

杜芳课题组叶片形态分析流程

Leaf morphology analysis process in the Du Fang's lab



杜芳, 李月娟, 祁敏, 苏蔚, 张艺 本资料仅限分子生态实验室内部交流, 有问题 请邮件至杜芳: <u>dufang325@bjfu.edu.cn</u> 最后更新: 4/24/2025 The analysis software obtained from open sources, please cite the relevant websites if you need to use them (relevant references are at the end of the documents).

software download address: Image J: <u>https://imagej.net/ij/download.html</u>

Morpho J: https://morphometrics.uk/MorphoJ_page.html

R: <u>https://www.r-project.org/</u>

—、	叶型扫扫	苗(CanoScan 5600 F 扫描仪)	1
二、	叶型标志	志点(ImageJ 软件)	1
	1.功能		1
	2.下载地	包址	1
	3.操作步	, 骤	1
	3.1	导入图像(Import image)	1
	3.2	标志点采集(Landmarks picking)	2
	3.3	标志点坐标数据处理(Data processing for land	mark
	c00	ordinates)	3
三、	叶型分枝	近(Morpho J 软件)	4
	1.功能		4
	2.下载地	5.址	4
	3.操作步	,骤	4
	3.1	导入文件格式(Input file formats)	5
	3.2	导入文件(Import file)	5
	3.3	前期数据处理(Preliminary data processing)	6
		3.3.1 普式拟合(Procrustes fit)	6
		3.3.2 查找异常值 (Find Outliers)	8
		3.3.3 分离对称和非对称性(Separate symmetry	and
		asymmetry)	8
		3.3.4 提取新分类层(Extract New Classifier)	11

目	录
-	-1-

	3.3.5 生成新的子样本集(Generate new subsamples)	.13
	3.3.6 构建线框(Create wireframes)	.15
	3.4 多变量分析(Multivariate analysis)	.16
	3.4.1 主成分分析(Principal component analyses)	.16
	3.4.2 典型变量分析(Canonical variate analysis)	.26
	3.4.3 判别分析(Discriminant analysis)	.29
	3.4.4 偏最小二乘法(Two-Block Partial Least-Squ	are
	analysis, 2B-PLS)	.31
	3.4.5 普式方差分析(Procrustes ANOVA)	.34
	3.5 导出数据(Export data)	.36
四、	检测叶形差异是否显著	.38
五、	地理、遗传和气候对叶形的影响	.39
	1.获取叶形、地理、遗传及气候因子	.39
	2.叶形与地理、遗传和气候的冗余分析	.39
六、	导入协变量(Import Covariates)	.43
	1.导入文件格式(Input file formats)	.43
4	2. 导入协变量	.44
	3.双标图(Biplot)	.46
七、	映射到系统发育树(Map Onto Phylogeny)	.48
	1. 导入系统发育树	.48
	2.独立变量在系统发育上映射	.49
八、	假定模块化评估(Modularity: Evaluate Hypothesis)	.53

	1.选择标志点子集	53
	2.修改邻接图	54
	3.比较分区(Comparing partitions)	56
九、	深度学习(Deep Learning)	57
	1.选择体系结构	58
	2.选择框架	58
	3.选择模型	59
	4.获取训练集	59
	5.具体操作步骤:	59
	5.1 准备软件	59
	5.2 数据准备及处理	50
	5.2.1 准备用于分类的图片	50
	5.2.2 修改运行命令	51
	5.3 数据上传及运行	53
	5.3.1 数据上传	53
	5.3.2 数据运行	53
+、	测量叶片传统性状指标	54
	1. 基础指标	54
	1.1 批量导入图像	54
	1.2 测量叶长	55
	1.3 像素转换	56
	1.4 其余叶性状指标	57

2. 叶脉密度(Leaf Vein Density).	68
2.1 一级叶脉密度(1° v	vein density)69
2.2 二级叶脉密度(2° v	vein density)69
2.3 小脉密度(Minor ve	ein density)70
3. 比叶面积(Specific Leaf Area)	75
3.1 导入图像(Import in	nage)76
3.2 参数设置(Paramete	r Setting)76
3.3 叶片干重(Dry Leaf	Weight)78
3.4 比叶面积(Specific)	Leaf Area)78
3.5 批量计算图片中叶	十片大小 (叶面积、叶最大直径)
十一、气候因子与叶性状的线性关系	<u></u>
1. 获取叶性状数据	79
2. 获取气候数据	79
3. 叶性状与气候因子的相关分;	析80
十二、R语言常用的作图方法	82
1. 直方图	82
2. 箱型图	
3. 散点图加折线图	
参考文献	

一、叶型扫描(CanoScan 5600 F 扫描仪)

步骤:先将叶片进行压制,干燥,然后使用扫描仪以 600dpi 分辨 率叶背朝下进行扫描,并保存扫描的叶片图像。

二、叶型标志点(ImageJ 软件)

1.功能:通过对样本进行标志点标记来测量面积、平均值、标准偏差、 最大值、最小值、角度以及长度。

2.下载地址: <u>https://imagej.net/ij/download.html</u>

3.操作步骤:

3.1 导入图像(Import image)

选择"File"→"Import"→"Image Sequence"选定并打开一张叶 片图像(图 1, 2)。



图1选择准备导入的叶片图像

🛓 Sequence Option	15	×	
Number of images: Starting image: Increment: Scale images: File name contains: (enclose reg	20 1 1 100 % ex in parens)		
Convert to RGE Sort names nu Use virtual state 2480 x 3507 x 20 OK	3 merically ck (663.6MB) Cancel Helt	2	The second secon
图2编辑导	·入图像数	Y	

3.2 标志点采集(Landmarks picking)

打开"Point Picker":选择"Plugins"→"PointPicker",在出现的 工具条上选择带"+"钢笔头形状的工具在叶片的右键标点(图 3)(可用 放大镜工具调整图像大小)。注意:给每一片叶子标点的时候务必按照 统一的顺序,如果想调整顺序可以选择带"▼"的钢笔头形状的工具 (图 3)。当标好所有的点之后,点击输出图标(纸张形状),然后点击 "show"选项,选中所有数据(图 4)后,复制并粘贴到 Excel 中,空出最 上一行和最下一行。

👱 ImageJ	-	×
File Edit Image Process Analyze Plugins Window He	elp	

图3钢笔工具

	🛓 Lo	og					
	File	Edit	Font				
	0	1197	628	1	11	12	
	1	1321	728	1	12	13	
	2	1501	872	1	13	14	
	3	1549	1136	1	14	15	
	4	1621	1685	1	15	16	
	5	1553	2033	1	16	17	
	6	1313	2414	1	17	18	
5	7	1088	2321	1	18	19	
	8	916	1893	1	19	20	
	9	728	1385	1	20	21	
	10	824	1040	1	21	22	
	11	1004	852	1	22	23	
	12	1173	724	1	23	24	

图4 标志点标记结果

3.3 标志点坐标数据处理(Data processing for landmark coordinates)

在电子表格中,标志点的 x 坐标值和 y 坐标值分别为第 2 列和第 3 列,在 x 坐标值数列的上方一行输入文本"LM=",其后为标志点的 数目(比如:LM=13)并在 x 坐标值数列的下方一行输入文本"ID=",后 跟分类群的名字(图 5,比如: ID=LZD_T_T01_L01,LZD 代表种群, T 代表物种或分类组(当研究同一个物种时,按不同组进行分类),T01 是树的棵数,L01 叶片数)。

	А	В
1	LM=13	
2	1260	2878
3	1276	2847
4	1232	2289
5	1283	2297
6	1346	2297
7	1511	2546
8	1487	2685
9	1459	2743
10	1207	2293
11	1188	2295
12	1052	2546
13	1140	2753
14	1176	2794
15	ID=LZD_T_	T01_L01

图 5 标志点坐标数据整理

关闭数据窗口,点击带"-"钢笔头形状的工具在叶片的右键标点 清除所有标点(图 3)。向下滑动鼠标翻到下张图片,重复以上步骤直 到标完所有图像。

三、叶型分析(Morpho J 软件)

1. 功能: 基于普式叠印法去除二维或三维标志点坐标的非形态变异

后,可以进行许多重要类型的几何形态分析 (geometric morphometric analyses, GMMs),包括识别离群值,生成协方差矩阵,进行主成分分析(Principal component analyses (PCA))、典型变量分析(Canonical Variate Analysis (CVA))、偏最小二乘法(Two-Block Partial Least-Square (2B-PLS))等分析操作,同时可以对形状变化的统计结果形成可视化的图形 (Klingenberg 2011)。

2.下载地址: <u>https://morphometrics.uk/MorphoJ page.html</u>

3.操作步骤:

3.1 导入文件格式(Input file formats)

用于后续叶形分析软件 Morpho J 的数据输入格式为 txt.格式,所以 在进行分析数据之前需要将标志点的坐标数据文件转化成 Morpho J 支持的文件类型(图 5)。

3.2 导入文件(Import file)

安装并打开 Morpho J 后,在"File"中选择"Create New Dataset",出现图 6 对话框,选中"TPS"格式,找到标志点数据保存位置,最后点击 Create Dataset 导入数据。

	earrow Select the appropriate options and the data file $ imes$	Ň	
	Dimensionality of the data: 2 dimensions Object symmetry? yes		
	○ 3 dimensions		
	Name for the new dataset newDataset		
	File type: _ text _ NTSYSpc		
	查找心: □叶片数据 🗸 🔊 🗇 🗇 🔛 🗄		
	☐ 1365∱ŀŀf, xisx ☐ ALLtxt		
	S.txt		
	T.txt		
	□ 全部的叶片标记款据.xlsx		
	文件名(N): ALL.bt		
	文件类型(1): 所有文件		
	Create Dataset		
1	图 6 文件导入		
如果数据成功-	导入出现图7对话框:		
秒 MorphoJ 1.06d		-	×
File Preliminaries Vari	ation Covariation Comparison Genetics Help		
Project Tree Results	Graphics Reports		
Reading from TPS file 'Al	LL. txť.		-
Finished reading the file	ē.		
	图 7 数据导入成功提示		

如果数据导入失败,会出现图 8 对话框,具体信息如下: Problems with reading the file.

The number of dimensions in some entry/entries in record 1471 appears to be inconsistent with the number of dimensions. Errors occurred in the reading process.

The new dataset could not be created.

查找数据集出现错误的方法是用给出的行数 1471*15=22065,

(乘 15:一片叶子 13 个标志点加 LM 和 ID 共 15 行), 然后用 EXCEL 打开原始数据, 找到 22065 行, 即可找到出现错误的数据, 改正后重新导入。



图 8 数据导入失败提示

点击 Project Tree 查看导入数据,单击右键,出现图9对话框,然后选 Rename 可以给导入数据命名,这里我们命名为:All data。

🐠 MorphoJ 1.06d	_	×
File Preliminaries Variation Covariation Comparison Genetics Help		
Project Tree Results Graphics Reports		
new project		
Rename		
Remove Item		
秒 MorphoJ 1.06d	-	\times
File Preliminaries Variation Covariation Comparison Genetics Help		
Project Tree Results Graphics Reports		
new project		
All data		
		ļ
图 0 终于数据文称		
图 9 修风级据石称		

3.3 前期数据处理(Preliminary data processing)

3.3.1 普式拟合(Procrustes fit)

叶片标本上标点的位置、取向和比例等非形态的变化会直接导致 分析结果出现偏差,因此必须除去非形态变异对分析的干扰。本方法 使用 普氏 叠 印 法 去 除 非 形 状 变 异 (Morpho J 手 册, http://www.flywings.org.uk/MorphoJ_guide/frameset.htm?comparison/cv a.htm)。

步骤为:注意一定要首先在"Project Tree"中选定要处理的数据, 然后选择"Preliminaries"→"New Procrustes Fit",在弹框图 10 中选择 "Align by principal axes"后点击"Perform Procrustes Fit"。

new project - All data Procru Note: Thi Align	tes fit for All data	 on will affect the presentation			
All data	tes fit for All data	 on will affect the presentation (
Note: Thi	tes fit for All data .	n will affect the presentation			
Note: Thi	choice of orientatio	on will affect the presentation			
Note: Thi	choice of orientatio	on will affect the presentation			
O Align			of data, but not the	eresult of sta	itistical tests et
	vith first specimen ((not recommended if there is (object symmetry)		
		•			
Align	y principal axes				
🔾 Align	sing specific landm	arks			
Indicate	the landmarks or la	ndmark combinations for the	alignment of the m	iean shape,	
e.g. "L5	ог "0.3*L2 + 0.7*L7"				
Le	t point for first axis:				
Rig	t point for first axis:				
Upper p	int for second axis:		(3D data only!)		
			1 (1
		Perform Procrustes Fit		Cancel	
🐠 Mor	noJ 1.06d			-	
File Pre	minaries Variation Covar	riation Comparison Genetics Help			
Project	ree Results Graphics	Reports			
All data					
Axis 1	;2				
		1833年			
7					
		1 1 1			
		1998 - 1998 - 1998 - 1998 - 1998 - 1998 - 1998 - 1998 - 1998 - 1998 - 1998 - 1998 - 1998 - 1998 - 1998 - 1998 -			
		 EXT (CP) = 1 	and the second sec		

图 11 普式叠印结果图展示 图 11 显示所有输入数据经过平移、旋转、缩放最终最大限度的

3.3.2 查找异常值 (Find Outliers)

可在"Preliminaries"菜单中选择"Find Outliers"在图 12 弹框的左下方"Included"中显示每个个体偏离平均值的数值,此时可根据偏差大小找到离群值并选择排除或保留。



图 12 离群值筛选

3.3.3 分离对称和非对称性(Separate symmetry and asymmetry)

选择"Preliminaries"→" Select Landmarks ", 如图 13 所示:

J I	MorphoJ 1.05d								
ile	Preliminant Variation Covariation C	Comparison	Genetics Help						
Pro	New Procrustes Fit	ts							
n I	Find Outliers								
	Include or Exclude Observations								
	Generate Covariance Matrix								
	Extract New Classifier From ID String								
	Edit Classifiers								
	Edit Covariates								
	Select Landmarks								
	Subdivide Dataset By								
	Combine Datasets								
	Average Observations By								
	Link Datasets								
	Create or Edit Wireframe								
	Set Options for Shape Graphs								
出 M	3现以下对话框:						Y		
出 Ma p	9现以下对话框: prphoJ 1.06d reliminaries Variation Covariation Trace Results Granbice Reput	Compariso	on Genetics Help	-		~/		_	
	9. 现以下对话框: prphoJ 1.06d reliminaries Variation Covariation Tree Results Graphice Report 矛 New Data Set with a Subset of La Enter the name for the new data set:	Compariso orts indmarks Object	on Genetics Help symmetry?		Include	ed:		Exclu	uded
	9. 现以下对话框: prphoJ 1.06d reliminaries Variation Covariation tree Passute Carabics Paper ● New Data Set with a Subset of La Enter the name for the new data set: new data	Compariso orte indmarks Object	on Genetics Help symmetry?		Include L1	ed:		Exclu	uded
H No ojev	g 现以下对话框: reliminaries Variation Covariation t Tree Posulte Granhics Power ♥ New Data Set with a Subset of La Enter the name for the new data set: new data	Compariso orts indmarks Object © Yes	on Genetics Help symmetry? s ◯ No		Include L1 1 L2	ed:		Exclu	uded
	g 现以下对话框: reliminaries Variation Covariation t Troe Results Graphics Reputer ♥ New Data Set with a Subset of La Enter the name for the new data set: new data	Compariso orts indmarks Object © Yes •1	on Genetics Help symmetry? S () No		Include L1 - L2 L3	ed:	clude	Exclu	uded
	g 现以下对话框: reliminaries Variation Covariation t Tree Pasults Graphics Peru ▼ New Data Set with a Subset of La Enter the name for the new data set: new data	Compariso orts indmarks Object • 1 • 2	on Genetics Help symmetry?) O No		Include L1 c L2 L3 L4	ed:	Iclude	Exclu	uded
	3. 现以下对话框: reliminaries Variation Covariation tree Results Graphics Reputer ♪ New Data Set with a Subset of La Enter the name for the new data set: new data 	Compariso orts indmarks Object I Yes 1 2	on Genetics Help symmetry?) No		Includee L1 2 L3 L4 L5	ed:	cclude	Exclu	uded
	g 现以下对话框: reliminaries Variation Covariation t Tree Passults Graphics Park ♪ New Data Set with a Subset of La Enter the name for the new data set: new data 	Comparise orte indmarks Object • Yes • 1 • 2	on Genetics Help symmetry? S O No		Include L1 - L2 L3 L4 L5 L6	ed:	iclude cclude	Exclu	uded
	g 现以下对话框: reliminaries Variation Covariation tree Results Graphice Rend ♪ New Data Set with a Subset of La Enter the name for the new data set: new data _13 _12	Comparise orte indmarks Object @ Yes 1 2	on Genetics Help symmetry? S O No •8 •7		Include L1 - L2 - L3 - L4 - L5 - L6 - L7 - L9 -		aclude cclude	Exclu	uded
	g 现以下对话框: reliminaries Variation Covariation tree Results Graphice Rend ▼ New Data Set with a Subset of La Enter the name for the new data set: new data _13 _12	Comparise orte indmarks Object • Yes •1 •2	on Genetics Help symmetry? S No •8 •7		Include L1 - L2 - L3 - L4 - L5 - L6 - L7 - L8 - L9 -		actude cctude	Exclu	uded
	g 现以下对话框: reliminaries Variation Covariation tree Results Graphice Rend Prove Data Set with a Subset of La Enter the name for the new data set: new data 13 12	Comparise orte indmarks Object • Yes •1 •2	on Genetics Help symmetry? s O No •8 •7		Include L1 L2 L3 L4 L5 L7 L7 L9 L10	d:	aclude cclude plicate	Exclu	uded
	9. 现以下对话框: reliminaries Variation Covariation tree Results Graphice Rend Prevention Covariation New Data Set with a Subset of La Enter the name for the new data set: new data 13 12	Comparise orte indmarks Object • Yes •1 •2	on Genetics Help symmetry? S O No •8 •7		Include L1 2 L3 L4 L5 L6 L7 L8 L9 L10 L11	d:	iclude kclude plicate	Exclu	uded
	g 现以下对话框: reliminaries Variation Covariation troe Results Graphice Read Prevention Covariation New Data Set with a Subset of La Enter the name for the new data set: new data 13 12	Comparise orte indmarks Object • Yes •1 •2	on Genetics Help symmetry? ONO 8 8 7		Include L1 2 L2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12	d:	iclude kclude plicate	Exclu	uded
	g现以下对话框: reliminaries Variation Covariation troe Results Graphics Read Prevention Covariation New Data Set with a Subset of La Enter the name for the new data set: new data 13 12 11	Comparise orte indmarks Object • Yes • 1 • 2	on Genetics Help symmetry? ONO 8 7		Include L1 2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12		nclude cclude plicate	Excl	uded
	g.现以下对话框: reliminaries Variation Covariation troe Results Granbics Read Prevention Covariation New Data Set with a Subset of La Enter the name for the new data set: new data 13 12 11	Comparise orte indmarks Object • Yes •1 •2	on Genetics Help symmetry? No •8 •7 •6		Include L1 2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12 Mo	d:	nclude cclude plicate	Excl	uded
	g 现以下对话框: porphoJ 1.06d reliminaries Variation Covariation to Troo Passutts Graphics Peore ♥ New Data Set with a Subset of La Enter the name for the new data set: new data 	Comparise orte indmarks Object • Yes •1 •2	on Genetics Help symmetry? No 8 7 6		Include L1 1 L2 L3 L4 L5 L6 L7 L8 L9 L10 L11 122 Mo	e Dourn	nclude cclude plicate	Excl	uded
	g 现以下对话框: porphoJ 1.06d reliminaries Variation Covariation to Troo Results Graphics Repute I New Data Set with a Subset of La Enter the name for the new data set: new data 13 12 11	Comparise orte indmarks Object • Yes •1 •2	on Genetics Help symmetry? No 8 7 6		Include L1 2 L3 L4 L5 L6 L7 L8 L9 L10 L11 122 Mo Move	ed:	nclude cclude plicate	Excl	uded
	g 现以下对话框: porphoJ 1.06d reliminaries Variation Covariation to Troo Passute Graphice Reputer P New Data Set with a Subset of La Enter the name for the new data set: new data 13 12 11	Comparise orts indmarks Object I 2	on Genetics Help symmetry? No 8 7 6		Include L1 L2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12 Mo Move	e Down	nclude cclude plicate	Exclu	uded
	g 现以下对话框: porphoJ 1.06d reliminaries Variation Covariation to Troo Passute Graphice Reputer P New Data Set with a Subset of La Enter the name for the new data set: new data 13 12 11	Comparise orte indmarks Object • Yes • 1 • 2	on Genetics Help symmetry? No 8 7 6		Include L1 - L2 - L3 - L4 - L5 - L6 - L7 - L8 - L9 - L10 - L11 - Move	d: In E) Du ve Up e Down	nclude cclude plicate	Exclu	uded
上 P Ojev ■	g.现以下对话框: porphoJ 1.06d reliminaries Variation Covariation Troe Results Graphics Reputer P New Data Set with a Subset of La Enter the name for the new data set: new data 13 12 11 10.9	Compariso orte indmarks Object I 2	on Genetics Help symmetry? ONo •8 •7 •6		Include L1 - L2 - L3 - L4 - L5 - L6 - L7 - L8 - L9 - L10 - L11 - Mov	ed:	clude cclude plicate	Exclu	uded ept cel

图 14 标志点筛选以及新数据命名

给新数据集命名,在图 14 中,选择 Yes(观察对称非对称结构),标志点筛选结束后选 Move Down,最后点击 Accept,出现图 15 对话框:

梦 Procrustes fit for A		×			
Note: This choice of orientation	n will affect the presentation	of data, but not the result of statistical tests etc.			
○ Align with first specimen (not recommended if there is	object symmetry)			
Align by principal axes					
Align using specific landma	arks				
Indicate the landmarks or lan e.g. "L5" or "0.3*L2 + 0.7*L7"	dmark combinations for the	alignment of the mean shape,			
Left point for first axis:					
Right point for first axis:					
Upper point for second axis:		(3D data only!)			
	Perform Procrustes Fit	Cancel			
图 15 区分对称和非对称结构的普式叠印					

点击 Perform Procrustes Fit, 弹出图 16 对话框:

秒 MorphoJ 1.06d			- 🗆
File Preliminaries Variation	Covariation Comparison Genetics Help		
Project Tree Results Gra	phics Reports		
Find outliers All data new	data Choose paired and median landmarks .		
	•1 •2 •8 •13 •7 •12		Landmark Lmk 1 and Lmk 1 Pair
	° 6	_ •11	
	<u>•5 •4 •9</u> •10 •3		Accept

图 16 两边成对的标志点和中轴上的标志点选择 可以看出浅蓝色线条相连的为对称的标志点,该软件可以自动识别对 称结构,点击 Accept,出现以下对话框,对话框中可以点击 Symm1vs2 和 Asym1vs2,查看对称(图 17)和非对称(图 18)结构的普氏叠印图:



图 18 非对称结构的普式叠印图

3.3.4 提取新分类层(Extract New Classifier)

如果需要单独分析每个种群,就要对数据库中的信息进行提取并

重新分类。选择"Preliminaries"→"Extract New Classifier from ID String"。弹出图 19 对话框,在"Name for new classifier"框中给新的分 类命名,这里命名"leaf": 根据"Identifier strings"中的字符所显示叶片 水平有 13 个字符长,从第 1 个第 13 个字符,因此"First character"为 1,"Last character"为 13,点击"Execute"。同样的方法可以命名为"tree": 根据"Identifier strings"中的字符所显示树水平有 9 个字符长,从第 1 个字符到第 9 个字符,因此"First character"为 1,"Last character"为 9。

梦 MorphoJ 1.06d		
File Preliminaries Variation C	Covariation Comparison Genet	tics Help
Project Tree Results Grap	hics Reports	
 ¬ new project ↑ ■ All data └─ ■ new data 	👽 Extract new classifier from	ID strings X
	Dataset: new data	
	Existing classifiers:	Identifier strings:
	Name for new classifier:	LZD_M_T01_L01
	First character: Last ch 1 13 (Positive - count from left; ne	egative - count from right) Cancel Execute

图 19 提取新的分类器

当提取新的分类后可以在"Preliminaries"下选择"Edit Classifier" 检查并编辑结果(图 20)。

Project Tree	Res	sults Gra	phics R	eports	1						
Find outliers	All	data nev	v data E	dit class	ifiers in 'I	new data'					
Identifier		leaf	tree	(aroup	population			 	 	
ZD M T15 L0	5 1	ZD M T1	LZD M T	15 M		LZD	-				
ZD M T16 L0	1 1	ZD M T1	LZD M T	16 M		LZD	-				
ZD M T16 L0	2 1	ZD M T1	LZD M T	16 M		LZD					
ZD M T16 L0	3 I	ZD M T1	LZD M T	16 M		LZD					
ZD_M_T16_L0	4 l	ZD_M_T1	LZD_M_T	16 M		LZD					
ZD_M_T16_L0	5 I	ZD_M_T1	LZD_M_T	16 M		LZD					
BX_T_T09_L0	1 (GBX_T_T0	GBX_T_T	09 T	(GBX	1				
BX_T_T09_L0	3 (GBX_T_T0	GBX_T_T	09 T	(GBX	1				
BX_T_T09_L0	4 (GBX_T_T0	GBX_T_T	09 T	(GBX					
BX_T_T09_L0	5 (GBX_T_T0	GBX_T_T	09 T	(GBX					
BX_T_T09_L0	6 (GBX_T_T0	GBX_T_T	09 T	(GBX					
BX_T_T11_L0	1 (GBX_T_T1	GBX_T_T	11 T	(GBX					
BX_T_T11_L0	2 (GBX_T_T1	GBX_T_T	11 T	(GBX					
BX_T_T11_L0	3 (GBX_T_T1	GBX_T_T	11 T	(GBX					
BX_T_T11_L0	4 (GBX_T_T1	GBX_T_T	11 T		GBX					
BX_T_T11_L0	5 (GBX_T_T1	GBX_T_T	11 T		GBX					
BX_T_T14_L0	1 (GBX_T_T1	GBX_T_T	14 T	-	GBX					
BX_T_T14_L0	2 (GBX_T_T1	GBX_T_T	14 T		GBX					
BX_T_T14_L0	3 (GBX_T_T1	GBX_T_T	14 T		GBX					
BX_T_T14_L0	4 (GBX_T_T1	GBX_T_T	14 T		GBX					
BX_T_T14_L0	6 (GBX_T_T1	GBX_T_T	14 T		GBX	_				
BX T T15 L0	1_(GBX T T1	GBX T T	15 T	1	GBX					-
leaf			Add	classifie	er		Rename cla	ssifier			
tree											
aroup			Delete	classifie	er(s)		Replace val	lies			
population							,				
			Combin	o oloo - ii	fioro				Conocl	Accont	1
			Combin	e classi	ners				Cancer	Ассері	

图 20 编辑分类器

3.3.5 生成新的子样本集(Generate new subsamples)

根据分组情况将叶片标志点的数据做平均,首先选择 "Preliminaries→Average Obsevations By..."(图 21), 然后弹出图 22 对 话框,选择"Average by..."确定要数据平均的水平,最后点击 Execute.

J	MorphoJ 1.06d	-	×
ile	Preliminaries Variation Covariation Comparison Genetics Help		
Pro	New Procrustes Fit ts		
l n	Find Outliers		
- 6	Include or Exclude Observations		
	Generate Covariance Matrix		
	Extract New Classifier From ID String		
	Edit Classifiers		
	Edit Covariates		
	Select Landmarks		
	Subdivide Dataset By		
	Combine Datasets		
	Average Observations By		
	Link Datasets		
	Create or Edit Wireframe		
	Set Options for Shape Graphs		

图 21 平均数据到不同的分组水平

秒 MorphoJ 1.06d

Project Tree	Results	Graphics	Reports				
new project All data	Jata						
		7 A	verage obs	servations l	oy		×
		Data	iset: new da	ita			
		Nam	ie for the ne	w dataset:			
		new	[,] data, avera	ged			
		Ave	age by			Data types:	
		leaf			•	Centroid size Symmetric component	
						Asymmetry component	
						Cancel Execute	

图 22 选择平均数据的分组水平

返回"Project Tree"中将新的数据集改名,如图 23 所示:

MorphoJ 1.0	6d	
Project Tree	Results Graphics Reports	
 new project 	ta r data, averaged	
	Rename item 2 Please enter the new name for the item currently named "new data, averaged". leaf 确定 取消	× 1
3.6 构建线相	图 23 修改数据名称 重(Create wireframes)	

线框是标志点之间的连线的集合,可帮助研究者观测标志点的结 构配置。使用 Morpho J 软件创建线框图的步骤如下:选择 "Preliminaries" → "create or edit wireframe", 弹出图 24 对话框, 在出 现的图形背景中右击鼠标可以调整图形的方向。编辑线框,即在标志 点之间进行划线,有两种方法:一是可以点击图中的一个标志点,然 后拖拽到另一个标志点上;二是在右边下滑菜单中选择要连线的两个 点然后点击"Link landmarks"。点击"Delete link"可删除线段。



3.4 多变量分析(Multivariate analysis)

3.4.1 主成分分析(Principal component analyses)

PCA 分析是将多个指标化为少数几个综合指标的一种多元统计分析方法。由于这种方法既可消除各指标不同量纲的影响,也可消除由各指标间相关性所带来的信息重叠,从而起到降维的作用,简化了指标的结构,使分析问题简单、直观、有效(杨维权等,1989)。主成分分析作为多变量分析中使用最广泛的方法之一,可用于检测数据集中形状变化的主要特征并可视化分析组之间的形状差异(Vujić et al., 2016)

步骤:进行 PCA 之前首先要创建叠印后的叶形态坐标的协方差 矩阵(covariance matrix)。点击"Preliminaries" → "Generate Covariance Matrix"; 在 "data types" 框 中 选 择 "Symmetric component" 或 者 "Asymmetry component"后点击"Execute"按钮。出现图 25 对话框:

💀 Morphol 1 06d	_	П	×
File Dreliminaries Variation Covariation Comparison Genetics Help		_	~
Project Tree Results Graphics Reports			
new project All data P I rew data P CovMatrix, leaf, Symmetric component F CovMatrix, leaf, Asymmetry component F covpatrix, leaf, Asymmetry component F group group group group			

图 25 创建协方差矩阵

然后选中分析数据的协方差矩阵进行 PCA 分析以显示不同的叶 片在几何空间上的异同。点击"Variation" \rightarrow "Principal Components Anaysis",出现图 26:

秒 MorphoJ 1.06d	– 🗆 X
File Preliminaries Variation Covariation Comparison Genetics H	elp
Project Tree Results Graphics Reports	
outliers All data new data Edit classifiers in 'new data' Creat	e wireframe PCA: CovMatrix, leaf, Symmetric component
PC shape changes Eigenvalues PC scores	
<mark>√</mark> 11	
	12
	13
~1 0	
•*	
3	2 1
4	
<⁵	
	* ⁸
	P ⁷
- 6	
N C1	
PUT	
下 D O C D C 1 4	しなっトルホル
图 26 PC1 轴	上的叶形变化

在出现的"graphics"窗口上点击"PC scores", 然后出现散点图(图 27):





Print ...

图 28 散点图的设置选项

"Label Data Points",可对所有标志点名字显示出来; "Color the Data Points" 选中"Use a classifier variable to determine the color for each category", 然后弹出图 29, 选择物种 spe 作为分类单元,



图 29 设置 PCA 散点图的分组水平

选中颜色,点击物种名可以人为更换颜色,如图 30 所示:



图 30 设置各组散点颜色

点击"Use Color"最后点"OK",即可出现不同颜色的散点图,如图 31 所示:



图 31 分组完成的 PCA 散点图 此外,单击右键出现图 32 对话框中阴影部分:



图 32 PCA 散点图的设置选项

弹出图 33, 选中"Draw ellipse(s)", 点击 OK, 图 34 是设置完成的 PCA 散点图

🛷 Choose set	ttings for the co	nfidence ellips	es		×
✓ Draw ellips	se(s)				
Type: 🔘 Eq	qual frequency e	llipse(s) 🔾	Confidence e	ellipse(s) for	mean(s)
Probability:	0.9		(Must be	between 0.0) and 1.0.)
🔽 Use a clas	sifier as a criter	ion for grouping	observation	s	
spe		ion for grouping	goborration		-
Use this cl	assifier to deter	mine the colors	s of ellipses a	nd data point	ts
(Caution: t	this may override	e earlier choice	s of colors.)		
Clip the elli	ipse(s) at the ma	argins of the gra	aph		
C Chautha a	lata painta				
Show the d	tata points		C	ancel	ОК
	图 33 P	CA 散点图	的置信区	间设置	
步 MorphoJ 1.06d					- 🗆 ×
File Preliminaries Van	riation Covariation Co	mparison Genetics	Help		
outliers All data ne	ew data Edit classifier	s rs in 'new data' Crea	te wireframe PC	A: CovMatrix, leaf, S	ymmetric component
PC shape changes	Eigenvalues PC score	es			
					м
	0.30				S T
	0.20		•	•	
	ut 7	1.1	1 . ¹		
	0.10 -				
	ă E				
	<u> </u>				
	-0.10				
	² rino				
	-0.20 -	•••••	••		
	-0.30		, ,	-, 1	
	-0.30 -	0.20 -0.10 -0	.00 0.10	0.20 0.30 1	
			Sinponent	1	

图 34 设置完成的 PCA 散点图 增加散点的大小,单击右键出现图 35 对话框中阴影部分:

	Choose Principal component for the Horizontal Axis						
	Choose Principal component for the Vertical Axis						
	☑ Use Same Scaling for Both Axes						
	Use Same Scaling for ALL Axes						
	Color the Data Points						
	Resize Data Points						
	🗆 Label Data Points						
	Confidence Ellipses						
	Print						
	Export Graph to File	The					
	Close Panel	YX/					
在图 36 中,	图 35 PCA 散点图的设置选项, 输入1到10可以改变散点大小, Resize data points ×						
	Diameter of points, in screen pixels						
	5.0						
	确定 取消						
	图 36 修改散点图中点的大小						
然后选择图	37 中的"Export Graph to File",将设置完	成的散点图					
导出							
	Choose Principal component for the Horizontal Axis						
	Choose Principal component for the Vertical Axis						
	☑ Use Same Scaling for Both Axes						
	Use Same Scaling for ALL Axes						
X	Color the Data Points						
	Resize Data Points						
	🗆 Label Data Points						
	Confidence Ellipses						
-	Print						

Export Graph to File ...

Close Panel

图 37 导出散点图选项

其中,图 38 展示了所有可导出的文件格式:



图 38 导出图的文件类型

另外还可以通过点击 PC shape changes, 查看不同 PC 轴上叶片形态的变化,图 39 是 PC1 轴上叶形变化的棒棒糖图:



图 39 PC1 轴上叶形变化的棒棒糖图

选中图像,右击鼠标,改变 PC 轴,设置 PC 值的大小,改变图 形类型等,以修改图形类型为例:

选择"Change the type of Graph"→"Wireframe Graph"(图 40),弹出





图 41 PC1 轴上叶形变化的线框图



我们可以通过右击鼠标点击 Export to File 导出线框图与网格图:

Clear Window Export to File..., 然后在作图软件中叠加起来(修图软件 Ai), 例如图 43, 更形象的表示叶片形态的变化。



图 43 最终合成的重组图

除了散点图和叶型轮廓图,在"result"中还显示了分析结果(图 44),右 击可选择输出结果,显示各个 PC 占总变异的比例:

步 N	NorphoJ 1.06d						-	
File	Preliminaries	Variation Covari	ation Compariso	n Genetics	Help			
Ргој	ect Tree Res	ults Graphics	Reports					
Princi	ipal Component A	nalysis: PCA: Co	vMatrix, leaf, Sym	metric comp	onent			
	Eigenvalues	% Variance	Cumulative %					
1.	0.00667139	42.591	42.591					
2.	0.00321160	20.503	63.094					
3.	0.00243967	15.575	78.669					
4.	0.00152800	9.755	88.424					
5.	0.00057340	3.661	92.084					
6.	0.00040910	2.612	94.696					
7.	0.00033209	2.120	96.816					
8.	0.00016292	1.040	97.856					
9.	0.00013167	0.841	98.697					
10.	0.00011857	0.757	99.454					
11.	0.00008557	0.546	100.000					
Total Varian Eigeny	variance: 0.01 nce of the eigen	566397 values: 0.00000 caled by total v	37704407 ariance: 0.01537					
Eigen	value variance s	caled by total v	ariance and number	of variable	es: 0.18594			
Note:	throughout all	calculations of	eigenvalue variand	es, the dim	ensionality used	was 11.		
Please	e double-check b	ecause this dime	nsionality may not	be appropri	iate for all situ	ations.		
	Principal Com	ponent Coefficie	nts					
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
		0.005505	图	44 PC	A结果显	.示	0.00010	0.0010

3.4.2 典型变量分析(Canonical variate analysis)

CVA 分析是用于检测具相同或相似特征的某一类样品总体之间 的差异的一种多元统计方法。不仅可以消除数据变量之间的相关性, 而且可以解决变量自身相关性(胡坤,1992)。CVA 产生的是散点图, 输出的散点图可以通过 Morpho J 添加相同频率椭圆(Equal frequency ellipses)或者椭圆置信区间(Confidence ellipses for means)。

操作步骤:选择"Comparison"→"Canonical Variate Analysis",在 出现的对话框中的"Classifier variables"选择之前提取的新的分类单元 group(组)用于分类, "Classifier variable(s) to use for grouping"选择 "group"作为分类单元,并且勾选执行 10000 次(图 45)。

秒 MorphoJ 1.06d	
File Preliminaries Variation Covariation	n Comparison Genetics Help
Project Tree Results Graphics	lanarta
 new project All data Image: Second strain str	 Canonical Variate Analysis × Name: CVA Dataset: tree Data type: new data, symmetric component Classifier variable(s) to use for grouping: tree group population
	Permutation test for pairwise distances Number of iterations: 10000 Cancel Execute

图 45 CVA 数据设置

执行后会出现叶型轮廓图(图 46):



之后和 PCA 一样单击右键对图片进行颜色、散点大小、椭圆频

率的制作及修改,同时也要通过 CV shape changes 将 CV 轴上叶片形态展现出来。

最后在"Results"菜单中可以显示 CVA 分析结果。

3.4.3 判别分析(Discriminant analysis)

DA 分析通过建立一个或多个判别函数,用多元数据确定判别函数中的待定系数,并计算判别指标(陈德钊,1998)。据此即可确定某一新样品属于哪一类。在形态学研究中一般用于分析两个类群(Viscosi and Cardini, 2011; Vieira *et al.*, 2014)间的变异。DA 分析通过十字交叉 函数(Cross-validated scores)和判别函数(Discriminant scores)两种方法 分别进行分析(T^2 *P*-value for tests with 1000 permutations < 0.0001),提供可靠的种群分类信息,输出的结果为条形图。

步骤: DA 分析不区分对称和非对称组分,首先在"Project Tree" 选中原始数据(All data)重做"select landmarks"(图 14),选择 No(不观 察对称和非对称结构),得到新的数据集(命名为"no data"),新数据同样需要进行分类及做平均,然后进行 DA 分析,在菜单栏点击"Comparision"选择"Discriminant Fanction"(图 48)。



图 48 DA 选项

选择好数据集,数据类型,数据分类水平以及成对比较的组,重 复执行1000次,点击"Execute"(图 49)。

Project Tree	Results Graphics Reports		
new project All data	🥩 Discriminant function analysis		×
	Name for the discriminat function analysis: Discriminant	function	
4	Dataset: no data, averaged tree		
	Data type: no data, Procrustes coordinates 🔻		
P-los	Classifier(s) to be used as grouping criterion Pairs	s of groups to be included:	
• € 	Classifier(s) to be used as grouping criterion Pairs tree M	s of groups to be included: S	
-	Classifier(s) to be used as grouping criterion Pairs tree M group M	s of groups to be included: S T	
-	Classifier(s) to be used as grouping criterion Pairs tree group S - 1	s of groups to be included: S T T	
- = gr = nc - = nc	Classifier(s) to be used as grouping criterion Pairs tree M group S 1	s of groups to be included: S T T	
P I P I No I I I I I I I I I I I I I I I I I I I	Classifier(s) to be used as grouping criterion Pairs tree group M S 1	s of groups to be included: S T T	
- - - - - - - - - -	Classifier(s) to be used as grouping criterion Pairs tree M group S 1	s of groups to be included: S T T	
- = = n - = = n - = = n - = = n - = = n	Classifier(s) to be used as grouping criterion Pairs tree group S - 1 S - 1 Include all pairs of groups	s of groups to be included: S T T	
- = gr - = no - = no - = no - = no	Classifier(s) to be used as grouping criterion Pairs tree M group S 1 ✓ Include all pairs of groups	s of groups to be included: S T T	

图 49 DA 数据设置

结果显示如图 50, 可以根据 DA 轴的值改变叶片形态。

	🌮 MorphoJ 1.06d	-							
	File Preliminaries Variation Covariation Comparison Genetics Help								
Project Tree Results Graphics Reports									
data' Y PCA: CovMatrix, leaf, Symmetric component / Create wireframe / CVA / CVA / no data / Discriminant fu									
	M-S M-T S-T								
	Shane difference Discriminant scores Cross validation scores								
	•1								
	•2	•2							
	● ¹ 3 ● ⁸								
	12 7								
	•								
	-11 -6								
	●1♥ ⁹ ●3 ● ⁴ ● ⁵								
	S1								

图 50 DA 的棒棒糖图

点击"Cross-validation scores"结果如图 51 所示, 再单击右键, 改变条
形图颜色, 菜单栏第四行, M-S, M-T, S-T 分别为不同组间两两比较的结果:



3.4.4 偏最小二乘法(Two-Block Partial Least-Square analysis, 2B-PLS)

2B-PLS 分析是几何形态测量学中用于测试两个定量数据源之间 的协方差的多变量测量工具(Marugán-Lobón, 2010),该方法可用来分 析植物的异速生长。异速生长指的是因为大小的变化从而影响形态发 生的变化(Klingenberg and McIntyre, 1998; Viscosi 2015)。叶片的大小 和形状是可以相互作用相互影响的,从统计学上可以通过测量叶片大 小比例的变化衡量异速生长。每一个物种叶片形状的对称和非对称结 构均可通过测量实验证明其在统计学上是否为的显著的形态学变异。 通过回归分析 2B-PLS 将形态性状映射到大小的性状上来确认不同的 形态变异是否由叶片大小的变化引起,计算结果将给出 P-value 值以 及 RV 值来确定结果是否显著。

步骤: 该分析同样可以通过 Morpho J 软件计算, 在菜单栏选中 "Covariation"然后选择"Partial Least Squares"再选择"Two Sequrate Blocks"(图 52)。

Preliminaries Variation	Covariation Comparison Genetics Help	
oject Tree Results Gr	aphics Reports	
new project	₱ PLS analysis: two separate blocks	E I
newDataset		
🗣 🗐 sym	Name for the analysis: PLS	
 sym, averaged lea 		
- Sym, averaged tre	Block 1	Block 2
PCA: Covil	(Please choose Block 1 before Block 2.)	
- 🗐 PC sco	Detenster	
- 😽 Modularity	h Datasets:	
- 😽 Modularity	no 📤	sym, averaged leaf
- Modularity	sym, averaged leaf	
- Modularity	eum averaned tree	
	Data matrices:	·
PCA: Covil		aum controld also
🗌 🗐 PC sco	sym, centroid size	sym, centrold size
•- 🏡 CVA	sym, symmetric component	sym, symmetric component
🖵 🗐 CVA, sco	sym, asymmetry component	sym, asymmetry component
PLS	Variables	
- PLS score	s valiables.	
PLS score	Centroid Size	Symmetric component
- PLS	Log Centroid Size	
E PLS score	s	
PLS		
- E PLS score	s	
 sym, averaged sp sym wireframe 	e	
- I sym averaged po	Perform permutation test Number	r of rounds: 10000
- R PLS		
E PLS scores	Pooled analysis within subgroups	
• 🗐 no	loaf	
👇 🗐 no, averaged leaf	free	
P Discriminant f	spe	
- I no averaged tree	3 000	Cancel Execute
no, averaged area		Cancer

图 52 2B-PLS 数据设置

Block1 和 Block2 的数据类型、矩阵、变量根据自己的问题来设定, 本示例中我们以个体水平为单位,选择叶片的对称组分作为 Block1, 选择叶片的大小作为 Block2,进行回归分析,以物种来分组,通过 10 000 次迭代来检验其统计学意义。

结果如图 53 所示,可以单击右键,



图 54 2B-PLS 分组后的散点图

图 54 是对称结构与叶片大小的回归分析结果,之后再做非对称 结构与叶片大小的回归分析,步骤与以上相似,只有选择数据矩阵时 选择非对称结构。

3.4.5 普式方差分析(Procrustes ANOVA)

Procrustes ANOVA 是一种量化不同水平的相对变异量的方法,在 叶片几何分析中用于误差测量和测试不对称变异的相对量 (Klingenberg and McIntyre, 1998; Klingenberg *et al.*, 2002; Klingenberg, 2003; Viscosi and Cardini, 2011)。它主要使用基于 Procrustes 距离的方 法来评估影响的相对大小,并使用 MANOVA 方法进行统计推断 (Morpho J 手册)。

	(1) 八托雨西巴) 并的数据(人夕为		
ANO	VA 分析需要于八制的数据(钾石为	LM=13	
"ANOVA	lata"),每个叶片需要重复标记两次,	1219	2134
数据格式	如右图, 每个叶片命名(ID=种群名)	1229	2105
the sta	卅 叶 香 台) 例	1208	1578
177 17	树_「_里友」,例	1244	1585
ID=MXO_	_S_T01_L01_1,表示种群 MXO 中 S	1265	1573
物种的第一	一棵树的第一片叶子第一次标记的	1415	1808
叶形粉 捉((卡士占从行)	1423	1904
「10致1石(你心思主你)。	1371	1998
🐠 MorphoJ 1.06d		1156	1600
File Preliminaries	Variation Covariation Comparison Genetics Help	1120	1588
Project Tree Re	Principal Component Analysis	1037	1814
new project	PC Scores From Other PCA	1065	1929
P→ ■ new data	Matrix Correlation	1081	2024
P-	Convert Covariance Matrix to Landmark Coordinates	ID=MXO_S	_T01_L01_1
<u>م</u>	Procrustes ANOVA	LM=13	
- 🔂 Cov	Matrix, leaf, Asymmetry component	1221	2135
- 🛆 leaf	i wireframe	1232	2102
	PLS scores	1210	1576
P ≣ tree P 80 CVA	\	1246	1587
	CVA, scores	1267	1571
P no data		1415	1807
P ■ no data	i, averaged leaf i, averaged tree	1423	1903
r 🔂 Dis	criminant function Scores: Discriminant function	1374	1996
- 🗐 no data	i, averaged group	1155	1600
ANOVA data		1121	1589
Į	图 55 ANOVA 洗项	1035	1815
开场 注制	HEN (SET O IN DI IN)	1067	1929
元 付 该 致 打	据于八("File→Create New Dataset"),	1081	2024
然后对娄	女据进行预处理,即普氏叠印	ID=MXO_S	T01_L01_2

(Preliminaries" \rightarrow "New Procrustes Fit);

检验离群值("Preliminaries" → "Find Outliers"); 分离对称和不对称("Preliminaries" → "Select Landmarks"); 数据分类("Preliminaries" → "Extract New Classifier from ID String); 将数据做平均("Preliminaries" → "Average observations by")。 开始 ANOVA 分析, 点击("Variation" → "Procrustes ANOVA")(图 55), 出现图 56 对话框, Individual(一般是指你的样本)选叶片水平, Side 是 系统指定的, Error 有两个, 一个是检验图像扫描误差, 另一个是检验 标记点标记的误差, 由于我们每片叶子只扫描了一次, 重复标记了两 次, 所以这只能检验标记误差, Additional main effects 就选剩下的几 个分类水平(物种, 种群, 树), 然后点击 Execute。

🐠 Procrust	es ANOVA	
Name: Dataset:	Procrustes ANOVA ANOVA	
Classifiers	for the effects:	Additional main effects:
Individual:	LEF 🗨	none
		CLADE
Side:	none 🔻	РОР
Error 1:	REP	
Error 2:	none 🔻	
		Cancel Execute

图 56 ANOVA 数据设置

ANOVA 结果如图 57, 叶片大小和形状变异被划分到物种, 种群, 树, 叶和误差几个水平。另外形状变异还被分解到定向不对称(Side)和波 动不对称(Ind*Side),分别表示叶片左右两侧之间的系统差异和微小 的随机偏差。最后各个水平对叶大小和形态的影响大小是根据 F 比率 和每种效应解释的方差百分比(将效果的平方和(SS)除以总 SS 并将该 比率乘以100)来衡量的。

秒 MorphoJ 1.06d

File Preli	minaries Variat	tion Covariation	Compar	ison Gene	tics Help
Project T	ree Results	Graphics Rep	orts		
inis anaiy	SIS CARES INCO A	ccount the object	symmetry	in the us	τα.
Classifier	s used for the P	rocrustes ANOVA:			
Extra main	effect(s):				
— Extra 1	: CLADE				
— Extra 2	: POP				
— Extra 3	: IND				
Individual	s: LEF				
Error 1: R	EP				
Centroid s	ize:				
Effect	SS	MS	df	F	P (param.)
Extra 1	1816138.286326	908069.143163	2	31.64	<. 0001
Extra 2	4719649.051126	168558.894683	28	5.87	<. 0001
Extra 3	2150643.436780	37730.586610	57	1.31	0.0870
Individual	5796534. 699348	28695.716333	202	7724.51	<. 0001
Error 1	1077.318013	3.714890	290		
Shape, Pro	crustes ANOVA:				
Effect	SS	MS	df	F	P (param.)
Extra 1	0.30989572	0.0140861691	22	7.45	<. 0001
Extra 2	2.77835461	0.0090206319	308	4.77	<. 0001
Extra 3	1.40306361	0.0022377410	627	1.18	0.0036
Individual	4.20229310	0.0018912210	2222	2.21	<. 0001
Side	0.00901947	0.0008199520	11	0.96	0.4837
Ind * Side	2.72297172	0.0008565498	3179	183.73	<. 0001
Error 1	0 02974303	0.0000046619	6380		

图 57 ANOVA 结果呈现

3.5 导出数据(Export data)

所有的分析都展示在 Project Tree 上(图 58),叶片形态数据可以 与环境数据或者遗传数据做相关性分析,所以我们需要获得叶片形态 数据,Morpho J软件可以将 PCA 分析之后的叶型数据导出。这部分 相关分析例如 RDA 未包括进来

步骤: PCA 分析后,如图选中 PCA 数据 SolvMatrix, leaf, Symmetric component M PCA: CovMatrix, leaf, Symmetric component □ PC scores, CovMatrix, leaf, Symmetric component, 选择"File"后"Export Dataset"(图

59)。

ſ



图 59 导出数据

注: 导出数据时, 若准备导出的数据之前进行了分层, 则可以导出不同的分层数据(图 60):

)ata types:		Classifi	ers:	
PC scores		leaf		
		ind		
		spe		
Include value	riable names as first lii	ie of the file		
保存: 📑 D	ocuments			
📑 Autodesk	Application Manager		EGA7	
📑 Download	s		My eBooks	
📑 Fax			My Music	
📑 Inventor §	Server SDK ACAD 2016		My Pictures	
•			•	11
文件名(<u>N</u>):	PC scores, CovMatrix,	A-D-H, Procrustes coordina	ates.txt	K 47.
☆仕米刑(工)・	所有文件			$\Gamma \propto$
XII X±\1/	рнал			
			保存 取消	

四、检测叶形差异是否显著

我们从 Morpho J 中导出了对称组分和非对称组分的 PCA 数据以 及叶片大小数据, 并利用 R 语言的 agricolae 包进行了多重比较分 析, 来检测不同物种间叶形对称组分, 非对称组分以及叶片大小变 异是否显著。运行代码如下: >library(agricolae) >data=read.table("clipboard",header=T,sep="\t") #导入数据 >model = aov(syPC1 ~ species, data) >summary(model) >out =scheffe.test (model,"species") >out #查看结果

五、地理、遗传和气候对叶形的影响

1.获取叶形、地理、遗传及气候因子

我们首先从 Morpho J 中导出了对称组分、非对称组分和叶片大 小的数据,作为叶形数据。微卫星分子(Simple Sequence Repeats, nSSRs)标记的 PCA 分析的前三个主成分(PCs)作为中性遗传变量。 为了获取气候变量,我们在 WorldClim 数据库 (http://www.worldclim.org)中获取了1970-2000近30年的气候数据, 共103个气候因子。并使用 ArcMAP(Environmental Systems Research Institute, 2014)将采样点的 GPS 坐标(经度和纬度)映射到 30 arcsecond 的 ESRI®气候数据网格中,获得19个采样点的气候因子。接 着,我们通过 R 语言的"usdm" 函数包(Naimi *et al.*, 2014)中的方差膨 胀因子分析(Variance Inflation Factor, VIF),从这19个气候变量中去 除部分相关性较强的因子,最终保留 3-5个气候因子。保留的气候因 子、SSR 分子标记的前三个 PCs 和经纬度及海拔数据作为叶形的预 测因子,被用于后续的冗余分析。

2.叶形与地理、遗传和气候的冗余分析

为了调查地理、遗传和气候因子是否解释了一些叶形变异(对称 组分、非对称组分和叶片大小),我们使用 R 语言的函数包 "usdm"(Oksanen et al., 2017)进行一系列基于距离的冗余分析 (distance-based redundancy analyses, dbRDAs),包括边际测试(Marginal tests)和条件测试(Conditional tests) (Legendre and Anderson, 1999)。我 们使用方差划分可以运行一个包括地理、遗传和气候解释变量的边际 模型(地理+遗传+气候)。另外也可以通过控制另外两种因素运行三个 不同的条件 dbRDAs 模型,分别用来测试纯的地理效应、纯的遗传效 应和纯的气候效应: (1)控制遗传和气候效应的有关地理变量的条件 模型(地理+Condition(遗传+气候)); (2)控制地理和气候效应的有关遗 传变量的条件模型(遗传+Condition(地理+气候)); (3)控制地理和遗传 效应的有关气候变量的条件模型(气候+Condition(地理+遗传))。最后, 通过方差分析函数(anova function)计算 dbRDA 的 F 值、解释变异的 百分比和 p 值来评估每个独立因素对叶形变异的影响(Oksanen et al., 2017)。

运行代码如下:

1. 气候, 地理和遗传对叶形的影响

###导入数据

##获取遗传数据

>library(vegan)

>data=read.table ("clipboard", header=F) #导入 SSR 归一化数据

```
> p=prcomp (data, retx = TRUE, center = TRUE, scale. = T, tol = NULL)
```

> summary(p) #PCA 结果作为后续分析的遗传数据

> predict(p) #计算标准化的主成分得分

> gene = read.table ("clipboard", header=T) #导入遗传数据 ##导入叶形数据

```
> sy= read.table ("clipboard", header=T) #导入从 Morpho J 中导出的 对称组分 PCA 数据
```

```
> asy= read.table ("clipboard", header=T) #导入从 Morpho J 中导出的非对称组分 PCA 数据
```

```
> size= read.table ("clipboard", header=T) #导入从 Morpho J 中导出的叶片大小数据
```

##筛选气候数据

>library(usdm)

>data=read.table ("clipboard", header=F) # 导入气候因子数据

```
>x=vifcor(data,th=0.5) #通过设置参数(th=0.*)来调节保留的因子个数 >x #查看保留了哪些气候因子
```

```
##导气候和地理数据
```

```
> envo = read.table ("clipboard", header=T)
```

###数据标准化

```
> sy1= decostand (sy, "max")
```

```
> asy1= decostand (asy, "max")
```

```
> size1= decostand (size, "max")
```

```
> envo1= decostand(envo, "max")
```

```
> gene1= decostand(gene, "max")
```

##建立模型

```
> mod1=varpart (sy1,~ BIO03 + PREC12 + WIND03 + BIO17, ~ long +
lat + alt, gene1, data = envo1)
```

```
> mm1 = model.matrix (~ BIO03 + PREC12+ WIND03 + BIO17, envo1)
[, -1]
```

> mm2 = model.matrix (~ long + lat + alt, envol) [, -1] > mod2 = varpart (sy1, mm1, mm2, gene1) ###运行模型 > sy dbrda = capscale (sy1 ~ mm1 + mm2 + as.matrix (gene1), dist = "euclidean") > summary (sy dbrda) #查看结果 > sy anova = anova (sy dbrda, by = "term", permutations = 999) #方差 分析 > sy anova #查看方差分析结果 > sy dbrda #得 total 结果 ##控制地理和遗传的影响 > sy dbrda1 = capscale (sy1 ~ mm1 + Condition(mm2) + Condition (as.matrix (gene1)), dist = "euclidean") > sy anova1 = anova (sy dbrda, by = "term", permutations = 999) > sy anoval > sy dbrda1 #得 total 结果 ##控制气候和遗传的影响 > sy dbrda2 = capscale (sy1 ~ mm2 + Condition(mm1) + Condition (as.matrix (gene1)), dist = "euclidean") > sy anova2 = anova (sy dbrda, by = "term", permutations = 999) > sy anova2 > sy_dbrda2 #得 total 结果 ##控制气候和地理的影响 > sy dbrda3 = capscale (sy1 ~ Condition (mm1) + Condition(mm2) + as.matrix (gene1), dist = "euclidean") > sy anova3 = anova (sy dbrda, by = "term", permutations = 999) > sy anova3 > sy dbrda3 #得 total 结果 ###注: 以上是基于对称组分的代码, 基于非对称组分和叶片大小的 代码,只需循环执行建立和运行模型,修改标黄的部分即可。标黄的 部分为叶形数据,这里是以 syl:从 Morpho J 中导出的对称组分 PCA 数据为例。 2. 具体的气候, 地理和遗传因子对叶形的影响 ###导入数据 > sy= read.table ("clipboard", header=T) #导入从 Morpho J 中导出的 对称组分 PCA 数据

```
> asy= read.table ("clipboard", header=T) #导入从 Morpho J 中导出
的非对称组分 PCA 数据
> size= read.table ("clipboard", header=T) #导入从 Morpho J 中导出
的叶片大小数据
> factors = read.table ("clipboard", header=T) #导入解释因子,包括气
候, 地理和 SSR 的前 3 个 PC 值
###运行模型
> sy dbrda = capscale (sy1 ~ BIO03 + PREC12 + WIND03 + BIO17 + long
+ lat + alt + PC1 + PC2 + PC3, factors, dist = "euclidean")
> summary (sy dbrda) #查看结果
> sy anova = anova (sy dbrda, by = "term", permutations = 999)
                                                       #方差
分析
>sy anova #查看方差分析结果
> sy dbrda
          #得 total 结果
##控制地理和遗传的影响
> sy_dbrda1 = capscale (sy1 ~ BIO03 + PREC12+ WIND03 + BIO17 +
Condition (long + lat + alt) + Condition (PC1 + PC2 + PC3), factors, dist =
"euclidean")
> sy anova1 = anova (sy_dbrda, by = "term", permutations = 999)
> sy anoval
            #得 total 结果
> sy dbrda1
##控制气候和遗传的影响
> sy dbrda2 = capscale (sy1 ~ long + lat + alt + Condition (BIO03 +
PREC12+WIND03 + BIO17) + Condition (PC1 + PC2 + PC3), factors, dist
= "euclidean")
> sy anova2 = anova (sy dbrda, by = "term", permutations = 999)
> sy anova2
> sy dbrda2
           #得 total 结果
##控制气候和地理的影响
> sy dbrda3 = capscale (sy1 ~ PC1 +PC2 +PC3 + Condition (BIO03 +
PREC12+ WIND03 + BIO17) + Condition (long + lat + alt), factors, dist
= "euclidean")
> sy anova3 = anova (sy dbrda, by = "term", permutations = 999)
> sy anova3
           #得 total 结果
> sy dbrda3
###注: 以上是基于对称组分的代码, 基于非对称组分和叶片大小的
```

代码,只需循环执行建立和运行模型,修改标黄的部分即可。标黄的部分为需要进行分析的叶形数据,这里是以 syl:从 Morpho J 中导出的对称组分 PCA 数据为例。

六、导入协变量(Import Covariates)

1.导入文件格式(Input file formats)

用于叶形分析软件 Morpho J 的数据输入格式为 .txt 格式,所以 在导入协变量之前需要将标志点的坐标数据文件转化成 Morpho J 支 持的文件类型。该文件可以在 Excel 中准备,并保存为以制表符或逗 号分隔的文本,文件格式如图 61:

🥘 popspe	经纬	度协变量.t	xt - 记事z	4
文件(<u>F</u>)编	辑(<u>E</u>)	格式(<u>O</u>)	查看(<u>∨</u>)	帮助(<u>H</u>)
ID	Lo	ngitud	e	Latitude
BYS_S	10)5.61	33.33	
DLT_A	11	2.69	33.87	
DSH_M	12	24.49	41.11	
FHS_F	10)6.92	27.7	
HPS_A	10)7.71	34.09	
HQL_S	11	1.81	33.81	
JFS_S	10)5.68	33.69	
JGS_F	11	4.09	31.86	
JLP_D	11	3.33	35.71	
JYS_F	10)6.39	29.84	
LSS_F	11	5.97	29.59	
MSS_D	11	7.97	35.55	
NNY_M	11	1.32	37.09	
QPS_M	12	23.67	41.93	
QSS_M	12	23.14	41	
SBP_A	11	2.83	35.22	
SBY_A	10)8.96	35.02	
SPS_D	11	9.52	40.06	
SYB_D	10	07.14	33.7	
TBB_S	10)5.37	32.23	

图 61 协变量数据的整理

第一行包含协变量的名称。第一个条目(上面的例子中的"ID") 将被忽略,但是剩下的条目将被用作协变量的名称。每个协变量都应 该有一个不同的名字,必须匹配精确,包括拼写和大小写字母,数字 可以是整数也可以有小数(例如: "34", "30.5"), 也可以用科学记数 法(例如: "1.753 e-09")。

2. 导入协变量

打开 Morpho J, 在对话框的顶部, 有一个标签指示已选择的数据集。协变量将只导入到这个数据集。在"Project Tree"选项卡中选择数据集, 然后在"File"中点击"Import Covariates", 导入数据, 如图 62:

1	MorphoJ 1.06d	
File	Preliminaries Variatio	n Co
Сгеа	ate New Project	
Сгеа	ate New Dataset	
Ope	n Project	
Оре	n Recent Project	•
Sav	e Project	
Sav	e Project As	
Exp	ort Dataset	
Ехр	ort Covariance Matrix	
Imp	ort Classifier Variables	
Imp	ort Covariates	
Imp	ort Covariance Matrix	
 Imp	ort Phylogeny File	
Imp	ort Phylogeny Tree Set	
Imp	ort Outline File	
Imp	ort Shape Change Vector	s
Clos	se Project	
Exit		

图 62 协变量导入

将调用以下对话框:

Data set: y5				
Match by 🔘 I	dentifier			
۱	Classifier variable:	popspe	-	
🖌 The first line	of the file contains	the names of t	the new variables.	
查找(): 📑	白栎组叶形			
	Advantation and a state		N M - M	
D popspe经	纬度协受里.txt		Y.pat	
S.pdf			🗋 白栎F.txt	
Swm1123.	nexus		🗋 白栎组论文.doc	
swnj123.r	nexus		☐ 白栎组论文.pdf	
wxCVA.ai	i			201215(1).doc
wxCVAas	ym.ai			201215_1(1).doc
•				•
	And the set of the set			
文件名(N):	popspe经纬度协变	堂.txt		
文件类型(I):	所有文件			-
				打开 取消

图 63 导入协变量至已经编辑好的分类器中

其中导入有两个分类选择,分别是 Identifier 和 Classifier variable, 用于选择协变量如何与数据集中的观察值匹配。如果选择了 Identifier, 则 Identifier 变量的值将用于匹配。如果选择了 Classifier variable,则 可以从下拉菜单中选择我们已经定义过的分类器,并使用其值进行匹 配。导入协变量后,通过在初步菜单中使用编辑协变量来再次检查结 果。导入成功后选择该数据集在 Preliminaries 中点击"Edit Covariates" 查看编辑协变量:

	∳	MorphoJ 1.06d					
	File	Preliminaries Variation Covariation C	Com	parison	Genetic		
	Pro	New Procrustes Fit	ts				
	🕘 n	Find Outliers					
	٩	Include or Exclude Observations					
<		Generate Covariance Matrix					
		Extract New Classifier From ID String					
		Edit Classifiers	nm	amatria aomananan			
\mathbf{N}		Edit Covariates	d in	d ind, Symmetric com			
\mathbf{X}		Select Landmarks	ō, a∖	averaged ind, Symm			
		Subdivide Dataset By	abte	abted: DO ecores O			
		Combine Datasets	sc	ores, Co	vMatrix, y5		
		Average Observations By					
		Link Datasets					
		Create or Edit Wireframe					
		Set Options for Shape Graphs					
		← III PLS 地理 scores ●- III PLS asvm deo	_				

图 64 编辑协变量

导入成功如下图所示,在此界面可进行复制粘贴,删除等操作, 并单击 Accept:

秒 Morph	oJ 1.06d						_			Х	
File Prelimin	aries Varia	tion Co	variation Com	parison Genetic	CS	Help					
Project Tree	Results	Graph	ics Reports								
PLS y5 sym	PLS asym	n geo	Edit covariates i	n 'y5, averaged i	nd'						
Identifier	Longitude	Latit	ude								
BYS_S_T01	105.61	33.33								- I.A.	
BYS_S_T02	105.61	33.33									
BYS S T03	105.61	33.33									
BYS_S_T04	105.61	33.33								- 14	
BYS_S_T05	105.61	33.33									
BYS S TO6	105.61	33.33									
BYS S TO7	105.61	33.33									
BYS S TO8	105.61	33.33									
DLT A T01	112.69	33.87									
DLT A T02	112.69	33.87									
DLT A T03	112.69	33.87									
DLT A T04	112.69	33.87									67.
DLT A T05	112.69	33.87									
DLT A TO6	112.69	33.87									
DLT A T07	112.69	33.87									
DLT A T08	112.69	33.87									
DLT A T09	112.69	33.87									
DSH M T01	124.49	41.11									
DSH M T02	124.49	41.11									
DSH M T03	124.49	41.11									
DSH M T04	124 49	41 11									
DSH M T05	124 49	41.11								-	
Longitude			Add covariat	•		Rename covariate					
Latitudo			Add Covariat	6		Nellanie Covanate					
Lautuue											
			Delete ccovaria	te(s)		Replace values					
										-	
							Cancel		Accept		
		_						_			

图 65 查看接受协变量

3.双标图(Biplot)

在 Morpho J 中, 双块偏最小二乘(2B-PLS)的实现使用了两块变量间协方差矩阵的奇异值分解(e.g. Rohlf and Corti, 2000)。Biplot (双标图) 是一种图形表示形式,最初开发它是为了在同一图中结合观测变量和观测结果的信息,但在 Morpho J 中仅用于说明不能以形状变化的形式来可视化的变量。

步骤:点击"Covariation"→"Partial Least Squares"→"Two separate blocks",出现图 66 对话框:先选 Block1 为协变量,再选 Block2 为 对称组分或非对称组分,点击执行

	🐠 PLS analysis: two s	separate blocks		×	
	Name for the analysis:	PLS			
	Block 1 (Please choose Block 1 be	fore Block 2.)	Block 2		
	Datasets:				
	v5, averaged pop		v5. averaged ind		
	y5, averaged ind		PC scores, CovMatrix, y5, ave	eraged i	
	v5 averaged leaf	▼		•	
	Data matrices:				
	y5, symmetric component	^	y5, centroid size		
	y5, asymmetry component	t 📃	y5, symmetric component	=	
	y5, averaged ind, Covariate	es 👻	y5, asymmetry component	•	
	Variables:				Jhr
	Longitude		Symmetric component		
	Latitude				
	- I		1		/
	Perform permutation te	st Number of ro	unds: 10000	7	
	Pooled analysis within	subgroups			
	рор				
	spe				
	ind		Cancel	xecute	
	图 66		inlat)参数选择	<u> </u>	
	图 00 2	~~小凶(B	ipioug今致远科	-	
结果如图 67:					
1	Morphol 1 06d	-		- п х	
File	Preliminaries Variation Covariation	Comparison Genetics I	lelp		
Pr	oject Tree Results Graphics Rej	ports			
P	LS y5 sym PLS asym geo Edit covar	riates in 'y5, averaged ind'			
y t	5, asymmetry component Block 2 A	mounts of covariance PL	S scores, block 1 vs 2 Group-centered	PLS scores S	
	10 -		, Latitude		
	0.8 -				
1					



-0.3 0.0 0.3 Block 2 PLS1 0.5 0.8 1.0

Block 2 PLS2

-0.5 --0.8 --1.0 --1.0

-0.8 -0.5

该图显示了两条 PLS 轴(水平方向的 PLS1 和垂直方向的 PLS2)。 可以看出, PLS1 主要是变量"经度", PLS2 轴主要是"纬度"。

七、映射到系统发育树(Map Onto Phylogeny)

1.导入系统发育树

Morpho J 可以使用平方变化简约法(squared-change parsimony)将 形状数据映射到系统发育树上(Maddison, 1991)。系统发育文件以 NEXUS 文件格式导入 Morpho J, 再导入数据之前确保 PCA 分析已 经完成。

NEXUS 文件是由 Mesquite、PAUP*、MacClade 等程序生成的, 格式如下图:

文件(E) 编辑(E) 格式(Q) 查看(⊻) 帮助(出) #NEXUS

BEGIN TREES;

Tree tree = ((F:0.00029610,M:0.00014854)1.0000:0.00026224,S:0.00014932, (A:0.00053028,D:0.00018217)0.9910:0.00008139);

END;

图 68 导入数据格式 在"File"中选择"Import Phylogeny File"导入文件:



图 69 导入上一步 NEXUS 文件

2.独立变量在系统发育上映射

将形态测量数据映射到系统发育树上,点击"Comparison"中的"Map Onto Phylogeny"。将出现以下对话框:

		🐠 Map da	ta onto ph	ylogeny		×	
		Phylogeny:	tree		•		
		Name:	tree1				
		Data:					
		Datase	et: y5, avera	aged ind	-		
		Data ty	vpe: Covaria	tes	-		
		Classi	fier for taxa:	spe	•		
		Weighting	by branch leng	gths (if available)			
		Mapping w	vith rooted tree	e. (This is the usual option.)		Vy/
		Permutati	on test for phy	logenetic signal			
		Number of ite	rations: 10	000			
				Cancel	Execute		
]	图 70 将	独立变	量映射到系约	充发育材	对上	
点击	"Excute"	得到如-	下结果:			_	~
	File Preliminaries Van	riation Covariation	Comparison	Genetics Help		- U	~
	Project Tree Results	Graphics R	eports				
	Tree Values						
			≪ ⊔				
				\mathbf{X}		*	
				Ĩ			
	e e	·					
			\checkmark				
	106	108 1	10 112	Longitude	118	120 12	2
	L						

图 71 系统发育树结果

横轴提供变量的值(这里是质心大小), 而垂直方向对应于从树的 根开始的累积分支长度(如果映射是通过加权的平方变化简化完成 的)。如果分支长度可以被解释为进化时间, 那么这个图就代表了进化 谱系沿着变量轴的运动。 在该界面任意地方左击鼠标,就会出现一个弹出式菜单。该弹出 菜单可用于改变图形的方向,比例因子, PC 轴显示,导出图形文件 并保存。



再进行映射, 依次点击"Comparison"→"Map Onto Phylogeny", 做如下图操作:



图 75 PC scores 文件映射系统发育树结果

八、假定模块化评估(Modularity: Evaluate

Hypothesis)

模块是由许多交互作用形成的高度整合的单元,但又相对独立于其他 这样的单元(Klingenberg, 2008)。在形态测量分析中,形态模块化是指 不同模块之间的弱共变和模块内部的强共变。注意:模块化分析适合 于单一物种。

为了评估模块化假设,可将假设模块之间的共变程度与整个结构 的可选分区进行比较。Klingenberg 等(2009)提出 RV 系数, RV 系数 是平方相关的多元类比(Escoufier, 1973),它的取值范围从 0 (完全不 相关)到1 (完全相关)。通过比较与可选分区子集之间的 RV 系数来 衡量两个或多个标志点子集之间相关性。如果一个假设划分与模块之 间的真正边界一致,那么子集之间的相关性应该是最小的。

1.选择标志点子集

在 Project Tree 选项卡中选择一个协方差矩阵(通过这个图标 ☑在 项目树中可以识别)。在"Covariation"中点击"Modularity: Evaluate Hypothesis", 如图 76:



图 76 选择 Covariation 菜单下的 Modularity: Evaluate Hypothesis 然后出现下图对话框:



图 77 Modularity: Evaluate Hypothesis

第一个元素是一个用于指定分析名称的文本字段(Name for the analysis),可自行定义名称。在它下面有一个复选框,用于指定在比较中只包含标志点的空间连续分区还是包含所有分区。是否选择这个选项取决于研究的生物学背景(Klingenberg, 2009)。下一个选择是分析应该使用所有可能分区的完整枚举还是大量随机分区;如果选择了Full enumeration of partitions 选项,则限制为空间连续的分区,如果选择Random partitions 选项,将激活用于输入随机分区数量的文本字段。Klingenberg 建议 20 个标志点以内可以实现完整枚举,除此之外推荐10,000 数量级的随机分区进行比较,提供 RV 系数分布的合理表征。

2.修改邻接图

点击"Accept"之后,我们在第三个选项 Modify adjacency graphs 中,将对称组分中的邻接点的线补充完整(可以用鼠标点击标志点直接进行连接,也可以在右侧操作界面点击标志点名称增加连接线段)并点击"Accept":



图 78 修改邻接图

点击"Accept"后我们在跳转页面中右上角(Number of subsets) 选择分区数量,每个分区至少包含2对标志点,左侧图是一个标志点 配置的图,显示了作为模块化假设指定的标志点子集(彩色点)和用于 建立空间连续分区的邻接图(仅在选择了该选择时使用)



图 80 选择分区子集

3.比较分区(Comparing partitions)

点击"Accept"之后出现 RV 系数分布直方图:



图 81 RV 系数的频数分布直方图

评估将标志点配置划分为两个模块的假设的最直接的方法是:计 算所有可能的划分为适当大小子集的分区的 RV 系数。如果模块化假 设成立,则根据该假设分区的 RV 系数应该是最低值,或者至少应该 接近所有分区的 RV 系数分布的下限。

当红色箭头在整个分区的接近左侧末端时,这一结果与划分的子 集是独立模块的假设一致。当箭头在右侧时则与假定的划分相反。将 设置不同分区的 RV 系数进行比较,可以提供比统计测试 P 值更多的 信息。当 RV 系数小于或等于设置的分区的 RV 值的分区比例,可以 解释为 P 值的模拟(The proportion of partitions for which the RV coefficient is less than or equal to the RV value for the partition of interest, which can be interpreted as the analog of such a P-value, is one piece of information that can be obtained. Klingenberg, 2009)

九、深度学习(Deep Learning)

机器学习指一类可以通过检测数据中的模式自动生成预测模型的 算法(Olden et al., 2008)。深度学习 (Deep Learning) 作为机器学习的 一个分支,指基于样本数据通过一定的训练方法得到包含多个层级的 深度网络结构的机器学习过程(Hinton, 2006)。使用一般的学习过程, 深度学习算法能够自动地从数据中检测和提取特征。这种自动学习过 程是通过将数据分解成多个层来实现的,每个层都有不同的抽象级别, 允许算法学习表示数据的复杂特征(Christin et al., 2019)。

1.选择体系结构

生态学中许多层面的研究通常提供高度维度的数据集,通过这些数据集,深度学习尤其准确和高效。在实践中,使用不同的深度学习 架构可以实现这些结果。其中,使用最广泛的是卷积神经网络 (convolutional neural network, CNN),由于其在图像分类中的性能,该架构有助于普及深度学习(Krizhevsky *et al.*, 2017)。CNN 是因其易于训练和良好的泛化能力而受到最多关注,是第一个真正意义上具有多层结构的深度学习模型,提高了数据量过于庞大时的处理效率,实现了特征自动提取的需求(LeCun *et al.*, 2014)。CNN 基本结构为:输入层,卷积层,池化层,全接接层和输出层。其中卷积层就是特征提取层,所包含的卷积核则是神经网络训练的核心。

2.选择框架

随着深度学习的快速发展,创建了大量框架。所有框架(如下图)在易用性、可用资源、架构支持、可定制性或硬件支持方面提供了不同的方案(Christin *et al.*, 2019)。

Framework	Language	ONNX Support	Implementation of popular networks	Pre-trained models available	URL
Tensorflow	Python, C/C++, R, Java, Go, Julia	Needs conversion from external tool	Yes	Yes	https://www.tenso rflow.org/
PyTorch	Python	Yes	Yes	Yes	https://pytor ch.org/
Keras ^a	Python, R	Needs conversion from external tool	Yes	Yes	https://keras.io/
Microsoft Cognitive Toolkit (CNTK)	C#, C++, Python	Yes	Yes	Yes	https://docs.micro soft.com/en-us/ cognitive-toolkit/
Deeplearning4J	Java, Scala	Basic support	Yes	Yes	https://deeplearni ng4j.org/
MATLAB + Deep Learning Toolbox	MATLAB	Yes	Yes	Yes	https://www. mathworks.com/ products/deep- learning.html
Apache MXNET	C++, Python, Julia, Matl ab, JavaScript, Go, R, S cala, Perl	Yes	Yes	Yes	http://mxnet.incub ator.apache.org/
PlaidML	Python	Yes	Yes (via keras)	Yes (via keras)	https://github.com/ plaidml/plaidml

图 82 不同框架对比图

3.选择模型

在 ILSVRC (imagenet large scale visual recognition challenge)比赛 上,研究者提出了一些新的 CNN 模型,其 2014 年的冠军 GoogLeNet 采用了 Inception 结构(在同一层中使用不同大小的卷积核),用以提 高分类效果。并且在 Inception 之后又增加了一些改进,Xception则为 Inception 的极限版本,在 ImageNet 数据库上测试的分类 top-5 正确率 高达 94.5%,是分类模型中比较出色的选择。

Xception 结构是一个由深度可分离的卷积层和残留连接组成的 线性堆栈。这使得架构非常容易定义和修改,使用像 Keras (https://github.com/fchollet/keras)或 TensorFlow-Slim 这样的高级库, 它只需要 30 到 40 行代码,这与 VGG-16 这样的架构没有什么不同, 而像 Inception V2 或 V3 这样的架构,它们的定义要复杂得多。

4.获取训练集

训练数据集通常包含数以百万计的例子,这取决于任务、检测项目的数量和期望的性能(Pondenkandath et al., 2018)。总的来说,训练数据集越大,分类准确率就越高。然而,重要的是要确保每个分类类别都有足够的测试,以避免识别偏差(Wearn et al., 2019)。还要注意,训练数据集通常分为两个子集,一个用于有效地训练模型,另一个用于评估其性能。因此,每个子集必须包含足够的例子代表所有分类类别,以便有效的训练。

5.具体操作步骤:

5.1 准备软件



用于连接服务器, 输入命令等操作, 下载地址:

https://xshell.en.softonic.com/



用于上传查看文件,下载地址: https://www.pcsoft.com.cn/soft/13932.html



用于编辑修改 python 文件, 下载地址: https://www.jetbrains.com/pycharm/

5.2 数据准备及处理

5.2.1 准备用于分类的图片

准备图片数据, 按训练集: 验证集: 测试集分别是 70:15:15 的 比例准备图片, 将图片随机排列并重命名 (BJFU_001594), 使每一张 图片有自己的编号, 放进如下图所示的位置相应的 input 文件夹



并在 GroundTruth 中 (.csv 格式) 对每个编号的图片进行如下图

定义(以三分类为例 DEN、ALI、HYB),1代表正确,0代表否定, 注意 input 的图片总数和 groundtruth 的总数要一一对应(某一编号可 缺失,但不可对应不上)

	A	B	C	D	
1	image	DEN	ALI	HYB	
2	BJFU_001926	1	0	0	
3	BJFU_001927	1	0	0	
4	BJFU_001931	0	1	0	
5	BJFU_001933	1	0	0	
6	BJFU_001934	1	0	0	
7	BJFU_001935	0	1	0	
8	BJFU_001936	1	0	0	
9	BJFU_001937	1	0	0	
10	BJFU_001938	1	0	0	
11	BJFU_001940	1	0	0	
12	BJFU_001941	0	1	0	
13	BJFU_001944	1	0	0	
14	BJFU_001945	0	1	0	
15	BJFU_001946	0	1	0	
16	BJFU_001947	0	1	0	
17	BJFU_001949	0	1	0	
18	BJFU_001950	1	0	0	
19	BJFU_001953	1	0	0	
20	BJFU_001954	1	0	0	
21	BJFU_001955	1	0	0	
22	BJFU_001957	1	0	0	
	图 84 Gro	undTru	th示例]	

5.2.2 修改运行命令

图 85 运行命令位置图

修改在上图位置的 python 文件 (__init__、train_new), 只需要改 相应的路径和命名

在__init__.py 中如下更改:

图片文件夹 task3_img = 'BJFU_Training_Input' task3_validation_img = 'BJFU_Validation_Input' task3_test_img = 'BJFU_Test_Input'

图 86 更改位置示例

图片文件夹如果没有重命名无需改动

分类的名称需要以物种为基础进行更改,如下图所示(以三分类为例 DEN、ALI、HYB)

```
# 标签文件夹名字
task3_gt = 'BJFU_Training_GroundTruth'
task3_val_gt = 'BJFU_Validation_GroundTruth'
task3_test_gt = 'BJFU_Test_GroundTruth'
DEN = 0_# Melanoma
ALI = 1 # Melanocytic nevus
HYB = 2 # Basal cell carcinoma
classes = [DEN, ALI, HYB]
class_names = ['DEN', 'ALI', 'HYB']
```

图 87 更改位置示例

在 train new 中如下更改:



图 88 更改位置示例

文件路径需按所存文件夹更改

5.3 数据上传及运行

5.3.1 数据上传

通过Xftp上传图片、groundtruth、_init_.py、train_new.py到相应 位置(一千张图片约7小时)

5.3.2 数据运行

在 Xshell 中进入到 runs 文件夹直接输入 train_new.py 进行图片 的训练与验证(一千张图片约 10 小时),运行训练集输出结果(以三 分类为例 DEN、ALI、HYB)如下:

50/50 [========================] - 22s 447ms/step - loss: 0.9094 - accuracy: 0.8706 val_loss: 3.2083 - val_accuracy: 0.4355 - val_sensitivity_at_specificity: 0.9395								
1	True\Pred	DEN	ALI	HYB	TOTAL			
Ì	DEN		θ.22	θ.12	170			
Ì	ALI	0.64		θ.12	75			
Ì	HYB	0.68	0.25		65			
Ì	TOTAL	204	72	34	Ì			
1	I	DEN	ALI	HYB	MEAN			
	precision				0.315			
	recall	0.659			0.325			

图 89 运行训练集结果

训练后输入 cls_predict.py 文件进行预测, 输出结果如下:

310/310 [=======] - 2s 6ms/step [5.7973313331604, 0.44516128301620483]								
1 1	DEN	ALI	HYB	ME AN				
precision	0.545	0.224	0.238	θ.335				
recall	0.682	0.227	0.077	θ.329				

图 90 运行测试集结果

若重复操作将下面位置 cache 文件夹内容清空

» bjfu > bjfu > datasets > ISIC2018 > cache

图 91 清空数据位置

十、测量叶片传统性状指标

1. 基础指标

这里以叶片长为例,此方法适用于测量叶长(lamina length)(叶顶 端到叶基部),叶宽(lobe width)(叶片最宽处),窦宽(sinus width)(最 宽叶片下边的凹陷点),叶柄长(petiole length),叶基部到叶最宽处的 长度(wp: length of the lamina base to the widest part)等(图 92)。



具体步骤:

1.1 批量导入图像

选择"File"→"Import"→"Image Sequence"选定并打开叶片图像 (图 93, 94)。



1.2 测量叶长



图 95 选取直线工具测量叶长

¢	Results										\times
File	Edit	Font I	Results								
	Area	Perim.	Angle	Feret	FeretX	FeretY	FeretAngle	MinFeret	Length		^
1	1599	1598	-90	1598	1644	1216	90	0	1598		
	图 96 获得结果										

1.3 像素转换

转换之前查看每张图片的分辨率,以图 97 为例,400 dpi (dot per inch)表示每英寸长度内有400 个像素点数(dot),换算一下,一个像素 点(dot)=(1/400) inch,其中1 inch=2.54 cm,则1 dot=(2.54/400) cm。 实际叶长 = 测得的长度像素点数 × (2.54/分辨率)
NUT_A_T02_L05 属性

常规 安全	详细信息 以前的版本	
属性	值	
分级	****	
标记		
备注		
来源 ——		
作者		
拍摄日期		
程序名称	IJ Scan Utility	-/
获取日期		1h7.
版权		· · · · · · · · · · · · · · · · · · ·
图像 ——		
图像 ID		
分辨率	3307 x 4677	
宽度	3307 像素	
高度	4677 像素	
水平分辨率	400 dpi	
垂直分辨率	400 dpi	
位深度	24	
压缩		
分辨率单位	2	
\$54.+二	-000	

图 97 图片属性

1.4 其余叶性状指标

其余指标如裂叶数量(lobe number), 夹层脉数量(number of intercalary veins) (an intercalary vein was a secondary vein irrigating a sinus and extending at least half way from the midrib to the base of the sinus), 叶基部形状(basal shape of the lamina)直接计数即可。其中叶基 部形状记录为 1-9 (图 98):



图 98 叶基部形状 (Kremer et al., 2002)

2. 叶脉密度(Leaf Vein Density)

叶脉密度(mm mm⁻²)是指单位面积的叶脉长度。

每片叶子从叶片的底部到顶部都有一条主脉(1° vein),还有许 多平行的二级脉(2° vein)分支。在目前的研究中,一级和二级叶脉, 也称为主要叶脉(major vein),从二级叶脉分支的叶脉被定义为三级叶 脉(3° vein)。在栎属的叶片中,三级及以上的叶脉无法清楚地区分, 因此将三级及以上级别叶脉的叶脉密度统称为小脉密度(Minor vein density)。

测量叶脉密度,包括一级叶脉密度(1° vein density)、二级叶脉密度(2° vein density)和小脉密度(Minor vein density)。



图 2 直线工具测量一级叶脉长度

2.2 二级叶脉密度(2° vein density)

用 ImageJ 软件的直线工具测量所有二级叶脉的长度总和, 除以 叶面积, 即为该叶片的二级叶脉密度。



图 3 直线工具测量二级叶脉长度

2.3 小脉密度(Minor vein density)

为了得到小脉密度,我们需要使用带有图像软件(徕卡应用套件)的计算机连接的显微镜拍摄小脉图像,再结合 ImageJ 软件进行测量。

具体操作步骤:

2.3.1 叶片样本准备

选择生长良好、具有代表性的叶片,尽量保证叶片完整且无病 虫害损伤。避开主脉剪下一小片叶片放到载玻片上(叶背朝上),在 叶片上滴加适量的清水,盖上盖玻片。

2.3.2 显微镜观察与图像采集

(1)打开仪器,顺序为①荧光→②显微镜→③电脑,关机顺序相反



图 4 仪器开关机顺序

(2) 将制备好的叶片切片放在显微镜载物台上,在电脑上打开LAS





(4) 选择明场"TL-BF"、"color mode",通过调节显微镜的焦距和载物台的位置,在10倍物镜下调节 Exposure[ms]、Gain、Intensity、Aperture 等参数至合适亮度,使叶片小脉清晰地呈现在视野中



图 5 参数设置

(5)单击右键,选择"Export Image",可保存为 JPEG、TIFF 等格式,点击"save"保存



图6保存图片

2.3.3 ImageJ 软件测量

(1)批量导入图像

选择"File"→"Import"→"Image Sequence"选定并打开叶片图 像。

(2)设置添加比例尺

①用直线工具沿比例尺画线



图7比例尺画线

②设置比例尺,选择"Analyze"→"Set Scale",在"Known distance"和"Unit of length"设置该直线的已知长度和单位,勾选"Global"对所有照片应用这一比例。

	🛓 Set Scale	×
XXX	Distance in pixels: 32. 3333 Known distance: 20 Pixel aspect ratio: 1. 0 Unit of length: µm]
	Click to Remove : Global Scale: 1.6167 pixels/µm	Scale
	图 8 设置比例尺长	度

(3)测量小脉长度

用直线工具测量所有小脉长度

19/33 (00H4F-T07-L0F); 1152 00x864 00 µm (1920x1440); ROB; 348MB	# Results					
	File Edit Font Results					
	Area	Mean	Min	Max	Angle	Length
	19 216.00	0 82.404	38.492	195.174	-26 266	359.152
	20 171.72	0 69.871	44.562	159.122	-96.590	285 538
	21 68.760	54.195	40.570	102.530	5.440	113.703
	22 97.920	56.813	44.399	112 703	-86.621	162.746
	23 159.13	0 74 255	43.886	201.258	-136.469	264 705
The second s	24 119.50	0 60.513	38 669	137.851	-9.234	198.365
	25 120.60	0 64.615	40.540	114.914	-73.137	200.531
	26 85.680	54.624	42.902	119.787	-65.158	142.408
	27. 83.520	59.959	30.000	178.208	-19.454	138.590
	26 73.060	79.100	43.212	170.406	1.701	121 259
	29 101.52	0 76.550	40.956	164.411	-51.340	168.681
	30 124.20	0 70.864	37.549	105.709	27.085	206.685
	31 109 80	0 53.868	34,723	126.764	-102.161	102 592
	32 114.12	0 65 842	36.766	155.072	-94.357	189.681
	33 202.68	0 55.295	27.358	147,196	-28.757	337.362
	34 54,360	71.068	35.450	169.839	-142.826	90.297
	35 56.160	41.531	25.324	107.309	-103.410	93.035
	36 86.760	81.857	25.915	199.171	-24.108	144.283
	37 91.440	56.519	44.616	91.285	-82.959	151.643
	38 87.840	67.478	51.347	96.592	-60.372	145.531
	39 156.60	0 52 813	40.451	02.533	-98.897	260.128
	40 201.96	0 43,670	28.526	72.812	-76 262	335.931
	41 143.64	0 54.923	38.114	84 194	21.400	238 662
	42 126.36	0 74.169	38.636	194.005	-52,441	209.964
	43 144.36	0 67.213	32.971	161.975	30.669	240.052
	44 95.760	50.924	34.602	112.033	-78.435	159.064
	45 111.96	0 55 006	33.717	191.252	16 275	195.865
	46 00.260	51.701	34.623	177.936	-77.779	133.479
	47 88.200	38.096	19.386	83.743	13.468	146 648

图 9 小脉长度测量

(4)计算小脉密度

将所有小脉长度求和,除以视野面积,即为小脉密度。(需将单位换算为微米,1µm=0.001mm)



图 10 小脉视野面积

3. 比叶面积(Specific Leaf Area)

比叶面积是指单位叶面积与干重的比值(cm²/g),是叶功能性状指标,与叶片光合作用和产量相关。

为了获得叶面积,我们使用扫描仪以300 dpi的分辨率批量扫描

带叶柄的完全伸展的叶片,并使用 ImageJ 软件 (https://imagej.nih.gov/ij/download.html)估算叶面积。

具体操作步骤:

3.1 导入图像(Import image)

批量导入图像:选择"File"→"Import"→"Image Sequence"选定并 打开叶片图像。

3.2 参数设置(Parameter Setting)

(1) 依次选择"Image"→"Type"→"8bit" (图 99), 分解 RGB 图 像为灰度图,所得结果最为精确。

(🛓 In	nageJ							—	
	File	Edit	Image	Process	Analyze	Plugir	ns Window	Help		
		0, 0	Туре			•	✔ 8-bit		3 8	
	Oval	', elliptio	Adjus	t		•	16-bit			
		Ę	Show	Info	Ctrl+I		32-bit			
			Properties		Ctrl+Shift+P		8-bit Color			
		•	Color			•	RGB Colo	r		
			Stack	s		•	RGB Stac	k		
			Нуре	rstacks		•	HSB Stack	k		

图 99 分解 RGB 图像为灰度图

(2) 选择"Image" → "Adjust" → "Threshold"顺序调整待测对象
 为 Red (图 100), 点击 Apply。





Distance in pixels: 分辨率单位 dpi 指的是每英寸所打印的点数, 因为 1 英寸(in)=2.54 厘米(cm), 同时我们的分辨率 300dpi=300 pix per in, 则像素距离(Distance in pixels) = 300 pix / 2.54cm = 118.11 pix/cm。这里有一点需要注意: 测叶面积之前要看一下图片的分 辨率具体是多少(点击图片右键→属性→详细信息)。 Known distance: 这个值不用设置, 自动生成为1 Unit of length: 设置单位为 cm

Global: 对于所有打开的照片应用该比例关系。如果有同批次同镜头 视野的多张照片,可以勾选此选项,一次性将所有照片中像素所 对应的实际距离进行设置。

- 有了这样的关系之后,我们利用 Image J 对位图以像素为单位进行测 量或者计算的时候,就与实际长度联系起来了。
- (4) 选择"Analyze"→"Analyze particles", 导出叶面积值(图 102)。

🛓 Analyze Particles 🛛 🗙	
Size (pixel^2): 10.89-Infinity Circularity: 0.00-1.00	743
Show: Outlines 💌	X
✓ Display results ✓ Exclude on edges	b
Clear results 🗌 Include holes	
Summarize Record starts	
🗖 Add to Manager 🛛 🗖 In situ Show	
OK Cancel Help	

图 102 过滤非目标的小杂质

Size 可以设置 10 或者更大起始,以过滤非目标的小杂质;我 们的叶子不是圆形, Circularity 的最大值可以设置的小一些; Exclude on edges 可以去除边界的非目标,但扫描时需要注意叶片不 能接触边界; show 可以选择 Outlines 看一下计算是否正确。由于我 们的扫描叶片数(count)是非常明确的,所以可以依次来检查匹配是 否正确。Result 界面可以连续出现数值,最后将所有结果统一 "Save as" 导出成 Excel 表格。

3.3 叶片千重(Dry Leaf Weight)

新鲜叶片在80℃通风烘箱中放置3d后,用天平称重。

3.4 比叶面积(Specific Leaf Area)

每个个体的比叶面积: 4-5 片叶片的总面积/4-5 片叶片的总干重。

3.5 批量计算图片中叶片大小(叶面积、叶最大直径)

(1) 下载 imgeJ 加载程序包 RosetteTracker (图 103);
 <u>https://telin.ugent.be/~jdvylder/RosetteTracker/</u>

W/ 🏛 👋	Jonas De Vylder					
GENT						
Home	RosetteTracker					
Education Publications	Image analysis of Arabidopsis thalana rosettes is an important non destructive method for studying plant growth. Some work on automatic rosette measurement using image analysis has been					
Downloads	proposed in the past, but is generally restricted to be used only in combination with specific high throughput monitoring systems. We introduce Rosette Tracket, a new oper source image analysis foo for evaluation of plant shoot phenotypes. This tool is not constraint to one specific monitoring system and can be adapted to different low budget imaging set ups and requires minimate input.					
Contact	contrast to previously described monotoring toos, isosette tracker anows to simutaneously quartury plating flowing, photosynthesis and least temperature related parameters through the analysis of VEs, chlorophyll fluorescence and/or thermal information and time tapes sequences. Autobalde for free, positive Tracker facilities gaining range of Arabidopsis tabilities generations and the sequences.					
	This software tool is discribed in: J. De Wylder, F. Vandeebusscher, Y. Hu, W. Philips, D. Van Der Straten, "Rosett Tracker: an open source image analysis tool for automatic quantification of					

图 103 RosetteTracker 下载

(2) 打开 imgeJ, 点击 Plugins 加载 Rosette Tracker, 点击 import images,
 选择 good contrast (目标叶片与背景颜色差异较大的图片), 导入图片;

(3) 点击 settings, 选择 diameter/are (最大直径/叶面积), a single image 中 segmentation 填写植物数量和列数;

(4) 图片标尺的换算:先在有标尺的图片上用直线工具描绘标尺,选择 Anysis 中的 set scale 换算 1plex=?mm,换算后的数据填入 set scale 中的 set manual;

(5) 点击 Day&night 页面, 取消 holes, 选取需要识别的叶片颜色, clutter 选择 3 默认, 点击 OK,开始 run!

十一、气候因子与叶性状的线性关系

1. 获取叶性状数据

从 Morpho J 中导出叶片大小和叶片形状 PC 的数据(参见三、 3.5), 作为叶形输入数据。

2. 获取气候数据

采样点气候数据的获取参考第五节中从 WorldClim 数据库中获 取气候因子的方法,第五节中使用方差膨胀因子分析(VIF)去除部分 相关性较强的因子,使用共线性低的具体气候因子做相关分析。

本节使用 R 语言的函数包 "vegan" (Oksanen et al., 2017) 将 19

个气候因子进行 PCA 降维分析. 使用 PC1 轴的数据来代表每个采样 点的气候(以下写为 cli PC1)。 PCA 分析代码: ####加载函数包 >library(vegan) ####导入数据 >envo=read.table("clipboard",header=T) ####建立模型 >pca princomp <- princomp(envo[,-19],cor=T) ####19个主成分的累积方差贡献率 >summary(pca princomp) ####前两个主成分的载荷值 >predict(pca princomp,loadings=TRUE) >pca princomp\$loadings[,1:2] ####每个采样点19个主成分的得分 >x=predict(pca princomp) >write.table(x,"output.txt")

3. 叶性状与气候因子的相关分析

为了调查不同物种之间叶性状与气候因子的相关关系,我们使用 R 语 言的函数包"ggplot2", "ggpubr", "lme4", "car"和"lmerTest" (Fox et al., 2013; Wickham, 2016), 进行线性分析和显著性分析。 运行代码如下: ####加载函数包 >library(ggplot2) #####设置 xy 轴的参数 >windowsFonts(RMN=windowsFont("Times New Roman")) >theme updatek=theme (>plot.title=element text(family="RMN", size=18. face="bold", colour="black", hjust = 0.5),>axis.title.x=element_text(family="RMN", size=14, face="bold", colour="black"), >axis.title.y=element text(family="RMN", size=14, face="bold", colour="black"), >axis.text.x=element text(family="RMN", size=10, face="bold",

```
colour="black"),
    >axis.text.y=element text(family="RMN", size=10, face="bold",
colour="black"),
    >axis.line = element line(size=1,colour = "black"))
theme legend=theme (
   >legend.text = element text(colour = 'black', size = 10, face = 'bold'),
   >legend.title =element text (colour = 'purple', size = 10,face =
'bold', angle = 0, hjust = 1, vjust = 1),
   >legend.background = element blank(),legend.key = element blank()
   )
####导入数据(个体水平, 一列 clipc1 数据, 一列叶型 PC1 数据, 一列
pop 数据 (用 1, 2, 3...来表示), 一列物种(species)数据)
    >data=read.table("clipboard",header=T)
####建立模型
    >p = ggplot (data, aes (x = cli pc1, y = leaf shape PC1, colour = species))
+ geom point () + scale x continuous (limits = c(-10,10)) + stat smooth
(method = lm, se = FALSE) + theme bw () + theme updatek +
theme legend
####得出散点图
    >p
#####每个物种气候因子与叶型的回归系数
####加载函数包
  >library(ggpubr)
####回归系数R值和P值
  >p+stat cor()
  >ggsave ("tmp4.pdf", width = 8, height = 8, units = "cm")
####显著性:
  >library(lme4)
  >library(car)
  >library(lmerTest)
####建立模型
    >m1 < -lmer (leaf shape PC1\sim cli PC1*species + (1|pop), data)
####显著性分析
    >anova (m1, type="I")
###注: 以上是基于叶片形状的代码, 基于叶片大小或其他性状的代
码,只需循环执行建立和运行模型。
```

十二、R语言常用的作图方法

1. 直方图

e.g., 三个组 A, D, H 的叶型 PC1 数据 (Du et al., 2022)



####加载包并导入数据
>library(ggplot2)
>data=read.table("clipboard",header=T)
#####作图 直方图重叠
>ggplot (data, aes (x = pc1, fill = spe)) + geom_histogram (position =
"identity", binwidth = 0.5, alpha = 0.4, colour = "black")

#####作图 直方图不重叠

>ggplot (data, aes (x = pc1, fill = spe)) + geom_histogram (position = 'dodge', binwidth = 0.1, alpha = 0.4, colour = "black")

2. 箱型图

e.g., 四个组 Aallo, Asym, Dallo, Dsym 的叶千重 (Du et al., 2022)



####加载包并导入数据

>library(ggplot2)

>library(ggpubr)

>data=read.table("clipboard",header=T)

####作图

>ggplot (data, aes (x = spe, y = mass, color = spe)) + geom_boxplot () + theme _ bw ()

####添加两两之间的显著性检验

>stat _ compare _ means (aes (label = .. p. signif ..), comparisons = list (c
('Asym', 'Aallo'), c('Asym', 'Dsym'), c('Aallo', 'Dallo'), c('Dsym',
'Dallo')))

3. 散点图加折线图

e.g., 三个组 A, D, H 的叶型 pc1 和 pc2 轴的数据 (Du et al., 2022)



参考文献

陈德钊. 多元数据处理. 化学工业出版社, 1998.

- 胡坤. 1992. 典型变量分析在化探样品分类中的应用. 物探化探计算 技术, ,14(3):196-199.
- *刘媛, 吕佳, 宋佳, 王玉垚, 王肖静, & 杜芳. 2017. 基于几何形态测 量学的植物鉴定方法. 植物科学学报, (6), 894-899.
- *宋佳, 侯盟, 鲁四海, 李俊清, & 杜芳. 2015. 一种基于标志点的叶片 几何形态分析方法. 兰州大学学报(自科版), 51(5), 705-710.

- *苏蔚, 宋以刚, 祁敏, 杜芳. 2021. 基于几何形态分析的栎属白栎组 叶片形态特征. 应用生态学报, 32, 2309.
- 杨维权等.1989. 多元统计分析. 高等教育出版社.
- Abràmoff MD, Magalhães PJ, Paulo J, Ram SJ. 2005. Image processing with IMAGEJ Part II. *Biophotonics International*, 11: 36-42.
- Christin S, Hervet É, Lecomte N. 2019. Applications for deep learning in ecology. *Methods in Ecology and Evolution*, 10(10): 1632-1644.
- *Du F, Qi M, Zhang YY, Petit RJ. 2022. Asymmetric character displacement in mixed oak stands. *New Phytologist*, 236, 1212-1224.
- Escoufier, Y. 1973. Le traitement des variables vectorielles. *Biometrics*, 29: 751–760.
- Hinton G E, Osindero S, The Y W. 2006. A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation*,18(7): 1572-1554.
- Klingenberg CP. 2003. A developmental perspective on developmental instability: theory, models and mechanisms. In: Polak M, ed. Developmental instability: causes and consequences. New York: Oxford University Press, 14–34.
- Klingenberg CP. 2011. MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources*,11:353-357.
- Klingenberg CP and McIntyre GS. 1998. Geometric morphometrics of developmental instability: analyzing patterns of fluctuating asymmetry with Procrustes methods. *Evolution*, 52:1363–1375.
- Klingenberg CP, Barluenga M, Meyer A. 2002. Shape analysis of symmetric structures: quantifying variation among individuals and asymmetry. *Evolution*, 56: 1909–1920.
- Klingenberg CP. 2008. Morphological integration and developmental modularity. *Annual Review of Ecology, Evolution and Systematics*, 39: 115-132.
- Klingenberg CP. 2009. Morphometric integration and modularity in configurations of landmarks: Tools for evaluating a-priori hypotheses. *Evolution & Development*, 11: 405-421.
- Krizhevsky, Alex, Sutskever, et al. 2017. ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*.
- Kremer A, Dupouey J L, Deans J D, Cottrell J, Csaikl U, Finkeldey R, *et al.* 2002. Leaf morphological differentiation between *Quercus robur* and *Quercus petraea* is stable across western European mixed oak

stands. Annals of Forest Science, 59(7): 777-787.

- Lecun Y, Boser B, Denker J, Henderson D, Howard R, Hubbard W, Jackel L. 2014. Backpropagation applied to Handwritten Zip Code Recognition. *Neural Computation*, 1(4): 541-551.
- Lele S, Richtsmeier JT. 1991. Euclidean distance matrix analysis: A coordinate-free approach for comparing biological shapes using landmark data. *American Journal of Physical Anthropology*, 86(3): 415-427.
- *Liu Y, Li Y, Song J, Zhang R, Yan Y, Wang Y, Du F. 2018. Geometric morphometric analyses of leaf shapes in two sympatric Chinese oaks: *Quercus dentata* Thunberg and *Quercus aliena* Blume (Fagaceae). *Annals of Forest Science*, 75: 90.
- *Li YJ, Zhang Y, Liao PC, Wang TR, Wang XY, Ueno S, Du F. 2021. Genetic, geographic, and climatic factors jointly shape leaf morphology of an alpine oak, *Quercus aquifolioides* Rehder & EH Wilson. *Annals of Forest Science*, 78, 1-18
- Maddison WP. Squared-change parsimony reconstructions of ancestral states for continuous-valued characters on a phylogenetic tree. 1991. *Systematic Biology*, 40(3): 304-14.
- Maddison DR, Swofford DL, and Maddison WP. 1997. NEXUS: an extensible file format for systematic information. *Systematic Biology*, 46: 590–621.
- Marugán-Lobón J. 2010. Combining Shape Data and Traditional Measurements with the 2B-PLS: Testing the Covariation Between Avian Brain Size and Cranial Shape Variation as an Example. *Morphometrics for Nonmorphometricians*.
- Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin P, O'Hara RB, Simpson G, Solymos P, Stevens M, Wagner H. 2017. vegan: Community ecology package. R package version.
- Olden, Jd, Lawler, et al. 2008. Machine learning methods without tears: A primer for ecologists. *Quarterly Review of Biology*,83(2): 171-193.
- Pondenkandath V, Seuret M, Ingold R, et al. 2018. Exploiting State-of-the-Art Deep Learning Methods for Document Image Analysis.
- *Qi M, Du F, Guo F, Yin KQ, Tang JJ. 2024. Species identification through deep learning and geometrical morphology in oaks (Quercus spp.): Pros and cons. *Ecology and Evolution*, 14(2): e11032.

- Rohlf FJ, Corti M. 2000. The use of two-block partial least-squares to study covariation in shape. *Syst. Biol*, 49: 740-753.
- Vieira M, Mayo SJ and Andrade IMD. 2014. Geometric morphometrics of leaves of Anacardium microcarpum Ducke and A. occidentale L. (Anacardiaceae) from the coastal region of Piauí, Brazil. *Brazilian Journal of Botany*, 37(3): 315-327.
- Viscosi V. 2015. eometric morphometrics and leaf phenotypic plasticity: assessing fluctuating asymmetry and allometry in European white oaks (*Quercus*). *Bot J Linn Soc*, 179: 335–348.
- Vujić V, Rubinjoni L, Selaković S and Cvetković D. 2016. Small-scale variations in leaf shape under anthropogenic disturbance in dioecious forest forb mercurialis perennis: a geometric morphometric examination. *Arch Biol Sci*, 68(4): 705-713.
- Wearn OR, Freeman R, Jacoby DMP. 2019. Responsible AI for conservation. *Nature Machine Intelligence*, 1(2): 72-73.
- *表示本实验室发表论文。

87