imcomplement(a1);
a3 = imfill(a2,'holes');
clean_ptoto=bwareaopen(a3, 5000);
[l,num] = bwlabel(clean_ptoto);
以最小矩形裁剪单个花粉, 旋转到水平位置, 计算每个花粉的各种参数, 并以 EXCEL 文件的形式保存。
Min-rectangle=regionprops(l,'BoundingBox');
for n=1:50
pollen= Min-rectangle(n).BoundingBox;
l1=imcrop(clean_ptoto,[pollen(1) pollen(2) pollen
(3) pollen(4)]);
Monad-pollen=bwareaopen(l1,5000);
lxie=bwlabel(Monad-pollen);
angle=regionprops(lxie,'Orientation');
angle1=angle.Orientation;
horizontal =imrotate(Monad-pollen,-angle1, 'bilinear');
L=bwareaopen(horizontal, 5000);
L1 = bwlabel(L);
ans11=regionprops(L1,'all');
pollenArea(n)=ans11.Area;
polleneccentricity(n)=ans11.Eccentricity;
pollenbox=ans11.BoundingBox;
pollenboxX(n)=pollenbox(3);
pollenboxY(n)=pollenbox(4);
end
aa(:,1)=pollenArea;
aa(:,2)=pollenEccentricity;
```

表 1 材料来源
Table 1 Source of materials

Species	Locality	Collector
<i>Populus canadensis</i> Moench	Hebei Normal University, Shijiazhuang	Li Yuecong, Jia Hongjuan
<i>P. tomentosa</i> Carrière	Hebei Normal University, Shijiazhuang	Li Yuecong, Jia Hongjuan
<i>P. cathayana</i> Rehder	Tuoliang, Shijiazhuang	Li Yuecong, Jia Hongjuan
<i>P. davidiana</i> Dode	Xiaowutai Mt., Hebei	Li Yuecong, Jia Hongjuan
<i>P. simonii</i> Carrière	Xiaojiao, Shijiazhuang	Li Yuecong, Jia Hongjuan
<i>Sabina chinensis</i> (Linn.) Ant.	Hebei Normal University, Shijiazhuang	Li Yuecong, Jia Hongjuan
<i>Juniperus formosana</i> Hayata	Hebei Normal University, Shijiazhuang	Li Yuecong, Jia Hongjuan
<i>Platycladus orientalis</i> (Linn.) Franco	Hebei Normal University, Shijiazhuang	Li Yuecong, Jia Hongjuan

```
aa(:,3)=pollenboxX;  
aa(:,4)= pollenboxY;  
xlwrite('simonii.xls',aa);  
取出最大、最小和平均值。  
max(pollenArea)  
min(pollenArea)  
mean(pollenArea)  
.....  
max(pollenboxY)  
min(pollenboxY)  
mean(pollenboxY)
```

计算机自动测量较手工作业精度和效率更高,只要准备好数字照片输入到程序中,几分钟就可以得到大量的数据。在光镜下目测最大分辨率一般是1 μm(王伏雄等, 1995)。在数字照片中1 μm 可以用9.4 个像素点来表示,这种方法的精度就可以达到0.106 μm,这样的精度可以使测量误差忽略不计。误差主要由于花粉的选择是随机的,所以要去掉畸形及折叠明显的花粉,以减少误差。利用获得的大量数据,在计算机上通过MATLAB 程序作量化分析,把对花粉定性的文字描述变为量化的数学分析。用数量表示比用文字描述能更好地对不同种花粉进行区分,为植物系统演化和分类学研究提供量化证据。

2 结果与分析

2.1 原始数据分析

由于根据文献中(王伏雄等, 1995; 王萍莉, 1997; 周忠泽等, 2004; 刘家熙和阎秀峰, 2005; 徐进和王章荣, 2005)普遍采用的20和30粒花粉计算的结果与80、90粒的结果相差很大,通过多次比较认为取50粒比较合适,所以最后统一取50粒计算花粉的面积、扁圆率、极轴长、赤道轴长和P/E等参数,取最大、最小及平均值。

由表2可见,杨树、柏树的花粉赤道轴、极轴的数量值很小,不同种之间的差值更小,以赤道轴、极轴的值为参数不能很好反映杨树、柏树不同种花粉之间的差别,而面积能较好地表现种类之间的差异。

柏树的花粉比杨树的花粉大很多,用其面积平均值的大小就可以区分这3种柏树、5种杨树。刺柏花粉的面积最大,为634.64 μm²,它的扁圆率为0.4257,是比较扁的花粉。侧柏花粉在柏树中比较小,与圆柏相差约50 μm²。在这8种花粉中最圆的扁圆率为0.3507,花粉个体的大小差异最明显,标准差为109.8710 μm²。圆柏花粉标准差为40.1268 μm²,是8种花粉中大小最均匀的。在8种花粉中加拿大杨花粉最小,平均面积只有339.2033 μm²,面积的标准差也只有40.4495 μm²,

表2 8种花粉的计算结果 (面积单位为 μm², 长度单位为 μm)
Table 2 Compute results of eight pollen species (area unit μm², length unit μm)

	Populus					Platycladus	Sabina	Juniperus
	canadensis	P. simonii	P. cathayana	P. tomentosa	P. davidiana	orientalis	chinensis	formosana
Max. area	533.0693	469.4206	460.7402	656.9149	855.9416	721.1747	825.6451	862.5396
Min. area	244.1489	257.9787	281.6772	294.4885	317.4287	398.7438	428.2028	438.1394
Mean area	339.2033	350.6338	363.5695	439.4862	520.6881	552.3766	610.2309	634.6424
Area standard deviation	40.4493	68.6974	79.2904	60.4708	84.8619	109.8710	40.1268	85.8047
Max. eccentricity	0.6075	0.6252	0.5610	0.6157	0.6152	0.5429	0.5677	0.6114
Min. eccentricity	0.1735	0.1517	0.1296	0.1859	0.2238	0.1446	0.1440	0.1380
Mean eccentricity	0.4158	0.4196	0.3563	0.4071	0.409	0.3507	0.4064	0.4257
Max. pole-length	29.1489	26.5957	26.1702	30.2128	35.5319	31.3830	34.1489	35.6383
Min. pole-length	19.2553	19.5745	19.6809	20.1064	21.2766	23.4043	25.0000	25.9574
Mean pole-length	22.1894	22.4787	22.6064	24.9894	27.1766	27.6489	29.5298	30.2787
Max. equator-length	23.5106	22.7660	23.6170	28.1915	31.2766	30.1064	33.2979	32.3404
Min. equator-length	16.0638	16.9149	18.0851	19.0426	18.8298	22.0213	22.2340	21.1702
Mean equator-length	19.8723	20.2596	21.0255	22.6532	24.6489	25.8553	26.9383	27.1319
P/E	1.1166	1.1095	1.0752	1.1031	1.1025	1.0694	1.0962	1.1160

表 3 8 种花粉面积的方差分析

Table 3 Variance analysis about pollen area of eight species

Source	SS	DF	MS	F	F _{0.01}
Columns	3.88086E+010	7	5.54409E+009	127.69**	3.0988
Error	1.70205E+010	392	4.34197E+007		
Total	5.58291E+010	399			

SS: 平方和; DF: 自由度; MS: 均方差; F: 均方值的比值; F_{0.01}: F 在 0.01 水平上的临界值

SS: square sum; DF: degree of freedom; MS: mean square sum; F: square sum among class/square sum inner class; F_{0.01}: the threshold value of F at 0.01

是 8 种花粉中大小比较均匀的。山杨的花粉是 5 种杨树中最大的, 大小也最不均匀。在杨树中, 加拿大杨、小叶杨和青杨的花粉比较小, 毛白杨和山杨的花粉比较大。

2.2 方差分析

为了验证用面积区分花粉是否可行, 我们用这 8 种各 50 粒花粉的面积作方差分析(表 3)。

由表 3 可知, F 值 127.69 远远大于 F_{0.01} 的值 3.0988, 检验结果是呈显著差异, 说明在 99% 的置信水平下用面积区分这 8 种花粉是可行的。

2.3 聚类分析

用加拿大杨、小叶杨、青杨、毛白杨和山杨花粉的面积平均值、扁圆率平均值、极轴平均值、赤道轴平均值和 P/E 值为参数作聚类分析(图 1)。

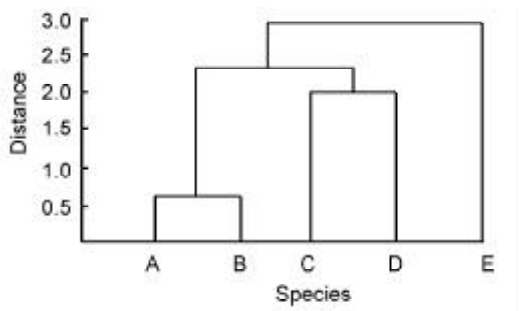


图 1 5 种杨树花粉聚类分析

A: 加拿大杨; B: 小叶杨; C: 毛白杨; D: 山杨; E: 青杨

Figure 1 Cluster analysis of five pollen species about populus

A: *Populus canadensis*; B: *P. davidiana*; C: *P. tomentosa*; D: *P. simonii*; E: *P. cathayana*

由图 1 可以看出, 5 种杨树可分成 3 组, 加拿大杨和小叶杨是一组, 毛白杨和山杨是一组, 青杨单独为一组。在电镜下观察, 加拿大杨和小叶杨呈不规则颗粒状纹饰, 颗粒比较密, 刺突不明显; 毛白杨和山杨也呈颗粒状纹饰, 与加拿大杨及小叶杨相比, 颗粒较稀, 刺突明显; 青杨则为网状纹饰, 与其它 4 种明显不同。电镜观察的纹饰特点与聚类分析结果一致。

以上分析只是在有限数据范围内得出的结果, 是否普遍适用, 有待获得大量数据后进一步验证。

3 结论与讨论

本文提供了一种利用光镜数字照片快速测量花粉形态参数的方法, 可以比目测得到更多的参数。由于花粉本身很小, 目测花粉极轴、赤道轴的误差比较大, 而利用 MATLAB 程序测量花粉形态参数提高了测量精度和效率。原始数据和方差分析表明, 用花粉面积可以区分 5 种杨树、3 种柏树的花粉。对 5 种杨树作了聚类分析, 分析结果与电镜观察纹饰特征一致。本研究把花粉研究从定性分析上升到定量分析, 为植物的分类和系统演化提供依据。电镜成本高, 花粉鉴定中不可能全部用电镜观察花粉, 光镜和数码照相机成本低, 利用光镜照片并借助计算机技术可以降低花粉研究的成本。

这种方法比较适用于圆形和椭圆形花粉, 其它形状的花粉还需要进一步研究。花粉较植物其它器官变异小, 通过计算大量的花粉样本, 建立花粉样本库, 可为进一步的形态分析、系统演化、分类和识别提供量化数据。

致谢 感谢李月丛老师帮助采集花粉, 于和平老师

处理花粉。

参考文献

- 陈桂明, 戚红雨, 潘伟 (2002). MATLAB 数理统计(6.X). 北京: 科学出版社. pp. 136-198.
- 成芳, 应义斌 (2004). 基于 MATLAB 平台的稻种图像分析系统. 浙江大学学报(农业与生命科学版) 30, 572-576.
- 董长虹, 赖志国, 徐啸海 (2004). MATLAB 图像处理与应用. 北京: 国防工业出版社. pp. 196-204.
- 刘家熙, 阎秀峰 (2005). 西藏产四种卷柏科植物的孢子形态观察. 植物学通报 22, 44-59.
- 王开发, 王宪曾 (1983). 孢粉学概论. 北京: 北京大学出版社. pp. 53-54.
- 王伏雄, 钱南芬, 张玉龙, 杨惠秋 (1995). 中国植物花粉形态. 北京: 科学出版社. pp. 9-9.
- 王萍莉 (1997). 中国实用花粉. 成都: 四川科学技术出版社. pp. 18-18.
- 徐进, 王章荣 (2005). 鹅掌楸属种间杂种与亲本花粉壁的超微结构的比较观察. 植物学通报 22, 35-38.
- 杨云, 黄亮 (2006). 运用 MATLAB 实现对显微图片的数字化分析. 机械设计与制造 2, 135-137.
- 曾志将, 王开发, 俞小明, 张卫东 (2000). 花粉形态计算机量化分析研究. 植物学通报 17, 372-374.
- 周忠泽, 张小平, 许仁鑫 (2004). 中国冰岛蓼属植物花粉形态的研究. 植物分类学报 42, 513-523.
- Erdtman G (王伏雄, 钱南芬译) (1962). 花粉形态与植物分类. 北京: 科学出版社. pp. 5-18.

Quantitative Analysis of Pollen Morphology with MATLAB

Hongjuan Jia¹, Junwen Zhang^{2*}

¹ College of Resource and Environment, Hebei Normal University, Shijiazhuang 050016, China

² Ordnance Engineering College, Shijiazhuang 050003, China

Abstract Pollen grains of three species of Cupressaceae and five species of *Populus* were examined in detail by use of a digital photographic technique and MATLAB software. For each pollen species, 50 measurements were taken of maximum, minimum and mean values of area, eccentricity, pole-length and equator-length. P/E and area standard deviations were computed. Variance analysis of the areas of 8 pollen species showed that the method can distinguish the 8 pollen species with reliability of 99%. Cluster analysis of mean eccentricity, pole-length, equator-length and P/E parameter of 5 species of *Populus* showed that the 5 species can be divided into 3 groups. *P. canadensis* Moench and *P. simonii* Carrière form one group; *P. tomentosa* Carrière and *P. davidiana* Dode form another group; *P. cathayana* Rehder stands alone as the third group. This outcome is consistent with scanning electron microscopy results. The method offers a speedy and convenient way for quantitative analysis of pollen morphology, provides more useful data for pollen analysis and improves the accuracy, by about 10 times, than with estimation by the naked eye.

Key words cluster analysis, digital photo, MATLAB, pollen morphology, variance analysis

Jia HJ, Zhang JW (2007). Quantitative analysis of pollen morphology with MATLAB. *Chin Bull Bot* 24, 511-515.

* Author for correspondence. E-mail: jhjzjw@sina.com

(责任编辑: 白羽红)