遗传多样性的取样策略

金燕 卢宝荣*

(教育部生物多样性与生态工程重点实验室, 复旦大学生物多样性研究所, 上海 200433)

摘要:合理取样是生物多样性有效保护、利用和研究所面临的最基本问题,它在很大程度上受到植物自身的生物学特性、环境条件和取样目的的影响。遗传多样性的取样策略是指对一定地理分布范围内的生物个体取样时,使样本具有代表性和包含尽可能多的遗传变异的最佳取样方法,包括了取样数目(一个给定区域的居群数和一个居群的个体数)以及取样方式。包括“哈迪-温伯格平衡(Hardy-Weinberg Equilibrium)”定律在内的居群遗传学基本原理是研究取样策略的理论基础,在此基础上可以对居群内的取样个体数及应获取的居群数进行理论计算,同时还可以根据物种居群的遗传结构特点和环境条件的异质性来决定取样的方式。因此,应该依据研究对象本身的特点和取样的目的来确定某一特定区域的居群取样数,以及某一居群内的样本数及取样方式。

关键词:样本数,取样方式,遗传变异,居群遗传结构,等位基因频率

中图分类号:Q943.Q16

文献标识码:A

文章编号:1005-0094(2003)02-0155-07

Sampling strategy for genetic diversity

JIN Yan, LU Bao-Rong*

Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, Institute of Biodiversity Science, Fudan University, Shanghai 200433

Abstract: An appropriate sampling method not only serves as a prerequisite for conservation and utilization of biodiversity occurring in nature, but also affects the results of population genetic research. A strategic sampling method for genetic diversity can be defined as the optimum sampling method to achieve representation of the full genetic variation represented in a species or a population within a manageable number of samples. Therefore, strategic sampling must consider two factors: 1) the number of samples (including number of populations in a given area and number of individuals within a given population), and 2) the methods of sampling in relation to the genetic structure of the populations under consideration. The well-known Hardy-Weinberg Equilibrium and other principles of population genetics provide the theoretical bases of sampling strategy design. The theoretical models of sampling based on the Hardy-Weinberg Equilibrium offer a general guideline for the number of individuals to be sampled within a population and the number of populations to be sampled in a given region if there is no detailed genetic and environmental information available. Based on this, approximately 30 individuals from a population could be considered as a recommended size for collection under such conditions. When designing a sampling strategy, the biotic and abiotic factors that can influence the genetic diversity captured in the organisms under study should also be considered. The taxonomic status, life cycle, geographical distribution, breeding system, and environmental heterogeneity of a target species will significantly affect the pattern of genetic variation and consequently the design of sampling strategy. In general, the optimum sampling strategy should be designed on the basis of the biological features of the target species, environmental conditions, and objectives for conservation, utilization, or research.

Key words: number of samples, sampling methods, genetic variation, population genetic structure, allele frequency

基金项目:上海教委上海市优秀学术带头人资助计划(00XD14006)

收稿日期:2002-11-28;接受日期:2003-02-19

作者简介:金燕,女,1980年出生,硕士研究生,研究方向为种群与进化生态学。

* 通讯作者 Author for correspondence. E-mail: brlu@fudan.edu.cn

1 取样策略的概念及其重要性

取样策略(sampling strategy)的问题最早产生于种质资源的野外采集和保护。Allard 于 1970 年提出了最佳取样策略的概念。他强调许多物种内存在着大量的遗传变异和不同的基因型 ,但在资源的采集过程中由于受到人力和物力的限制 ,只能采集到其中的一部分基因型或遗传变异 ,因此存在着如何使一定数量的样本包含尽可能多的有用变异的问题 (Marshall ,1989)。生物多样性的研究同样存在取样策略的问题。与种质资源的野外采集和保护相比 ,生物多样性的取样策略不仅要考虑在某一生物居群或研究空间中如何使获取的样本包含尽可能多的遗传变异 ,而且还要根据研究的目的使获取的样本具有一定的代表性和统计学意义 ,同时还要求该样本的数量应该在可操作的范围内。因此 ,遗传多样性的取样策略是指对一定地理分布范围内的生物个体取样时 ,使样本具有代表性和包含尽可能多的遗传变异的最佳取样方法 ,包括取样数目(一个给定区域的居群数和一个居群的个体数)以及取样方式。

无论是对生物多样性进行研究、保护还是利用 ,首先面临的问题便是如何取样。在种质资源采集和生物多样性保护过程中 ,除了个体数和居群数都很少的一些濒危物种外 ,由于人力和物力的限制 ,通常不可能对每个物种的所有居群和个体都进行采集或保护 ,而只能采集或保护其中的一部分。然而我们又希望在可操作的前提下 ,获取的样本能包含尽量多的变异 ,能够反映该物种或居群在进化过程中形成的特定时空遗传结构式样。如果没有一个合理的取样策略 ,保护的资源和研究的结果便很难客观反映保护和研究对象的遗传多样性。例如 Jennifer *et al.*(1997)对 *Baptisia arachnifera* 的 10 个居群的同工酶研究表明 ,大部分的变异(90%)存在于居群内部 ,但居群之间的变异水平不一致 ,10 个居群的预期杂合度(H_e)在 0.049 和 0.097 之间变动。根据居群间基因分化系数 $G_{st} = 0.096$ 而估算得出 ,两个居群就可以包括这 10 个居群 99% 的变异 ,因而他们根据上述的居群遗传结构参数选择了两个包含最多遗传变异的居群进行保护。

在群体遗传学研究中 ,合理的取样策略是研究的基础 ,关系到所得出的研究结果是否可靠。李昂

等(2000)对细距堇菜(*Viola tenuicornis*)的研究表明 ,从空间分布式样来看 ,样本所覆盖的面积对遗传参数的估算有相当重要的影响 :样本覆盖面积占整个群体面积的比例越小 ,遗传参数估算产生偏差的可能性越大 ,而且样本在群体中所处的空间位置对遗传参数的估算也有影响 ,根据来自面积相同而位置不同的样本估算出的遗传参数会有很大差异。从样本数量上看 ,对大群体取样时样本数应大于 30 才能得到较准确的遗传参数估算值 ,而样本数小于 20 时的估算可能有较大的偏差。Sjögren & Wyöni (1994)通过计算机模拟运算也发现 ,样本数和等位基因位点数过少可能导致检测到的变异偏低或者甚至检测不到变异。因此 ,选择最佳的取样策略在生物多样性的有效保护、准确研究和合理利用方面都具有十分重要的意义。

2 取样策略的理论基础和计算

2.1 居群遗传学原理

物种居群各世代之间稳定的基因和基因型频率是进行合理取样的前提。包括“ 哈迪-温伯格平衡 (Hardy-Weinberg Equilibrium) ”定律在内的居群遗传学原理是研究取样策略的理论基础。“ 哈迪-温伯格平衡 ”定律认为 ,在没有选择、突变、遗传漂变和迁移的情况下 ,在一个很大并随机交配的居群内 ,各等位基因的频率将世代保持恒定 ,或基因型频率至多在第二代以后将达到恒定 ,即保持一个稳定的平衡。这是研究居群遗传最简单和最基本的模型 ,也是取样策略的理论基础。

2.2 取样策略的理论计算

2.2.1 居群内的取样个体数

在不了解居群遗传结构的情况下 ,根据群体遗传学原理可计算出取样数目对样本所包含的等位基因的影响。Ewens(1972)为此建立了中性无限等位模型(neutral infinite allele model):

$$K \approx \theta [\log_e (S + \theta) / \theta] + 0.6 \tag{1}$$

其中 $\theta = 4Nu > 0.1$, $S > 10$, K 指某一位点选择中性的等位基因数目 , N 指居群的有效个体数 , S 指取样个体数 , μ 指该位点的突变率。

此模型强调一定数目的样本所包含的等位基因数和样本数的对数成正比。这个结论是在假设平衡的条件下根据中性等位模型得到的 ,并已被证实也适用于其他情况 ,如杂合模型(heterotic model)、非

平衡状态(non-equilibrium situations)和经验分布(empirical distributions)等。这个模型有两个显著的优点：首先，该模型普遍适用于对具有中性性状的群体进行取样；其次，该模型中只有一个不定参数 θ ，它提供了一种简单并且较为可靠的计算方法。因此，这是一个比较折衷且简单易用的模型。尤其是在不知道居群遗传结构的情况下，通过这个模型计算出的取样个体数可以作为取样的一个参考值(Brown & Briggs, 1991)。

但是在实际取样过程中往往对取样有一定的要求，例如要求样本应包含居群 95% 以上的遗传变异，或者要求样本包含基因频率小于 0.05 的稀有等位基因。Brown & Moran (1981) 为此提出了下面的计算公式：

$$S \approx 3 \lceil (F - 2) \log_e (1 - p) \rceil \quad (2)$$

S 指样本数， F 指居群的近交系数，即个体在某个基因位点上从上代获得两个相同的等位基因的机率，通常它的值在 -1 和 $+1$ 之间(王中仁, 1998)， p 指基因频率。根据给定居群的近交系数(F)以及取样要求包括的等位基因的频率(p)，可以计算出居群取样的个体数(S)。该取样数可保证 95% 的基因频率为 p 的等位基因至少被保留一份拷贝(Frankel *et al.*, 1995)。这个模型要求已知目标居群的近交系数，对一个已知其遗传结构的居群来说很容易计算出最佳的样本采集数。

Sjögren & Wyöni(1994)提出了一个针对有限群体的假设模型。假定居群内共有 N 个个体，该居群不显著偏离哈迪-温伯格平衡，具有两个等位基因的位点上的常见等位基因频率为 p ，稀有等位基因频率为 q ，并且 $p > q$ ， n 为样本数，

$$P^- = \prod_{i=0}^{2n-1} [(p2N - i) / (2N - i)] \quad (3)$$
$$P^+ = 1 - P^-$$

本式中 P^- 指样本中包含常见等位基因 p 的概率， P^+ 指样本中包含稀有等位基因 q 的概率。作者用计算机程序模拟计算了在不同 N 值和 p 值情况下，使 $P^+ > 0.95$ 的 n 值，当 p 值为 0.95 时，随着居群内个体数 N 的增加，样本数 n 也不断增加，当居群增加到无限大时($N \rightarrow \infty$)， n 值被固定为 30。因此按照 Sjögren & Wyöni 的假设，要检测到频率为 5% 的稀有等位基因，取样数至少要达到 30 个个体($P^+ \geq 0.95$)。

2.2.2 居群内取样的空间样式

一旦居群内的取样个体数确定以后，还要考虑样本的空间分布。异交尤其是风媒物种个体间的基因交流较充分，遗传变异的分布通常比较均匀，与空间距离没有显著的相关性。但是很多物种由于受其繁育系统、种子和花粉传播机制的影响，居群内的遗传变异并不是完全随机分布的。通常自交和无性繁殖的物种个体间的基因交流受到限制，遗传变异的分布呈斑块状，同时生境异质性也会导致居群内遗传变异存在一定的空间分布格局。

Epperson(1990)，Sokal & Oven(1978)，何田华(1999)等将空间自相关分析用于居群遗传结构分析。空间自相关是一种统计方法，可以用来检测和量化从多个标定点中取样值变异的空间依赖性，能够较好地揭示遗传变异的空间分布格局，从而确定遗传变异的空间分布是否随机，如果居群内的遗传变异呈斑块状分布，那么斑块的大小如何等等(何田华, 1999)。这些研究为遗传多样性的取样策略提供了重要的参考。李昂等(2002)用该方法分析了两种兰科植物的居群遗传结构，其中硬叶兜兰(*Paphiopedilum micranthum*)的居群中遗传变异在空间上的分布不均匀，呈 3 ~ 4 m 大小的斑块状，而独花兰(*Changnienia amoena*)居群内的遗传变异不存在明显的空间结构。因此，为了达到合理取样的目的，还应对不同物种居群遗传变异的空间结构有一个大致的了解。

2.2.3 取样的居群数

Hamrick *et al.* (1991) 提出用基因分化系数(G_{st})来估算一个物种内应取样的居群数。他们提出，如果居群间的遗传多样性水平相同，居群间的基因分化系数为 0.6，那么 6 个居群基本上就能够包含该物种 95% 的变异。公式如下：如果取样的居群数为 n ，则所包含的遗传变异的比率为 $1 - (G_{st})^n$ 。那么当 G_{st} 值趋于 1 时，说明居群间的差异极大，即使 n 值很大时也只能包含物种很小一部分遗传变异；当 G_{st} 值趋于 0 时，说明居群间的差异很小，较低的 n 值就能包含大部分的遗传变异。但是这种算法的问题在于：随着 G_{st} 值的增加，居群间的差异增大，可能有的居群变异水平很高而另一些居群的变异水平却较低，这就可能导致采样的居群不能很好地代表该物种整体的遗传多样性。在这种情况下就需要进一步了解哪些居群的变异较大。另外这里用的 G_{st} 值是

各位点的平均值,因此如果 G_{st} 值在不同的位点之间存在差异,就可能导致估计值与实际值之间的误差 (Jennifer *et al.* , 1997)。

3 影响遗传多样性和取样的因素

3.1 物种的生物学特性

物种自身的生物学特性在很大程度上影响其整个基因组的进化和遗传变异的格局。这些特性包括物种的生活史、种子散播和传粉机制、交配系统、生殖模式、突变率和突变机制等。这些特性及其对物种和居群遗传多样性的影响,在很大程度上会左右我们的取样策略。葛颂(1994)引用了前人对 165 属 449 种植物等位酶数据进行的统计处理和分析,指出了等位酶水平上的变异与物种的分类地位、生活型、分布范围、繁育系统等特性之间的关系 (表 1)。

统计学检验表明,在物种水平上影响遗传变异大小的因素依次为:分类地位 > 分布范围 > 生活型 > 繁育系统 > 种子散播机制;而在居群水平上同样是这 5 个因素,但重要性变为:繁育系统 > 分布范围 > 生活型 > 分类地位 > 种子散播机制。对这些关系的认识是我们制定特定物种取样策略的重要参考。

繁育系统是影响居群遗传结构的最重要的因素之一,也同样影响到物种水平上的遗传变异。自交和混交物种的总体遗传变异水平低于异交物种,而自交植物的总体变异虽小,但居群间的分化大,以无性繁殖为主的物种的遗传变异也较小。将不同的 F 值代入 Brown & Moran 提出的公式 $S \approx 3 [(F - 2) \log_e (1 - p)]$,如果要求保留 95% 的基因频率为 0.05 的稀有等位基因,在假定居群总体遗传变异水平一致的情况下,近交系数不同的居群的取样个数分别如下:

表 1 不同类型植物的遗传变异水平和居群分化程度(引自葛颂,1994)
Table 1 Genetic variation and population differentiation in different types of plant species

	物种数 No. of species	P_s	P_p	H_s	H_p	G_{st}
分类地位 Taxon		***	***	**	***	***
裸子植物 Gymnosperm	55 ~ 80	0.709	0.577	0.173	0.160	0.068
双子叶植物 Dicotyledon	246 ~ 338	0.448	0.290	0.136	0.096	0.273
单子叶植物 Monocotyledon	80 ~ 111	0.592	0.403	0.181	0.144	0.231
生活型 Life cycle		***	***	***	***	***
一年生 Annual	146 ~ 190	0.507	0.302	0.161	0.105	0.357
短寿多年生草本 Short-lived perennial herb	119 ~ 159	0.413	0.280	0.116	0.096	0.233
短寿多年生木本 Short-lived perennial tree	8 ~ 17	0.418	0.313	0.097	0.094	0.088
长寿多年生木本 Long-lived perennial tree	110 ~ 131	0.647	0.500	0.177	0.149	0.076
分布范围 Distribution		***	***	***	***	NS
特有种 Specially distributed	52 ~ 100	0.400	0.263	0.096	0.063	0.248
窄域种 Narrowly distributed	82 ~ 115	0.451	0.306	0.137	0.105	0.242
地区种 Locally distributed	180 ~ 193	0.529	0.346	0.150	0.118	0.216
广布种 Widely distributed	85 ~ 105	0.589	0.430	0.202	0.159	0.210
繁育系统 Breeding system		***	***	**	***	***
自交 Inbreeding	78 ~ 125	0.418	0.200	0.124	0.074	0.510
混交(动物传粉)Mixed breeding(animal pollinated)	60 ~ 85	0.400	0.292	0.120	0.090	0.216
异交(动物传粉)Out-crossing(animal pollinated)	124 ~ 172	0.501	0.359	0.167	0.124	0.197
异交(风媒)Out-crossing(wind pollinated)	102 ~ 134	0.661	0.497	0.162	0.148	0.099
种子散播 Seed dispersal		***	***	***	**	***
重力 By gravity	161 ~ 199	0.457	0.298	0.136	0.101	0.277
附着 Attaching	52 ~ 68	0.688	0.421	0.204	0.137	0.257
破裂 Exploding	23 ~ 34	0.304	0.213	0.092	0.062	0.243
吞食 Swallowed	39 ~ 67	0.457	0.324	0.176	0.129	0.223
风传播 Wind mediated	105 ~ 121	0.554	0.429	0.144	0.123	0.143

P_s 和 H_s 分别代表在物种水平上的多态性和杂合度, P_p 和 H_p 分别代表在居群水平上的多态性和杂合度。 P_s 和 H_s represent the polymorphism and heterozygosity at species level, P_p and H_p represent the polymorphism and heterozygosity at population level.
*** $P < 0.001$, ** $P < 0.01$, NS: 不显著 Non-significance

$F = 1$ 时： $S \approx 58$
 $F = 0$ 时： $S \approx 30$
 $F = -1$ 时： $S \approx 19$

近交系数 F 值可以在 -1 到 1 之间变动： $F = -1$ 指全部杂合(如雌雄异株)， $F = 1$ 指全部纯合(如严格自交)， $F = 0$ 指实际等位基因频率符合哈迪-温伯格平衡，该居群是随机交配的(王中仁，1998)。由此可见，居群内的杂合程度越高，等位基因在个体间的分布越均匀，较少的个体就能代表整个居群的遗传变异。反之，居群内的杂合程度越低，等位基因在个体间的分布越不均匀，就需要较多的个体才能代表整个居群的遗传变异。

Ellstrand & Roose (1987)统计了 21 种克隆繁殖植物的数据，发现大部分克隆植物的总体遗传多样性水平较低，有两个物种完全没有变异。但居群之间的遗传多样性变异较大，75.8% 的基因型只存在于 1 个居群，也有所有基因型都只存在于 1 个居群的极端情况。当然也有一些分布范围较广的克隆植物例外。Xu *et al.* (2003)对来自中国 5 个省的喜旱莲子草(*Alternanthera philoxeroides*) RAPD 的研究发现，这些样本的所有位点在 8 个居群之间几乎没有变异，说明了该物种克隆繁殖的特点。

通常物种的分布区大小对遗传多样性有显著的影响。不管是从居群水平上还是从物种水平上来看，遗传多样性与物种分布范围大小成正比，因为广布种分布的环境差异大，个体和居群的数目多，适应幅度大，基因交流频繁，维持遗传变异的能力强，产生新的遗传变异的机会也多。Lu *et al.* (2002)对来自亚洲 36 个野生稻(*Oryza rufipogon* 和 *O. nivara*)居群的 RFLP 研究表明，地理隔离在野生稻的遗传分化中具有非常重要的作用。

植物种子的传播机制是影响遗传变异在居群内和居群间分布的另一个重要因素。靠重力、附着和果实破裂散播种子的物种，其 G_{st} 值较高，而靠风力散播种子的物种，其 G_{st} 值较低。这是因为后者不同居群间的基因交流比率较高，故居群间的分化水平较低，而靠重力、附着和果实破裂散播种子的物种，其后代向外扩展的能力有限。动物参与传播种子的情况比较复杂，因为动物传播种子不仅方式多样，而且不同动物的迁移能力相差很大，有些动物虽然能将种子携带到很远的地方，但这种作用不一定能显著影响居群间的遗传变异(葛颂，1994)。

3.2 环境条件的影响

物种的遗传多样性水平及分布还受到环境条件的影响，这是一个重要的植物生态学事实。影响居群分化的环境因子包括：温度、光照、白昼长度、湿度、土壤特性及营养状况等(Frankel *et al.*，1995)。物种在其分布范围内的不同生境条件下面临不同的选择压力，由于地理距离会造成居群之间的分化和隔离，因此广布种通常由于适应当地的生境条件而形成在遗传上有区别的不同生态型。Gustavsson (2001)，Gram & Sork (2001)和 Apist & Dixon (2001)的研究都表明，在大尺度上，居群间的遗传变异和环境异质性之间存在显著的相关性。

在小尺度上，物种内的居群间也存在与生境相关的遗传分化。Liu & Godt (1983)总结了居群间存在遗传分化的现象，居群间的遗传分化存在于相邻、相距 3 m、10 m、100 ~ 200 m、1000 m 的居群间。造成居群间遗传分化的原因是生境条件的不同(如环境异质性或人类活动引起的环境条件变化)及居群间的隔离。

4 遗传多样性的取样策略

依据上述分析我们不难看出，取样策略的制定与许多因素有关，但有两方面的信息是最为重要的：一是研究对象本身的特性，如物种的遗传多样性水平和分布、居群遗传结构、基因流和繁育系统以及由环境因素引起的变化等；二是取样的目的，是为了遗传多样性的保护和可持续利用，还是为了进行某一方向的研究，是考虑整个物种，还是该物种在特定区域的居群之间或者是居群内的遗传多样性等。如果研究对象是整个物种，那么样本要能够代表该物种的整体遗传多样性水平，应该尽可能覆盖该物种的分布区；如果研究对象是物种在特定区域的居群间的遗传多样性水平，那么就應該在该区域合理地选取具有代表性的居群；如果同时还要考虑其他因素对该物种的影响，则取样时要设计各因素的梯度或对照。针对不同的研究对象和研究目的应该制定相应的取样策略。

4.1 居群内的样本数和取样方式

从前面的不同模型我们可以看出，在已知物种的遗传多样性水平和结构的情况下，比较容易根据居群的遗传结构参数来确定样本的数量，也可以根据居群遗传变异的空间分布格局来确定取样的方

式。但是在大多数情况下,尤其是对很多野生物种,很难获得其全面的遗传结构信息。在这种情况下,一种比较简单的方法就是根据理论公式和经验数据来确定样本数。

通常认为一个居群内最佳的取样数要求包括 95% 以上基因频率大于 0.05 的等位基因(Marshall & Brown, 1975)。按照这个标准,根据公式(2)计算得到近交系数 F 值在 -1 到 1 之间变动时,样本数应在 19 ~ 58 之间。如 $F = 0$, 则随机交配居群的样本数为 30, 这个样本数与 Sjögren & Wyöni(1994)根据公式(3)计算的结果基本一致。因此在不知道某一物种生物学特性和遗传背景的情况下,对一个居群进行资源保护的随机取样数目可以确定在 30 个左右。如果该植物为异交,样本的数量可以适当降低;如果该植物为自交,可以适当增加样本的数量。Jin *et al.*^① 用分子标记(ISSR)对野生大豆这一严格自交的一年生植物的居群遗传学研究和多次随机抽样统计的结果表明, 35 ~ 45 个样本能基本包含该居群 100 个样本中绝大部分的遗传变异。居群内取样的空间距离也应根据其繁育系统、居群内生境的异质性和遗传多样性的结构及空间分布来决定。Jin *et al.*^① 的研究还表明,对野生大豆取样时每一个个体应该相隔 10 m 以上,这样才能保证不对遗传变异完全相同的个体重复取样。

在居群取样的个体数和空间距离确定之后,还可以根据具体研究目的采取不同的取样方法,以达到更好的效果。居群内的取样方法有简单随机取样(simple random sample)、分层随机取样(stratified random sample)、系统取样(systematic sample)和偏差取样(biased sample)。在不知道居群的遗传多样性水平及其分布、繁育系统、种子传播及花粉传播等机制的情况下,采用随机取样的方法是比较可取的;而当居群内的生境异质性比较明显,个体分布于不同斑块时,采用分层取样的方法比较合适,即首先将居群按异质生境分成不同斑块,然后在斑块内进行随机取样。系统取样要求样本在居群内平均分布,这在操作上与严格的随机取样相比较为简单,并且能保证样本覆盖整个居群,但是如果遗传变异呈周期性变化,这样的取样容易产生偏差。偏差取样是将某一性状如花色和叶片形状等的变异作为标准进行取样,这种取样方法的问题是这些形态学变异不一定反映了遗传水平的变异(Brown & Briggs,

1991)。

4.2 取样的居群数及每一区域内居群的选择

确定某一给定地区居群取样数目要比确定某一居群内的样本数更加困难。居群本身就很难定义。陈家宽(1994)列举了居群的 8 种不同定义,这些定义从不同的侧面强调了居群的特点。文中采用的居群定义是:“在特定空间和时间里生活着的自然或人为的同种个体群”。在某些情况下,如在一个相对隔离的池塘中或小岛上,居群是一个比较自然的单位,但是在许多情况下,居群的划分很大程度上受人为因素的影响。如我们可以把上海市连续分布的野生大豆(*Glycine soja*)看作一个居群,也可以把上海江湾机场内的野生大豆看作一个居群。另外,除了濒危珍稀物种,通常一个物种的分布区内有很多居群,一些连续分布的居群由于生境的片段化有可能被分割成若干亚居群,因此居群的数量可以很大。重要的是根据我们研究的目的和可操作性来确定居群的大小。

在确定居群的取样数目时,应首先考虑所研究对象的分布区域,使所选择的居群尽可能覆盖整个分布区域。如果已知该物种居群间的分化系数,可根据 $1 - (G_{st})^n$ 来估算应该采样的居群数。在给定的区域内选择取样居群有两种最简单的方法:一种是均匀取样法,即所选的居群均匀分布于该区域。这种方法通常用于对未知的物种进行取样,它可以获取该区域内受土壤、气候等环境因素影响所形成的最大量的变异,但如果该区域的范围很大则工作量很大。另一种方法是聚集取样法,即取样居群呈聚集状分布,每组内有若干居群,组间大致呈均匀分布。这种方法可以获取大量受地理位置和微地形变化影响而产生的变异,因而更适用于野生物种,因为野生物种的遗传变异在长期的进化过程中主要受到自然环境因素的影响,野生物种通过花粉、种子传播的基因流受人为因素的影响相对较少,因此相当一部分变异是和环境异质性相关的(Marshall, 1989)。

5 小 结

遗传多样性的取样策略是与保护、研究和可持续利用生物资源紧密相关的重要命题,直接关系到

① Jin Y, Zhang W-J, Fu D-X and Lu B-R. 2003. Sampling strategy within a wild soybean population based on its genetic variation detected by ISSR markers. *Acta Botanica Sinica* (submitted)

保护和利用的对象是否包括了必要和尽可能多的多样性,还关系到我们的研究结果是否具有代表性和重要参考价值。由于研究对象自身复杂的生物学特性和受不同环境条件的影响,使得物种的居群间和居群内的遗传多样性水平和分布式样都有很大的变化,只有了解这些变化的规律并结合取样的目的来制定不同的取样策略,才能够在所采集的有限样本中包含尽可能多和能代表研究对象特点的遗传多样性。

参考文献

Apist V. J. and Dixon P. M. 2001. Genetic diversity and population structure in *Echinacea laevigata* (Boynton et beadle) Blake, an endangered plant species. *Natural Areas Journal*, **21**(1): 71 – 77.

Brown A. H. D. and Briggs J. D. 1991. Sampling strategies for genetic variation in *ex situ* collections of endangered plant species. In: Falk D. A. and Holsinger K. E. (eds.), *Genetics and Conservation of Rare Plants*. Oxford University Press, New York, 99 – 119.

Brown A. H. D. and Moran G. F. 1981. Isozymes and the genetic resources of forest trees. In: Conkle M. T. (ed.), *Isozymes of North American Forest Trees and Forest Insects*. Pacific South West Forest and Range Experiment Station Technical Report No. 48. US Department of Agriculture, Washington, D. C., 1 – 10.

Chen J-K (陈家宽). 1994. Population biology and evolutionary biology. In: Chen J-K (陈家宽) and Yang J. (杨继) (eds.), *Plant Evolutionary Biology* (植物进化生物学). Wuhan University Press, Wuhan, 1 – 47. (in Chinese)

Ellstrand N. C. and Roose L. M. 1987. Patterns of genotypic diversity in colonial species. *American Journal of Botany*, **74**(1): 123 – 131.

Epperson B. K. 1990. Spatial autocorrelation of genotypes under directional selection. *Genetics*, **124**(3): 757 – 771.

Ewens W. J. 1972. The sampling theory of selectively neutral alleles. *Theoretical Population Biology*, **3**: 87 – 112.

Frankel O. H., Brown A. H. D. and Jeremy J. B. 1995. *Conservation of Plant Biodiversity*. Cambridge University Press, London, 10 – 38.

Ge S (葛颂). 1994. Allozyme and plant evolutionary biology study. In: Chen J-K (陈家宽) and Yang J (杨继) (eds.), *Plant Evolutionary Biology* (植物进化生物学). Wuhan University Press, Wuhan, 153 – 208. (in Chinese)

Gram W. K. and Sork V. L. 2001. Association between environmental and genetic heterogeneity in forest tree populations. *Ecology*, **82**(7): 2012 – 2021.

Gustavsson B. A. 2001. Genetic variation in horticulturally im-

portant traits of fifteen wild lingonberry *Vaccinium vitis-idaea* L. populations. *Euphytica*, **120**(2): 173 – 182.

Hamrick J. L., Godt M. J. W., Murawski D. A. and Loveless M. D. 1991. Correlation between species traits and allozyme diversity: implication for conservation biology. In: Falk D. A. and Holsinger K. E. (eds.), *Genetics and Conservation of Rare Plants*. Oxford University Press, New York

He T-H (何田华). 1999. Spatial autocorrelation analysis of plant population genetic variation. *Chinese Bulletin of Botany* (植物学通报), **16**(6): 636 – 641. (in Chinese)

Jennifer F. C., Affolter M. J. and Hamrick J. L. 1997. Developing a sampling strategy for *Baptisia arachnifera* based on allozyme diversity. *Conservation Biology*, **11**(5): 1133 – 1139.

Li A. (李昂), Wang K-Q (王可青) and Ge S. (葛颂). 2000. Genetic diversity within and among population of *Viola tenuicornis* with reference to sampling strategy. *Acta Botanica Sinica* (植物学报), **42**(10): 1069 – 1074. (in Chinese)

Li A. (李昂), Luo Y-B (罗毅波) and Ge S. (葛颂). 2002. Spatial autocorrelation study of population genetic structure of two orchid species. *Biodiversity Science* (生物多样性), **10**(3): 249 – 257. (in Chinese)

Liu E. H. and Godt M. J. W. 1983. The differentiation of population over short distances. In: Schonewald-Cox C. M., Chambers S. M., MacBryde B. and Thoms W. L. (eds.), *Genetics and Conservation*. The Benjamin/Cummings Publishing Company, Inc., Menlo Park, 78 – 96.

Lu B. R., Zheng K. L., Qian H. R. and Zhuang J. Y. 2002. Genetic differentiation of wild relatives of rice as referred by the RFLP analysis. *Theoretical and Applied Genetics*, **106**: 101 – 106.

Marshall D. R. and Brown A. H. D. 1975. Optimum sampling strategies in genetic conservation. In: Frankel O. H. and Hawles J. G. (eds.), *Crop Genetic Resource for Today and Tomorrow*. Cambridge University Press, London, 53 – 80.

Marshall D. R. 1989. Crop genetic resources: current and emerging. In: Brown A. H. D., Clegg M. T., Kahler A. L. and Weir B. S. (eds.), *Genetics and Conservation of Rare Plants*. Oxford University Press, New York, 367 – 388.

Sjögren P. and Wyöni P. I. 1994. Conservation genetics and detection of rare alleles in finite populations. *Conservation Biology*, **8**(1): 267 – 270.

Sokal R. R. and Oven N. L. 1978. Spatial autocorrelation in biology. I. Methodology. *American Naturalist*, **138**: 518 – 521.

Wang Z-R (王中仁). 1998. *Plant Allozyme Analysis* (植物等位酶分析). Science Press. Beijing, (in Chinese)

Xu C-Y, Zhang W-J, Fu C-Z and Lu B-R. 2003. Genetic diversity of alligator weed in China by RAPD analysis. *Biodiversity and Conservation*, **12**(4): 637 – 645.