

分子系统学研究中分子位点数与遗传差异 信息可靠性的关系^①

周泽扬 夏庆友 鲁成 冯丽春 向仲怀

(西南农业大学蚕桑学农业部部级重点实验室, 重庆 400716)

摘要 用桑蚕(包括不同生态地理品种 59 个及野桑蚕)和桑属植物(包括 12 个种和 2 个变种)两类截然不同的材料,探讨了分子系统学研究中 RAPD 分子位点数与遗传差异信息可靠性间的关系。发现:(1)所分析的位点数多少与所能提供系统信息的量及可靠性之间有明显的关系;(2)当位点数在 20 以下时,得到的遗传差异的结果极不可靠,随着位点数的增加所提供的信息量及可靠性增加。当位点数超过 70 个时,所提供的信息可靠性趋于稳定;(3)对桑蚕和桑属植物两种截然不同的材料的分析,均得相似结果。由此推论:在用 RAPD 进行生物系统学研究中,以分析 70 个左右位点数为好;这一结论受研究对象的影响小,在其它类似的研究中或许具有一定的参考价值。

关键词 分子系统学, RAPD 位点数, 遗传差异, 蚕, 桑

中图分类号 Q38

The Correlation Between the Number of RAPD-loci and the Reliability of the Information on Genetic Variation in Molecular Phylogenetic Studies

ZHOU Zeyang XIA Qingyou LU Cheng FENG Lichun XIANG Zhonghuai

(The Key Sericultural Laboratory of Agricultural Ministry, Southwest Agricultural University, Chongqing 400716)

Abstract We studied the correlation between the number of RAPD-loci investigated and the reliability of the information on genetic variation analysis using quite different two-kind materials, silkworms (including 59 geographic varieties of *Bombyx mori* belonging to different ecotype, as well as *Bombyx mandarina* M.) and mulberries (including 12 species and 2 varieties). The results showed as follows: (1) There was obvious correlation between the numbers of RAPD-loci investigated and the reliability of the information on genetic variation analysis. (2) The data were not suitable for the genetic variation analysis when the number of RAPD-loci investigated was smaller than 20. The more RAPD loci were analyzed, the more accurate genetic diversity would be identified when the number of the locus was varied from 20 to 70. Furthermore, the stable information to be available is shown when the number exceeded 70. (3) The similar regularities mentioned above were observed in both silkworm and mulberry. In order to get reliable information of genetic variation in molecular phylogenetic studies, at least 70 RAPD loci would be investigated. And the results obtained in this study might be referential for AFLP and RFLP analysis.

Key words molecular phylogeny, RAPD loci, Genetic variation, Silkworm, Mulberry

^①国家自然科学基金(39570557)和博士点基金(950604)资助的项目。

在一项系统学研究中, 至少需要分析多少个性状(标记), 一直是一大难题。至今能够普遍接受的, 仍是愈多愈好。这在理论上当然是无可非议的, 但实践上则难度很大。某些性状(如形态特征等)客观上存在位点数有限的问题; 而另一些(如象 RAPD 等的位点数等)虽然理论上几乎是无限的, 但实际研究中却受研究规模、经费和时间的限制, 其可分析和利用的位点数也是有限的。以利用 RAPD 进行的研究为例, 在 Apostol 等^[4]、Castiglion 等^[3]的研究中, 分别分析了 40 个和 120 个位点; 而 Link 等^[6]、Jain 等^[5]则分别分析了 365 个和 500 个位点。这反映出在分子系统生物学研究中, 标记(位点)的数量选用存在着相当的不规范性。这种状况严重影响着研究结果的可靠性和可比性。

本文以家蚕(*Bombyx mori*)和桑属植物(*Morus*. L.)两种不同的物种为材料, 对 RAPD 分析中分子位点数与系统遗传差异信息可靠性的关系进行了探讨, 发现 RAPD 的位点数与所提供的系统信息量之间有明显的关系, 并对分子系统学研究中所需 RAPD 位点数作了分析探讨。

1 材 料 和 方 法

1.1 材料

家蚕材料取自西南农业大学家蚕基因库保存的典型地理生态代表品种 59 个以及重庆北碚蚕种场采集的野桑蚕(*Bombyx mandarina* M.), 计 60 种材料^[8]; 桑属植物材料采自西南农业大学桑树资源圃, 每一个种或一个变种选取一个代表性的材料, 共 12 个种及 2 个变种, 计 14 个研究材料^[9]。

1.3 模板 DNA 制备及 PCR 扩增

家蚕基因组 DNA 的制备和 PCR 扩增按夏庆友等的方法^[8]进行。桑属材料 DNA 按向仲怀等^[2]的方法制备并按冯丽春等的方法^[7]进行扩增。

1.5 实验数据分析处理

对于任何 RAPD 位点数 n , 任意两个品种 i 和 j 间的单匹配相似系数(Simple matching coefficients of similarity, SM 系数)^[11]:

$$SMn_{(ij)} = (a + d) / (a + b + c + d) \quad (1)$$

其中 a 为第 i 和第 j 品种间皆为“1”(有带)的位点数, d 为皆为“0”(无带)的位点数; b 和 c 为两品种的交错匹配带数(一个为 1 一个为 0)。

前期研究^[8]已知, 34 个引物对家蚕扩增出的总带数为 103 条(103 个位点)。对于任意 RAPD 位点数 n 和品种数 k , 所有品种组合的 SM 系数与其在位点数为 103 时的 SM 系数之差的平方和为:

$$SS(n, k) = \sum \sum (SM_{103}(ij) - SMn(ij))^2 \quad (2)$$

其中 n 为位点数; k 为品种数; $SM_{103}(ij)$ 为位点数等于 103 时, 第 i 和第 j 品种间的 SM 系数; $SMn(ij)$ 为位点数为 n 时, 第 i 和第 j 品种间的 SM 系数。

对于桑属材料, 前期研究^[9]得知 20 个引物能扩增出 237 条带, 亦按上述同样的方法进行实验数据处理。

2 结 果 与 讨 论

2.1 家蚕不同地理品种及野蚕间 RAPD 位点数与 SM 相似系数稳定性的关系

本研究共用 60 个蚕品种材料, 所有可能的两两材料组合数为 1 770 个。首先, 按位点数为 103 个时的 SM 系数值大小进行排序, 分别选取 SM 系数最大的、居中的和最小的组合中各 1 个品种组合(表 1), 分别根据式(1)计算位点数 n 从 1 到 103 时 $SMn(ij)$ 的值, 并相对于位点数作图, 得到如图 1 的结果。由图可见, 随 RAPD 位点数的不同, SM 的变化明显分为 3 个阶段: a: 剧烈变化阶段: 当位点数为 1 到 20 时, SM 值上下起伏甚大, 极不稳定; b: 缓慢变化阶段: 当位点数在 20 到 70 左右之间, SM 值随着位点数的不同有一定的波动; c: 趋于平稳阶段: 当位点数超过 70 左右, SM 值波动渐趋稳定。

表 1 三个典型蚕品种组合的 SM 及其在整个组合中的排序

品种名及组合	SM 103	按 SM 大小排列的位顺
野蚕——中江土种	0.563	1766
柘蚕(湖)——余杭2化(甲)	0.699	880
临海20——土白	0.874	1

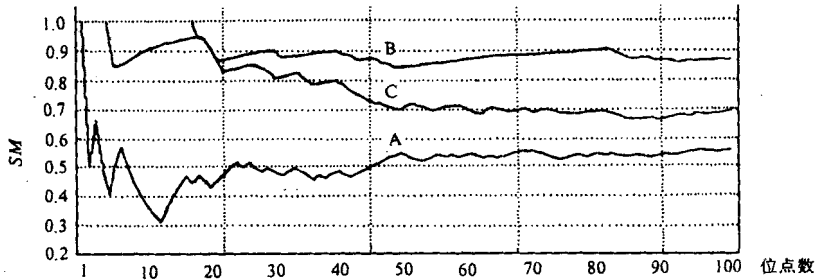


图 1 RAPD 位点数与不同家蚕组合间 SM 值的关系
A. 野蚕——中江土种; B. 临海 20——土白; C. 柘蚕(湖)——余杭 2 化(甲)。

2.2 桑属植物不同种及变种间 RAPD 位点数与 SM 相似系数稳定性的关系

同前述方法, 在 14 个桑属材料中所有两两组合中遗传关系最近者鲁桑与瑞桑($SM=0.870$), 最远者川桑与长穗桑($SM=0.618$)和居中者瑞穗桑与广东桑($SM=0.710$)三个典型的组合, 考察其 RAPD 位点数与 SM 系数值的波动关系, 如图 2 所示, 其 SM 值的稳定性随位点数的关系趋势与前述家蚕材料呈一致的倾向, 即 70 个左右位点数以前的 SM 剧烈变化, 稳定性差, 70 个左右以后趋于平稳。

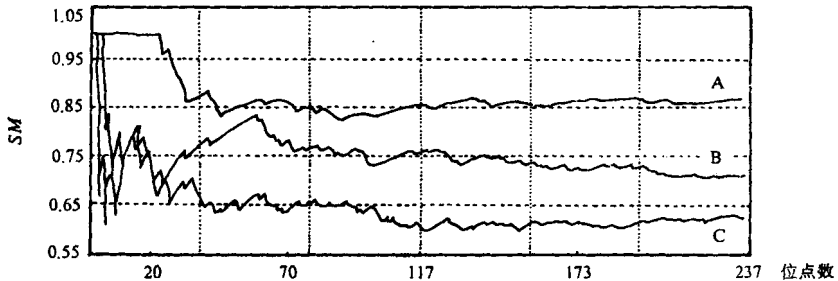


图 2 RAPD 位点数与不同桑属植物组合间 SM 的关系
A. 鲁桑与瑞穗桑; B. 广东桑与瑞穗桑; C. 川桑与长穗条。

2.3 RAPD 位点数与 SM 的方差值稳定性间的关系

以上就独立实例的情况进行了讨论。这里再就 SM 系数的总体变化趋势进行考查。如果将基因组 DNA 整个区域看成若干位点构成的“位点群体”, 那么一个 RAPD 位点就相当于这个“位点群体”的随机样本。于是, 采用类似于统计学上的随机抽样调查方法来研究 SM 系数的总体变化: a. 用所有位点 SM 系数的平均值作为参照; b. 用不同位点数 SM 系数的方差之和表示 SM 值逼近某个固定值的程度, 利用按式(2)计算所得的方差 $SS(n, k)$ 相对于位点数 n 作图, 得到如图 3 和图 4 的结果。如图可见, 无论是家蚕材料还是桑属材料, 其 SM 系数方差的变化均可分为 3 个阶段, 即位点数 1~20 的剧烈变化区, 位点数 20~70 左右的缓慢变化区和位点数 70 左右以后的平稳区。并发现即使研究材料数不同(家蚕材料数 k 分别为 10、20、40 和 60 四种情况, 桑树材料数分别为 2、5、10 和 14 四种情况), 其变化规律和趋势亦是一致的(图略), 由此表明, SM 值稳定性与位点数的关系, 无论在总体上还是在具体的材料组合, 均有同样表现。

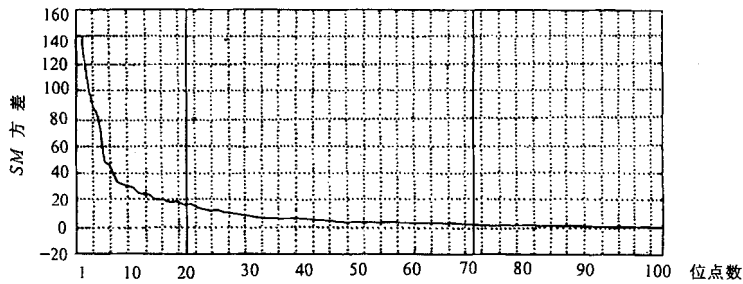


图 3 60 个蚕品种数场合下, RAPD 位点数与 SM 方差的关系

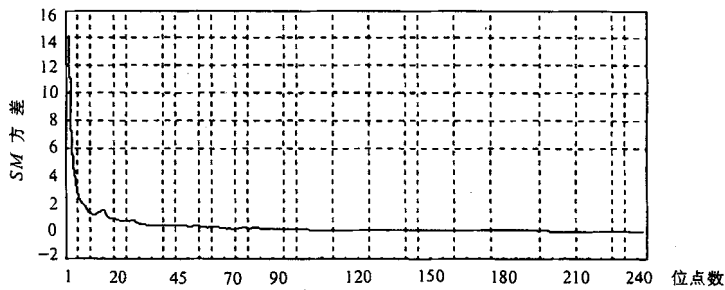


图 4 14 种桑属材料数场合下 RAPD 位点数与 SM 方差的关系

在分子系统学研究中, 所分析的位点数(性状)多少, 是关系到研究可行性和结果可靠性的至关重要的问题。究竟至少应该分析多少位点数所获得的信息才较为可靠? 本研究的结果表明: (1) RAPD 位点数与所提供的信息质量之间有明显关系, 位点数越多, 提供的信息越准确; (2) 当所调查的 RAPD 位点数在 20 以下时得到的信息极不可靠; 而当调查的位点数达到或超过 70 时, 即能获得比较稳定可靠的信息。

在本研究中, 家蚕和桑属植物两类截然不同材料均得到基本一致的结果, 并且由于 RAPD 引物碱基序列是随机的, 暗示本研究所得到的结论受研究对象的影响小。但是否有普适性则要考虑基因组的大小。因为基因组 DNA 越大, 就需要更多的位点才能覆盖到 DNA 的大多数区域, 需要的最少位点数也许就相应越多。

另一方面, RAPD、RFLP 和 AFLP 等分析技术, 它们的分析过程虽各不相同, 但原理上均是分析 DNA 链上碱基序列的差异, 因此, 本文所获得的结果对于 RAPD 以外的分子标记技术或许也有参考价值。

参 考 文 献

- 1 Sokal R R *et al.* A statistical method for evaluating systematic relationships. *Univ. Kan. Sci. Bull.*, 1958, 38: 1409~1438
- 2 向仲怀等. 采用随机扩增多态性 DNA 技术 (RAPD) 在桑属植物系统学中的应用初报. *蚕业科学*, 1996, 21(4): 203~208
- 3 Castiglioni S *et al.* RAPD fingerprints for identification and fortaxonomic studies of elite poplar (*Populus spp*) cloned. *Theor Appl. Genet.*, 1980, 32: 314~331
- 4 Apostol B L *et al.* Estimation of the number of full sibling families at an oviposition site using RAPD-PCR markers applications to the mosquito. *Theor Appl. Genet.*, 1993, 86: 991~1000
- 5 Jain A *et al.* Potential use of random amplified polymorphic DNA (RAPD) technique to study the genetic diversity in Indian mustard and its relationships to heterosis. *Theor Appl. Genet.*, 1994, 88: 116~122
- 6 Link W *et al.* Genetic diversity in European and Mediterranean faba bean germplasm revealed by RAPD markers. *Theor Appl. Genet.*, 1995, 90: 27~32
- 7 冯丽春等. 利用 RAPD 对桑属植物种间亲缘关系的研究. *中国农业科学*, 1997, 30(1): 52~56
- 8 夏庆友等. 家蚕分子系统学研究. *昆虫学报*, 1998, 41(1): 32~40

1997-08-04 收稿, 1998-03-09 修回.