

新疆野核桃种质资源数量分类研究

王磊 李霞 杨辽

张汉斐

林培均 许正

(新疆农业大学林学院·乌鲁木齐)

(新疆巩留县农技站)

(新疆伊犁地区园艺所)

摘要: 本文应用聚类分析和主座标分析方法对新疆野核桃 73 个样本的 59 个性状观测结果进行了计算机分析。结果表明: 聚类分析得到的树状图和主座标分析的样本座标排列图直观, 明确地给出了新疆野核桃种下类型间的距离及遗传差异: 它们可分为 14 个类型并可归并为 4 个类群、2 大分枝。类型间、类群间的遗传差异大于类型内、类群内样本间的遗传差异, 体现了新疆野核桃自然演化的阶段性及丰富的变异类型。也进一步证明了对新疆野核桃用经典形态分类学的研究结果^[1]是正确的, 为野核桃种质资源研究及育种材料的选择提供参考。

关键词: 新疆野核桃 种质资源 聚类分析

主座标分析

新疆野核桃是栽培核桃的直系祖先, 具有许多优良的遗传特性, 在形态和品质上与栽培核桃及其相似, 它们既可直接利用又可作为育种材料。在研究栽培核桃的起源、演化上具有重要价值。张剑等曾在 60 年代按新疆野生核桃果实的大小, 形状将其分为 3 类, 但没有从树体的各器官的植物学特征及生物学特征上进行综合分类。为了进一步利用及研究, 作者曾于 1985~1989 年在新疆巩留县野核桃沟中进行了五年实地调查, 对其中 150 株典型树的植物学特征、生物学特征等进行系统定点观测。根据观测结果初步进行了种下分类, 大致分为 14 个类型^[1]。为了全面、科学地探讨新疆野核桃种质资源种下分类及自然演化规律, 本文应用聚类分析及主座标分析方法就野核桃的种下分类进行了探讨, 并与^[1]中分类结果进行了比较和分析, 为野生核桃种质资源研究及育种工作者提供依据。

1. 材料和方法

对新疆巩留县野核桃林中有代表性的 150 余株样本的 100 余个性状进行了五年系统定点观测, 选择其中 73 株典型样本的 59 个性状(数量与非数量性状)输入计算机进行聚类分析和主座标分析。

2. 数据处理与计算

2.1 数据标准化

对新疆野核桃 73 株典型样本的 59 个性状按公式: $y_{ki} = \frac{x_{ki} - x_{kmin}}{x_{kmax} - x_{kmin}}$ (其中 $k=1, 2, \dots, 73$; $i=1, 2, \dots, 59$) 进行标准

化, 得到标准化数据阵(限于篇幅省略)。

2.2 聚类分析^[3]

对标准化数据阵以其数量性状为指标求各样本间的欧氏距离得到样本距离矩阵 C_{ij} , 其中 $d_{ij} = \left[\sum_{k=1}^n (x_{ik} - x_{jk})^2 \right]^{1/2}$, $k=1, 2, \dots, 59$; 然后使用中间距离法聚类: $d_{ij} = \left(\frac{1}{2} D_k^2 + \frac{1}{2} D_r^2 + \beta D_{kr}^2 \right)^{1/2}$ 其中 $\beta = -\frac{1}{4}$ 根据聚类结果画出聚类树状图。

2.3 主座标分析^[3]

为了有效利用部分定性因子, 直观表达样本间距离, 本研究采用主座标分析法进行计算。其步骤为:

2.3.1 以样本距离阵 M 为原始数据求距阵 A

$$a_{ij} = \frac{1}{2} \left(\frac{1}{m} \sum_{j=1}^m m_{ij}^2 + \frac{1}{n} \sum_{i=1}^n m_{ij}^2 - m_{ij} - \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m m_{ij}^2 \right)$$

即 a_{ij} 等于第 i 行平均数加第 j 列平均数减 m_{ij} 再减去总平均数所得数的一半。

2.3.2 求 A 的非零特征根 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ 和单位特征向量 V_1, V_2, \dots, V_m 则第 j 号样本的主座标为

$$\begin{pmatrix} C_{1j} \\ C_{2j} \\ \vdots \\ C_{mj} \end{pmatrix} = \begin{pmatrix} \sqrt{\lambda_1} V_{1j} \\ \sqrt{\lambda_2} V_{2j} \\ \vdots \\ \sqrt{\lambda_m} V_{mj} \end{pmatrix}, \text{ 株}$$

根据如上计算得到前 6 个主成份特征根及贡献率表和样本点在前两个座标上的投影图即样本座标排列图。

3. 结果与讨论

属国家自然科学基金资助课题, 参加野外调查工作的还有赵永生等同志。

现为南京林业大学硕士学位课程学习班学生。

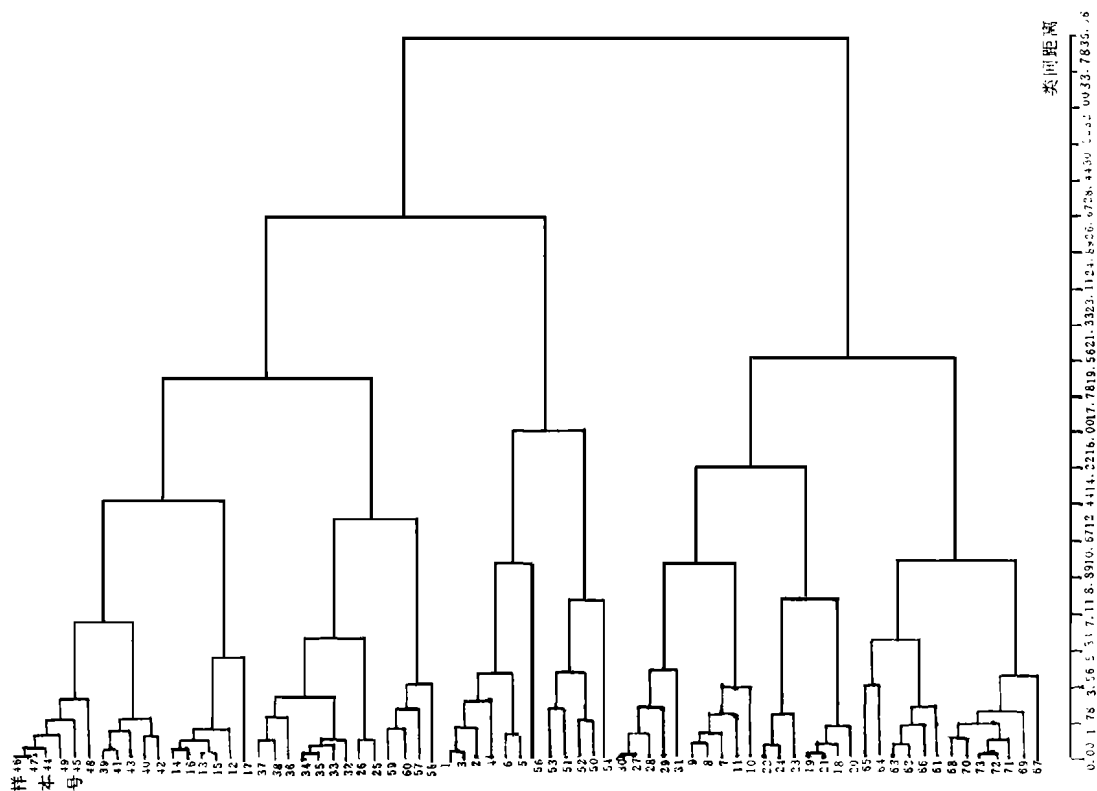


图 1 系统聚类图

3.1 聚类分析是根据材料间相似进行归类合并的研究方法。聚类分析结果得到树状图（图 1）上所有野核桃样本联结成树状，其树根在上，树梢向下，随着树枝向下伸长扩展，野生核桃的分化类型增多。愈向树梢同一枝上的分枝间同质性增强；愈向根部，相邻枝（类）间的差异愈大。图 1 给出了新疆野核桃 14 个类型间的距离及遗传差异，其种下类型的分类结果和〔1〕中的分类是完全一致的（见表 1）。这不仅证明经典形态分类学的研究结果是正确的，而且证明研究分类与演化的性状指标的选取是合理的。

3.2 从聚类分析结果树状图（图 1 可知），新疆野核桃的 14 个种下类型是由果实形态及树体各部的综合因子作用而形成的，它们可归并为具有一定特性差异的 4 个类群和 2 大分枝。但从薄壳与铁壳核桃的距离较近这一点可以说明虽然它们在内果皮厚薄及软硬上差异较大，但只是同一类群内的不同类型而已，不能作

为分类群的主要特征。由于野核桃长期的自然杂交和实生繁殖产生了丰富的变异类型，其类型间、类群间的遗传差异大于类型内及类群内的差异，这对于研究栽培核桃的演化和以野核桃作育种材料时亲本的选择都是有意义的。

3.3 用主座标分析法得到的样本主座标排列图（图 2），直观地表现了各样本点在第一、第二座标系中的位置及点间距离。根据各样本点间距离的远近可大致归并为 14 个种下类型（例如 18—21 为一类；25、26 为一类；39—43 为一类等等）其中心型（32—38）与三棱型（57—60）距离很近，这说明这两个类型尽管在心皮数上（2 心皮与 3 心皮）及果形上（圆形与心形）差异较大，但它们的综合性状在第一、第二主座标上的差异不明显。

3.4 从样本主座标排列图（图 2）各样本点的聚集程度看，新疆野核桃也可大致分为 4 个类群，这与聚类

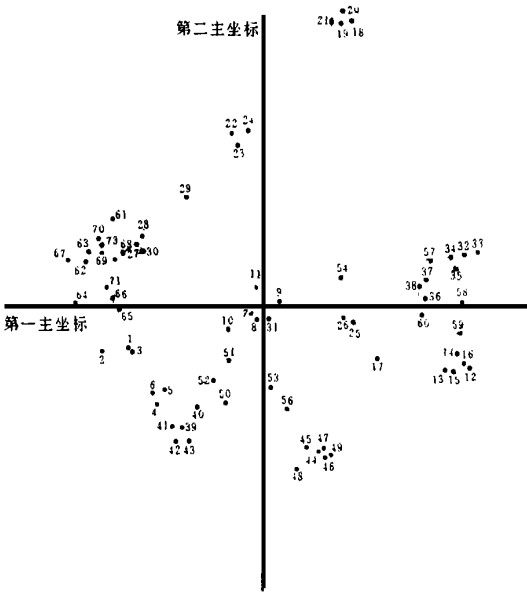


图 2 样本坐标排列图

树状图的方向是一致的。

表 1. 新疆野核桃样本号及其原种下类型表

样本号	原类型号	原类型	样本号	原类型号	原类型
1—6、56	1	平底圆型	32—38	8	心型
7—11	10	小矩圆型	39—43	9	纺锤型
12—17	11	小圆型	44—49	13	矩圆型
18—21	12	铁壳型	50—54	14	卵圆型
22—24	6	尖嘴型	57—60	2	三棱型
25—26	5	歪嘴型	61—66	4	长薄壳型
27—31	7	尖果型	67—73	3	薄壳型

表 2 主座标分析特征根及贡献

主成分	特征根 λ_i	贡献率 (%) $\lambda_i / \sum \lambda_i$	累积贡献率 (%)
1	67.4	23	23
2	37.07	13	36
3	30.15	10	46
4	24.36	8	54
5	18.05	6	60
6	16.08	5	65

3.5 由主座标分析得到前六个主成份的特征根累积贡献率(表 2)。从表 2 可知各主成份的特征根和累积贡献率都不是太大, 前两个主座标累积贡献率为 36%, 但各样点在两个座标间的分布图上的聚集情况却又和聚类树状图完全相吻合, 这说明此分布图基本符合实际情况。而主成份贡献的大小同所用性状数目反相关, 即所用性状增多每一主成份的贡献变小。本研究各主成份贡献率的大小与每一样本的 59 个性状有关。如何在众多的性状因子中选取有代表性的性状, 还需要对各性状因子进一步筛选。从贡献率大小来看, 这些二维标记只代表了部分器官的变异, 不能反映整体差异, 据此我们认为性状的全面性比信息量更为重要, 性状增多, 绝对信息量必然大从而在一定程度上保证了二维标记的可靠性。

3.6 无论使用主座标分析法还是聚类分析法, 总有个别点在类型内和其它点距离较大如 39、10、54、617 等等。这说明这几个样本和其它样本是有一定差异的, 但是否能划分为一个新的类型? 或者是一新的变异分枝? 还有待研究。

参考文献

[1]王磊、崔乃然、张汉斐等《新疆野核桃的研究》干旱区研究 1997·1·17—27
 [2]张鲁刚、王鸣《甜瓜种质资源 Q 型聚类分析及主成分分析》中国西瓜甜瓜, 1990·1·14—19
 [3]唐守正 多元统计分析方法, 中国林业出版社 1989
 [4] Herman, P·P·J, P·J·Robbertse and N·Groffelaar, 1987 A numerical analysis of morphology of the leaves of some southern African Pavetta species s·Abr·J· Bot 53 (1): 53—60.