

中国种子植物区系定量化研究

I 区域性种子植物区系研究实用程序工作原理

(QX—系列程序)*

傅德志 李 奕 薛艳红

(中国科学院植物研究所系统与进化植物学开放研究实验室, 北京 100093)

摘 要

本文根据吴征镒教授对中国种子植物属分布区类型研究结果, 研制出定量化研究区域性种子植物区系的电子计算机程序, 应用本程序可完成某区系的分布区类型统计分析, 科属组成分析和与其它地区以共有属关系构建的相似性系数的计算, 同时可大量节省研究人员的劳动强度和时间, 对提高研究水平有一定的效果。

关键词: 种子植物区系; 分布区类型; QX—程序

区域性种子植物区系的研究, 一般是对该区域内种子植物科、属、种的组成, 分布状况, 生物学特性, 以及反映这一植物区系自然条件及地质地理变迁历史的研究, 并在与有关的其他地区植物区系比较和分析的基础上, 研究该区系的基本特点及发生、发展和演变的规律。目前, 我国在区域性种子植物区系的研究工作中, 主要是依据吴征镒教授^[1]对中国产全部属按原始类型或现代分布中心划分的 15 个分布区类型及亚型, 对某地区的种子植物在属级水平上进行统计和分析。尽管很多研究人员提出, 要进一步在种级水平上进行深入的研究工作, 但就某一个地区的种子植物区系在属级水平的研究, 仍是十分必要和可行的, 这种方法对认识某个区系的性质及与其它区系的关系方面起到很重要的作用。由于研究一个或多个区系时, 涉及的原始资料数目很大, 计算量相当繁重, 因而, 在一定程度上, 限制了研究人员利用原始资料所能提供的更多信息, 进行深入研究和分析。另外, 有些作者在进行多区系的研究时, 将不同区系中各个分布区类型属(种)的绝对数字及有关的相对数字直接加以比较研究, 用于确定区系间的相似或关联程度, 甚至用于推测区系间的亲缘关系, 但由于这些数字仅代表了各种分布类型类群在当地的数目, 既不反映各种分布类型内属的组成, 也不反映种的组成, 同时又强烈受各区系面积的影响, 因此在表达不同

* 覃海宁、李振宇、李良千、杨亲二、李永红等同志参与大量原始资料核对工作, 特此致谢。

本文使用程序的开发和编写者为李奕。

本文介绍的各程序均存中国科学院植物研究所系统与进化植物学开放研究实验室, 并向有关研究人员提供服务。

区系间的关系方面,误差是很大的,甚至是完全错误的,因此,植物区系的定量研究中,应建立合适的方法和指标。工欲善其事,必先利其器,我们在学习和研究吴征镒教授植物区系研究思想的基础上,针对中国局部地区区域性种子植物区系的研究,设计和编制成电子计算机实用程序(QX—系列)。试用结果表明,该程序对摆脱繁琐的劳动,提高研究水平是非常有效的。本程序按吴征镒教授最新发表中国产全部属的分布区类型资料(根据最近期刊或新版《中国植物志》等资料,略有调整),可对某地区种子植物区系资料进行统计运算和分析,为研究人员深入研究,提供更多素材和方便。

本程序包括一系列统计运算程序及辅助程序,以保证在资料更新、数据改变等情况下,迅速向研究人员提供统计分析结果。本程序根据研究人员提供某区系植物组成的原始资料,以表格或名录形式输出如下结果:

1. **某地区种子植物区系类型统计分析表(QX01—表)**。 根据研究人员提供中国某地区种子植物名录,在属级水平上提供关于该区系的两个统计分析表。表1为包括广布属在内的统计表,表2为扣除广布属影响的统计表。上述2表中,均有对热带类型和温带类型的小结及总和统计(见附录3—1,表1—2)。同时,各表中也计算出用于区系性质分析和与其它区系比较研究用的区系指数(F)和种系分化指数(ID),对各分布区类型的计算公式完全一样,属(种)区系指数的公式为: $Fg(s) = (LG(s)/LG(s)T)/(CG(s)/CG(s)T)$,种系分化指数的公式为: $IDc = CS/CG$ (关于这两个指数的意义,见博德志、左家嘴,另文发表)

2. **某地区种子植物区系科属统计表(QX02—表)**。 根据同表1、2的原始资料,统计分析了该区系科属组成及与中国和世界有关的相应资料,表3为按科排列(本文使用的科号为恩格勒系统,见中国科学院植物研究所标本室科号表)。表4为按属大小排列,表5为按种数多少排列(表3—5式样见附录3—2)。同时,该项统计表可根据需要提供该地部分或全部属按字母,科号,及分布区类型等资料分别排序或组合排序的名录。

3. **两区系比较分析表(QX03—表)**。 本表根据研究人员提供两个地区种子植物区系原始资料,分别求算两个地区之间根据共有属总数计算的总体相似性系数,并同时分别求算出在各分布区类型内共有属数目及其相似性系数,本程序使用的相似性计算公式为较为常见的Jaccard系数,计算公式见表6的附注。同时,根据上述资料,亦可求出在两地区共有属基础上构建的区系指数 FG_{m-n} 。上述的结果形式见附录3—3,表6—7。本程序在提供表6的基础上,还可提供两区系共有属按各种方式排序的名录。种子植物区系组成相似性的研究和分析的问题很多,定量分析方法本身也将是今后值得注意的问题之一。

经初步使用表明,上述中国种子植物区系研究程序,对解除繁琐的计算,节约时间,提高研究水平有一定效果,但对于进行中国植物区系研究和建立中国植物区系学派的工作,该程序仅仅提供了一个有效工具,决定研究水平高低,在于研究者本身而不在于武器。这一点,应尤为本方法使用者所注意。

参 考 文 献

- [1] 吴征镒, 中国种子植物属分布区类型. 云南植物研究, 1991: 增刊 IV, 1--139.

附录 1. 本程序使用的数据库(主库)结构

字段	字段名	类型	宽度	小数
1	E 码*	数字型	4	
2	H 码*	数字型	4	
3	类群	字符型	1	
4	拉丁属名	字符型	45	
5	中文属名	字符型	16	
6	本地种数	数字型	4	
7	中国种数	数字型	4	
8	世界种数	数字型	4	
9	分布区类型	数字型	3	
10	亚型	数字型	1	

* E 码, H 码分别为恩格勒系统和哈钦松系统的科号, 参见中国科学院植物研究所植物标本馆科号表。

附录 2. QX—系列程序工作条件:

1. 使用 IBM286—486 个人电子计算机。
2. 硬盘应保留 2M 以上空间, 内存 1M。
3. 使用 DOS 3.3 和 FOXBASE V2.1 软件。
4. 与微机相匹配的打印机。

附录 3. 为节省篇幅, 本文仅介绍 QX—系列程序的各种输出结果形式, 各有关程序在中国科学院植物研究所系统与进化植物学开放研究实验室。为节省篇幅, 各种输出表格或目录, 一般只列出代表形式。

附录 3—1. 分布区类型统计程序(QX01.prg)及 QX01 子程序(PR02.prg)运行 QX01 程序时, 必须有此子程序。本程序输出形式见表 1—2。

表 1 (QX01—1 表)某地区种子植物区系分析统计计算结果(未修正)

分布区类型	1	2	T(2—7)	8	T(8—14)	15	Total
当地属数(LG)	49	102	246	108	271	21	587
当地种数(LS)	157	279	529	326	638	21	1345
中国属数(CG)	107	363	1523	299	1225	241	3096
中国种数(CS)	3430	3966	9801	7524	12784	472	26487
世界种数(WS)	28586	49613	89620	30474	48434	480	165120
LG/CG	0.45794	0.28099	0.16152	0.36120	0.22122	0.08714	0.18960
LS/CS	0.04577	0.07035	0.05397	0.04338	0.04991	0.04449	0.05078
CS/WS	0.12902	0.07994	0.10936	0.24690	0.26395	0.98333	0.16041
LG/LGT	0.08348	0.17376	0.41908	0.18399	0.46167	0.03578	1.00000
LS/LST	0.11673	0.20743	0.39331	0.24238	0.47435	0.01561	1.00000
CG/CGT	0.03456	0.11725	0.49193	0.09658	0.39567	0.07784	1.00000
CS/CST	0.12950	0.14973	0.37003	0.28406	0.48265	0.01782	1.00000
ID1	3.20408	2.73529	2.15041	3.01852	2.35424	1.00000	2.29131
IDc	32.05607	10.92562	26.43533	25.16388	10.43592	1.95851	8.55523
Fg	2.41532	1.48203	0.85192	1.90509	1.16680	0.45958	1.00000
Fs	0.90140	1.38536	1.06291	0.85326	0.98280	0.87617	1.00000

表 2 (QX01-2 表) 某地区种子植物区系分析统计计算结果(修正)

分布区类型	2	T(2-7)	8	T(8-14)	15	Total
当地属数(LG)	102	246	108	271	21	538
当地种数(LS)	279	529	326	638	21	1188
中国属数(CG)	363	1523	299	1225	241	2989
中国种数(CS)	3966	9801	7524	12784	472	23057
世界种数(WS)	49613	89620	30474	48434	480	138534
LG/CG	0.28099	0.16152	0.36120	0.22122	0.08714	0.17999
LS/CS	0.07035	0.05397	0.04333	0.04991	0.04449	0.05152
CS/WS	0.07994	0.10936	0.24690	0.26395	0.98333	0.16644
LG/LGT	0.18959	0.45725	0.20074	0.50372	0.03903	1.00000
LS/LST	0.23485	0.44529	0.27441	0.53704	0.01768	1.00000
CG/CGT	0.12145	0.50953	0.10003	0.40984	0.08063	1.00000
CS/CST	0.17201	0.42508	0.32632	0.55445	0.02047	1.00000
IDc	2.73529	2.15041	3.01852	2.35424	1.00000	2.20818
IDI	10.92562	6.43533	25.16388	10.43592	1.95851	7.71395
Fg	1.56112	0.89739	2.00676	1.22907	0.48411	1.00000
Fs	1.36533	1.04754	0.84092	0.96859	0.86350	1.00000

* 上述表中有关全国植物属的数据主要根据吴征镒最新发表资料,上表中 IDI=LS/LG, IDc=CS/CG, Fg=(LG/LGT)/(CG/CGT), Fs=(LS/LST)/(CS/CST)

附表 3-2. 科属统计分析程序(QX02.prg). 本程序输出形式见表 3-5.

表 3 (QX02-1 表) 按各科顺序排序

科号	LG	LS	CG	CS	WS	LG-CG	LS-CS	CS-WS	LG-LGT	LS-LST
1	1	2	1	8	17	1.0000	0.2500	0.4706	0.0018	0.0013
2	1	1	1	1	1	1.0000	1.0000	1.0000	0.0018	0.0007
3	1	2	4	12	22	0.2500	0.1667	0.5455	0.0018	0.0013
6	1	2	1	7	9	1.0000	0.2857	0.7778	0.0018	0.0013
7	3	5	10	89	237	0.3000	0.0562	0.3755	0.0053	0.0033
8	2	2	5	7	8	0.4000	0.2857	0.8750	0.0035	0.0013
13	3	3	3	1	5	1.0000	0.7500	0.8000	0.0053	0.0020
14	2	2	4	71	2011	0.5000	0.0282	0.0353	0.0035	0.0013
:										
296	2	3	11	112	1114	0.1818	0.0211	0.1275	0.0035	0.0020
299	1	2	1	80	250	1.0000	0.0250	0.3200	0.0018	0.0013
300	2	3	3	42	402	0.6667	0.0714	0.1015	0.0035	0.0020
302	2	3	22	145	1111	0.0909	0.0207	0.1305	0.0035	0.0020
306	15	23	171	1039	7959	0.0877	0.0221	0.1305	0.0265	0.0150

表 4 (QX2-2 表) 按各科内属数多少排序

科号	LG	LS	CG	CS	WS	LG-CG	LS-CS	CS-WS	LG-LGT	LS-LST
2	1	1	1	1	1	1.0000	1.0000	1.0000	0.0011	0.0004
49	1	1	1	1	6	1.0000	1.0000	0.1667	0.0011	0.0004
76	1	1	1	1	1	1.0000	1.0000	1.0000	0.0011	0.0004
137	1	1	25	55	688	0.0400	0.0182	0.0799	0.0011	0.0004
:										
144	2	14	2	62	320	1.0000	0.2258	0.1938	0.0023	0.0056
138	2	17	2	76	163	1.0000	0.2237	0.4663	0.0023	0.0068
:										
187	3	6	8	86	185	0.3750	0.0698	0.4649	0.0034	0.0024
198	3	7	10	91	1780	0.3000	0.0769	0.0511	0.0034	0.0028
:										
261	44	87	223	2254	17954	0.1973	0.0386	0.1255	0.0496	0.0349
273	58	109	238	1300	7695	0.2437	0.0838	0.1689	0.0654	0.0437

表5 (QX02-3表) 按各科内种数多少排序

科号	LG	LS	CG	CS	WS	LG-CG	LS-CS	CS-WS	LG-LGT	LS-LST
2	1	1	1	1	1	1.0000	1.0000	1.0000	0.0011	0.0004
49	1	1	1	1	6	1.0000	1.0000	0.1667	0.0011	0.0004
40	1	2	2	20	82	0.5000	0.1000	0.2439	0.0011	0.0008
218	1	2	11	27	711	0.0909	0.0741	0.0380	0.0011	0.0008
48	2	3	3	7	285	0.6667	0.4286	0.0246	0.0023	0.0012
30	2	3	5	9	114	0.4000	0.3333	0.0789	0.0023	0.0012
94	24	107	53	1020	5100	0.4528	0.1049	0.2000	0.0271	0.0429
273	58	109	238	1300	7695	0.2437	0.0838	0.1689	0.0654	0.0437

附录3-3. 两区系共有属统计分析程序(QX03.prg)。运行QX03程序时,请先准备好主库(即附录1)的数据库和由各地区原始资料建立的子库,子库结构和主库完全相同,是把各地区原始资料分别输入到主库后,根据本地种数<>0的条件分别建立的数据库。本程序输出形式见表6-7。

表6 (QX03-表)两区系相似性分析表

分布区类型	1	2	3	14	15	Total
中国属数 CG	107	362	68		303	240	3084
甲地属数 XG	64	128	16		107	37	800
乙地属数 YG	57	123	13		108	33	746
甲乙共有属 ZG	50	93	11		70	19	550
ZG/CG	0.4673	0.2569	0.1618		0.2310	0.0792	0.1783
S _{xy}	0.7042	0.5886	0.6111		0.4828	0.3725	0.5522

* $S_{xy} = ZG / (XG + YG - ZG)$ (Jaccard similarity coefficient), X, 甲地, Y, 乙地, Z, 甲乙共有, S_{xy} : 甲乙两区系相似性系数

表7. 两地区共有属名单。(略)

附录3-4. 辅助程序(XQ00.prg)。本程序用于将各子库资料分别输入到主库更新主库数据。

文中使用符号说明:

G(genus); 属; S(species); 种; L(local); 当地; C(China); 中国; W(world); 世界; T(total); 总和; LG: 当地某类型属之和; LS: 当地某类型种之和; CG: 中国某类型属之和; CS: 中国某类型种之和; WS: 中国产属中世界所产的种数; LGT: 当地产全部属之总和; LST: 当地产全部种之总和; CGT: 中国全部属之总和; CST: 中国全部种之总和; LG(S)/CG(S): 某地某类型全部属(种)占全国同类型属(种)之比例; LG(S)/LG(S)T: 某地某类型全部属(种)占该地全部属种总数之比例; CS/WS: 中国某类型各属的种之和与各属世界种之和的比例。

QUANTITATIVE STUDIES ON THE SEED PLANT FLORA OF CHINA
I. COMPUTER PROGRAMS FOR THE STATISTICAL
ANALYSIS OF THE SEED PLANT FLORA

Fu Dezhi, Li Yi and Xue Yanhong

(*Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Academia Sinica, Beijing 100093.*)

Abstract

In the present paper, practical programs for the quantitative study of the seed plant flora of any a given region of China are worked out according to the areal-type classification of seed plants of China proposed by Professor C. Y. Wu. By use of these programs, the statistical data of the seed plant areal-types, composition of families, genera and species of any a given seed plant flora as well as the Jaccard's similarity coefficient of genera of this flora with another one can be readily and exactly obtained.

Key words: Seed plant flora; Areal-type classification; QX program