

鳞翅目昆虫基因组中微卫星 DNA 的特征以及对 其分离的影响*

吉亚杰 张德兴**

中国科学院动物研究所农业虫鼠害综合治理国家重点实验室, 北京 100080

摘要 本文根据我们对鳞翅目昆虫棉铃虫和松毛虫以及其它动物(筱蜘蛛、朱鹮、鳕鱼和飞蝗)的微卫星富集性基因组DNA文库的筛选和分析结果, 结合其它实验室已发表的资料, 对鳞翅目昆虫基因组中微卫星DNA的丰度和结构特点进行了较为系统的分析。结果表明: 与其它类群相比, 尽管鳞翅目昆虫物种间存在差异, 但其基因组中存在明显偏多的侧翼序列重复的、以多拷贝形式存在的微卫星位点, 且其中相当一部分以基因家族的形式存在。微卫星DNA家族通常可以在序列分析阶段被识别出来, 但很多多拷贝位点只有通过一系列后续分析才能被检查出来。这是鳞翅目昆虫中微卫星位点的优化率相对偏低的主要原因。棉铃虫和松毛虫基因组中三相重复微卫星丰度相对较高, 从而从某种程度上补偿了这些物种微卫星分离过程中因丰度低、多拷贝位点比例高所带来的困难。棉铃虫微卫星DNA家族侧翼序列中多聚T/A序列的存在表明, 逆转录转座或逆转录侵染可能是在基因组中形成多拷贝微卫星位点和微卫星DNA家族的重要机制之一 [动物学报 50 (4): 608–614, 2004]。

关键词 鳞翅目昆虫 棉铃虫 松毛虫 微卫星DNA家族 逆转录转座

Characteristics of microsatellite DNA in lepidopteran genomes and implications for their isolation*

JI Ya-Jie, ZHANG De-Xing**

State Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, The Chinese Academy of Sciences, Beijing 100080, China

Abstract A comparative analysis of microsatellite DNA in lepidopteran insects was carried out, based on our data from microsatellite-enriched genomic libraries of lepidopteran insects (the cotton bollworm and the masson pine moth) and a variety of other organisms (the migratory locust, gadoid fishes, fen raft spider and the crested ibis), as well as published data on other lepidopteran insects. Our results suggest that: 1) Both the proportion of the positive clones in genomic libraries and the optimization efficiency of microsatellite loci are much lower in lepidopteran insects than in other organisms (Tables 1, 3); 2) A much larger proportion of microsatellite loci bear repetitive flanking region (s) in lepidopteran insects than in other organisms (Table 3) and some are present as multi-copy microsatellite DNA families (Fig.1). This means that although the overall abundance of microsatellite loci in lepidopterans is lower, the proportion of multi-copy loci is more pronounced. Therefore, the majority of positive clones from genomic libraries are not suitable for use as diagnostic markers in population genetic analysis. This considerably increases the difficulty in microsatellite isolation in lepidopteran insects. 3) There are relatively more trinucleotide microsatellite loci in the cotton bollworm, with the majority of them being a single copy locus (Table 2). This compensates somewhat the difficulty in the isolation of effectively usable microsatellite markers in this species. 4) Null allele appears to be a commonly occurring phenomenon for lepidopteran polymorphic microsatellite loci. It means that there possibly exists an elevated mutation rate in the flanking regions of microsatellite sequences in this group of insects. 5) The existence of poly (T/A) sequences in the repetitive flanking regions of microsatellites (Fig.1) suggests that retrotransposition or retrofection may be involved in the formation of microsatellite DNA families and other multi-copy microsatellite loci. [Acta Zoologica Sinica 50 (4): 608–614, 2004].

2004-01-06 收稿, 2004-03-23 接受

* 国家自然科学基金(No. 30270223)、国家杰出青年科学基金(No. 30025008)和中国科学院“百人计划”资助项目 [This research was funded by the National Natural Science Foundation of China (Grant Nos 30270223 and 30025008) and the Chinese Academy of Sciences (“Bai Ren Ji Hua” Program)]

** 通讯作者(Corresponding author). E-mail: dxzhang@ioz.ac.cn

© 2004 动物学报 Acta Zoologica Sinica

Key words Lepidopteran insects, Cotton bollworm, *Helicoverpa armigera*, Masson pine moth, *Dendrolimus punctatus*, Microsatellite DNA family, Retrotransposition

微卫星 (Microsatellite) DNA 有时也被称为短串联重复序列 (Short tandem repeat) 及简单串联重复序列 (Simple tandem repeats, 包括小卫星 minisatellite)，一般指以 1–6 个碱基为基本组成单元的串联重复序列。相对于其它分子标记而言，微卫星 DNA 具有分布广泛、选择中性、共显性遗传、多态性高、遗传检测时对模板 DNA 数量和质量要求低、分析方法简捷、实验结果稳定可靠、重复性好等优点，是当前种内变异研究中分辨率最高、揭示力最强的细胞核 DNA 标记 (Zhang and Hewitt, 2003)。微卫星 DNA 在亲缘关系鉴定、种群遗传结构、近缘种的系统进化关系、种群演化历史、繁殖行为、社群结构等遗传和进化研究中得到了广泛的应用 (Blouin et al., 1996; Bowcock et al., 1994; Dib et al., 1996; Favre et al., 1997; Queller et al., 1993)。然而，建立微卫星 DNA 标记系统对实验室综合实力的较高要求，成为其更广泛应用的重要限制性因素。鳞翅目昆虫微卫星 DNA 标记的分离则更是本领域公认的难点问题，在同样的建库方法下所得到的可用微卫星位点的比例往往少于其它物种 (Meglécz and Solignac, 1998; Nève and Meglécz, 2000)，尽管不同的学者对造成这种现象的原因进行了推测 (Nève and Meglécz, 2000; Ji et al., 2003)，但目前仍缺少这方面的系统研究。

我们通过构建基因组 DNA 文库先后为棉铃虫和松毛虫等六个物种建立了微卫星标记系统。本文依据我们的实验结果，结合其它实验室已发表的资料，对鳞翅目昆虫基因组中微卫星 DNA 的丰度和结构特点进行了较为系统的分析，并讨论了它们对于在鳞翅目昆虫中建立和应用微卫星标记的影响。

1 材料与方法

1.1 实验材料

供试棉铃虫 31 条，幼虫期浸入无水乙醇中保存；松毛虫 12 条，无水乙醇中保存；筏蜘蛛 2 只，采自英国 Middle Fen 地区的成年个体，腿部取下后立即进行 DNA 抽提；朱鹮血液样品采自北京动物园。

1.2 DNA 抽提

总量 DNA 提取采用 Zhang and Hewitt (1998)

的方法并做如下改进：组织块在室温条件下研磨，研磨前后置于冰上；所有样本研磨完成后置于 55℃ 水浴 2–16 h；将样本由水浴锅中取出，点转离心后加入等体积的 Tris-饱和酚 (pH 8.0)，轻轻混匀，置于冰上 2 min，中间再混合一次；在 4℃ 条件下以 8 000 r/min 离心 10 min，取上清液，用 100 μl Tris-饱和酚 (pH 8.0) 抽提 2 次或 Tris-饱和酚 (pH 8.0) 和 TE-饱和氯仿各抽提 1 次，若两相界面上仍有白色沉淀存在，则用 100 μl Tris-饱和酚 (pH 8.0) 继续抽提几次；以相对于上清液 1 倍体积的醋酸氨 (10 mol) 和 5 倍体积的无水酒精沉淀 DNA，室温放置 15 min 以上后 12 000 r/min 离心 15 min，弃上清液，将沉淀分别用 70% 酒精洗两次、100% 酒精洗一次，室温晾干后溶于 100 μl TE 中。

1.3 微卫星富集性基因组 DNA 文库的构建

本文所分析的微卫星文库的构建策略参见 Zhang et al. (2001) 和 Ji et al. (2003)。简述如下：基因组 DNA 经限制性内切酶 (BRL 或 Bio-labs) 消化后与接头相连接 (T_4 DNA 连接酶, Invitrogen)，经过一次富集后与 pBleuscript 噬菌粒载体 (Stratagene) 连接，导入大肠杆菌培养过夜，通过 α -互补原理选择含有重组序列的克隆，培养后转到尼龙膜上进行杂交筛选。杂交筛选所使用的探针共 8 种，包括：(GT)_n、(GA)_n、(GTA)_n、(CCA)_n、(TTC)_n、(CAG)_n、(GTGA)_n、(AGTT)_n。阳性克隆筛选采用非放射性检测方法，使用 Sigma 公司生产的 CDP-StarTM 通用检测试剂盒，按试剂盒所附说明书进行操作。同时为棉铃虫、松毛虫、筏蜘蛛、朱鹮等物种建立了微卫星富集性基因组 DNA 文库。

1.4 序列测定和引物设计

序列测定使用 Applied Biosystems 公司的测序反应试剂盒 (ABI BigDyeTM Terminators Cycle Sequencing Kit V2.0)，按试剂盒所附的说明书准备测序反应体系，反应在 Applied Biosystems 公司的 9700 型 PCR 仪上进行，序列检测在 ABI PRISM[®] 3100 遗传自动分析仪上进行。单个序列的分析和校正用 Sequence Analysis (V. 3.7, Applied Biosystems) 软件人工完成。序列间比较使用 Sequencher (Gene Codes Corporation) 软件。引物设计用 Oligo[®]

6.31 (National Biosciences Inc.) 软件。

1.5 引物优化

引物的 PCR 反应条件均在标准 PCR 反应条件的基础上通过调整 Mg^{++} 浓度、退火温度等因素而确定。标准 PCR 反应条件为: Mg^{++} , 1.5 mmol; dNTPs, 200 μ mol; 引物, 0.5 μ mol; 退火温度, 根据引物的 T_m 值确定。基因型检测在 ABI PRISM® 3 100 遗传自动分析仪上进行, 分子量内标为 GENESCAN® 400HD [ROX] Size Standard (Applied Biosystems)。

2 结果

2.1 微卫星 DNA 丰度

表 1 详细列出了本研究所涉及的几个物种以及目前所发表的其它鳞翅目物种的微卫星 DNA 位点的情况, 以及其中部分物种的阳性克隆比率和最终筛选出的多态性位点的比率。可以看出, 对大多数鳞翅目昆虫而言, 一次建库所筛选到的可用微卫星位点数为 2~5 个, 阳性克隆在全部重组克隆中所占的百分比多小于 4%, 最终筛选出的多态性位点的比率一般小于 1%。

2.2 微卫星的种类

表 2 列出不同物种微卫星 DNA 富集文库中各

类型微卫星位点的比例。本文所重点分析的六个物种中二相重复微卫星都占有较大的比例(从 56% 到 91.7%), 其中绝大多数都是 CA/GT 重复(54.5%~94.6%), 其次是 CT/GA 重复(5.4%~45.5%), AT/GC 重复最少, 仅占 0~8%; 三相重复微卫星的丰度为 0~28.6%, 且在不同物种中有较大差异。值得强调的是, 在棉铃虫和松毛虫中, 三相重复微卫星分别占多态位点总数的 80% 和 36%, 明显多于其它物种。

2.3 多拷贝微卫星位点

所构建的各物种的微卫星 DNA 文库的筛选结果总结于表 3: 在棉铃虫、松毛虫、筱蜘蛛、朱鹮、飞蝗以及鳕鱼中, 克隆效率为 3.8%~20%, 多态位点比率为 0.6%~5%。需要特别指出的是, 在大多数物种中观察到侧翼序列在基因组中重复的、很可能以多拷贝形式存在的微卫星 DNA 位点, 其比率为 24%~87.9%, 以棉铃虫基因组中筛选出的多拷贝位点的比例最高, 占 87.9%。此外, 在棉铃虫和松毛虫中我们还观察到微卫星 DNA 家族(即侧翼序列高度相似的一组微卫星 DNA 位点, 图 1) 的存在, 其比率为 30.7% 和 13.3%。

表 1 鳞翅目昆虫微卫星 DNA 的情况

Table 1 Microsatellite loci characterized in lepidopteran insects

| 物种 Species | 微卫星位点数 Number of polymorphic loci | 探针数 Number of probes | 阳性克隆 Positive clone(%) | 多态位点 Polymorphic loci(%) | 数据来源 Reference |
|-----------------------------------|--------------------------------------|-------------------------|---------------------------|-----------------------------|---------------------------|
| <i>Melitaea cinxia</i> | 2 | — | 0.0125 | 0 | Palo et al., 1995 |
| <i>Lymantria dispar</i> | 4 | — | — | — | Bogdanowicz et al., 1997 |
| <i>Parnassius mnemosyne</i> | 3 | 4 | 0.83 | 0.18 | Meglécz and Solignac 1998 |
| <i>Parnassius smitheus</i> | 4 | 1 | <6 | 0.6 | Keyghobadi et al., 1999 |
| <i>Bombyx mori</i> | 15 | 2 | 2.37 | 0.4~0.8 | Reddy et al., 1999 |
| <i>Lysandra bellargus</i> | 5 | — | — | — | Harper et al., 2000 |
| <i>Lycaeides melissa samuelis</i> | 4 | — | — | — | Anthony et al., 2001 |
| <i>Speyeria idalia</i> | 4 | 1 | 0.03 | 0.003 | Williams et al., 2002 |
| <i>Lymantria dispar</i> L. | 3 | 10 | 0.125 | — | Koshio et al., 2002 |
| <i>Heliconius erato</i> | 15 (3 个库, 3 libraries) | — | — | — | Flanagan, 2002 |
| <i>Helicoverpa armigera</i> | 5 (2 个库, 2 libraries) | 8 | 2.5 | 0.6 | Ji et al., 2003 |
| <i>Coenonympha hero</i> | 7 | 3 | 0.91 | 0.09 | Cassel, 2002 |
| <i>Zale galbanata</i> | >5 | — | — | — | Caldas et al., 2002 |
| <i>Bicyclus anynana</i> | — | — | 0.133 | 0 | Meglécz et al., 1998 |
| <i>Spodoptera frugiperda</i> | — | — | 0.077 | 0 | Meglécz et al., 1998 |
| <i>Coenonympha hero</i> | — | — | <1 | ND | Meglécz et al., 1998 |

表 2 不同物种微卫星富集文库中各类型微卫星位点的比例

Table 2 Percentage of different types of microsatellite loci in various species

| 重复类型 Microsatellite type | 二相重复 Dinucleotide | | | | 三相重复 Trinucleotide | | 其它 Other types (%) |
|--------------------------------|-------------------|--------------|--------------|-------------|---------------------------------------|--|--------------------------|
| | CA/GT (%) | CT/GA (%) | AT/GT (%) | 合计 Total | 占阳性克隆的百分率 % of all positive clones | 占多态位点的百分率 % of all characterized loci | |
| | | | | | | | |
| 棉铃虫 CBW | 55.5 | 3.2 | — | 58.7 | 28.6 | 80 | 12.7 |
| 松毛虫 MPM | 44.8 | 8 | 3.2 | 56 | 15 | 36 | 29 |
| 筏蜘蛛 FRS | 53.1 | 10.4 | 5.5 | 69 | 5.5 | 0 | 25.5 |
| 鳕鱼 PC | 50 | 41.7 | — | 91.7 | 0 | 0 | 8.3 |
| 飞蝗 ML | 58.0 | 28 | 4.7 | 90.7 | 7 | 12.5 | 2.3 |

CBW: Cotton bollworm, *Helicoverpa armigera*. MPM: Masson pine moth, *Dendrolimus punctatus*. FRS: Fen raft spider, *Dolomedes plantarius*.PC: Poor cod, *Trisopterus minutus*. ML: Migratory locust, *Locusta migratoria*.

表 3 各物种微卫星 DNA 文库的筛选结果

Table 3 Summary characteristics of microsatellite library screenings

| 物种 Species | 棉铃虫 | 松毛虫 | 筏蜘蛛 | 朱鹮 | 飞蝗 | 鳕鱼 |
|---------------------------------|-----------------|----------------|-----------------|--------------|------------------|----------|
| | Cotton bollworm | Mass pine moth | Fen raft spider | Crested ibis | Migratory locust | Poor cod |
| 克隆效率 | 3.8 | 10.5 | 14.62 | 18.99 | 14 | 20 |
| Cloning efficiency (%) | | | | | | |
| 多拷贝位点比率 | 87.9 | 42.3 | 39 | 24 | 35 | 0 |
| Multicopy loci (%) | | | | | | |
| 微卫星家族比率 | 30.7 | 13.3 | 0 | 0 | 0 | 0 |
| Microsatellite DNA families (%) | | | | | | |
| 多态位点比率 | 0.6~0.9 | 2 | 3 | 3 | 2.3 | >5 |
| Polymorphic loci (%) | | | | | | |

克隆效率指含有微卫星的阳性克隆占所有重组克隆的百分比, 多拷贝位点比率指多拷贝位点占所有阳性克隆的百分比, 微卫星家族比率指在序列分析阶段能够排除的以多基因家族形式存在的多拷贝位点占所有阳性克隆的百分比, 多态位点比率指最终所获得的多态性单拷贝微卫星位点占所有重组克隆的百分率。

Cloning efficiency refers to the percentage of true positive clones in a library. Multi-copy loci % refers to the percentage of multi-copy microsatellites in all true positive clones. Microsatellite DNA families % refers to the percentage of microsatellites existing in gene families among all true positive clones. Polymorphic loci % refers to the percentage of polymorphic microsatellite loci in a library, that is, in all recombinant clones.

图 1 显示棉铃虫的一组多拷贝 GT/AC 微卫星 DNA 家族。注意这些不同微卫星 DNA 位点的侧翼序列在序列结构上高度相似, 而且还有一个典型特征, 即 3' 端侧翼序列含有一段多聚 T [poly (T)] 序列。与此相似, 松毛虫中也有一组多拷贝非完全重复 (AY/RT) 和一组完全重复 (CATA) 微卫星 DNA 家族。在其它四个物种的文库中则未发现这样的多拷贝微卫星 DNA 家族。

2.4 无效等位基因

我们所优化出的 10 个棉铃虫的微卫星位点在种群中都存在无效等位基因, 其中两个位点的无效等位基因频率在个别种群中可达 50%, 另有四个位点在 3%~5% 之间。飞蝗的 8 个微卫星位点在飞蝗北方种群中也都存在不同程度的无效等位基

因, 但其频率都不超过 5% (实验室未发表的数据)。

3 讨 论

3.1 微卫星 DNA 丰度

迄今为止, 除家蚕 (Reddy et al., 1999) 外在成功建立微卫星 DNA 标记的十余种鳞翅目昆虫中, 一次建库所得到的可用微卫星位点一般不超过 5 个 (表 1)。从表 1 数据可以看出, 棉铃虫和松毛虫微卫星富集文库中微卫星的阳性克隆比例分别为 3.8% 和 10.5%, 低于用同样方法同时建立的筏蜘蛛和朱鹮微卫星文库及用类似的方法建立的鳕鱼 (Zhang et al., 2001) 和飞蝗 (Zhang et al., 2003) 的微卫星文库 (表 3)。我们在棉铃虫中所得到的

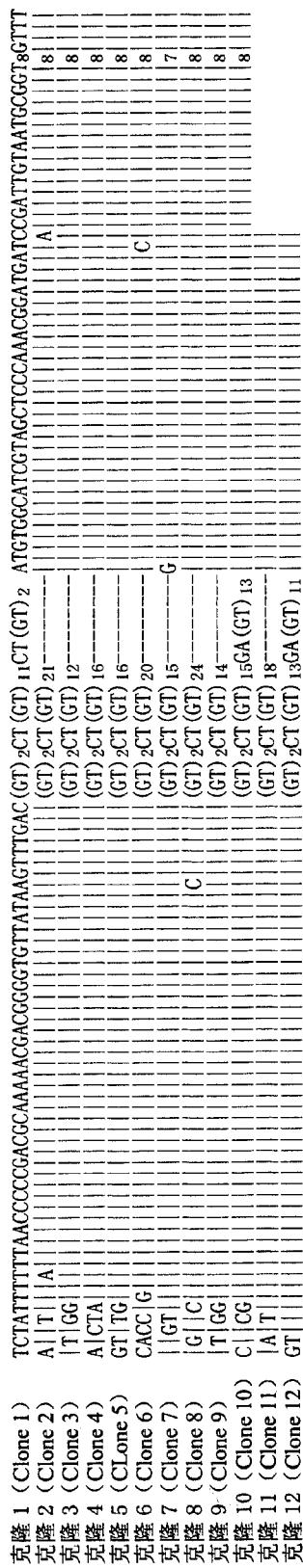


图 1 棉铃虫中的一组多拷贝微卫星位点

“|”，碱基与克隆 1 相同；“—”，多重排序引入的空隙。所有 12 个克隆的 5' 末端均有 50 bp 的序列几乎等同，克隆 1–10 的 3' 末端有 60 bp 的序列等同，克隆 11–12 的 3' 末端因被切断而缺乏序列信息。注意 3' 端侧翼序列有一段多聚 T [poly(T)] 序列。

Fig. 1 AN example of microsatellite DNA family in the cotton bollworm

Vertical bar in the sequences signifies nucleotide which are identical to those in clone 1. “—”alignment gap. The first 50 bp in the 5' end of all sequences are nearly identical. Sequence of the first 60 bp of the 3' end of clone 1 to clone 10 is identical. No sequence information of the 3' end of clone 11 and clone 12 is available. Note the poly (T) tract in the 3' flanking region.

结果与其它已发表的鳞翅目昆虫的结果相近，而松毛虫的克隆效率却高于其它鳞翅目物种，表明不同鳞翅目物种间存在一定差异，Reddy et al. (1999)

在家蚕中的结果也说明这一点。但即使对家蚕而言，与其它非鳞翅目物种相比，克隆效率和多态位点率仍处于较低的水平（表 1, 3）。总体而言，鳞翅目昆虫中微星 DNA 丰度相对偏低是一个普遍存在的现象。

3.2 多拷贝点与微卫星克隆效率

在我们所研究的大多数物种中观察到侧翼序列在基因组中重复的、很可能以多拷贝形式存在的微卫星 DNA 位点（表 3），其比例占阳性克隆的 24% – 87.9%，并以棉铃虫基因组中筛选出的多拷贝位点的比例为最高（87.9%）。而且在鳞翅目昆虫棉铃虫和松毛虫中都发现了以基因家族形式存在的多拷贝微卫星位点（图 1），而在笈蜘蛛、朱鹮、飞蝗、鳕鱼以及刚刚完成序列分析工作的黑斑蛙和泽蛙（未发表）的微卫星文库中则未观察到这种现象。所以，相对于其它物种而言，棉铃虫和松毛虫基因组中不仅可能含有较多的多拷贝微卫星位点，而且这些多拷贝位点中有相当一部分是以基因家族的形式存在。

从克隆效率来看，在使用同样方法所构建的文库中，棉铃虫的克隆效率最低，仅为其它物种的十分之一到三分之一；松毛虫的克隆效率虽高于棉铃虫，但仍不能达到其它物种的平均水平，仅是其它多数物种的 2/3 左右。因此，在同样的条件下，在棉铃虫和松毛虫中所获得的多态性微卫星位点的比例要少于其它物种（表 3），这应是其基因组中较低的微卫星 DNA 丰度和较高比例的重复序列交互作用的结果。鳞翅目昆虫基因组中微卫星 DNA 总体丰度低、多拷贝点多的特点，大大降低了微卫星 DNA 的克隆效率，显著增加了在这些类群中建立适于种群遗传分析的微卫星标记的难度。因此，在鳞翅目昆虫微卫星文库构建过程中需要准备足够多的重组克隆；同时，在引物设计前应对所有微卫星序列进行深入分析，以便尽可能地剔除重复序列，最大限度的减少不必要的后期投入。

大量多拷贝微卫星位点的存在意味着使用富集性克隆的方法并不能有效提高克隆效率，因为尽管富集性克隆能提高阳性克隆率，但不能同比例提高引物优化率和多态位点率。例如，在棉铃虫中，尽管在序列分析阶段我们已经从微卫星 DNA 文库中剔除了相当多的多拷贝位点（30.7%），但棉铃虫的引物优化率仍是所有建库物种中最低的，造成这一现象的主要原因正是由于很高比例（87.9%）的多拷贝位点的存在。我们的经验表明，在获得阳性

克隆序列并设计引物以后, 尚需通过系统性 PCR 优化来淘汰更多的多拷贝位点(在棉铃虫中为 57.2%), 而且绝大多数阳性克隆最终都会因不适用于分子标记而被淘汰(表 3)。

3.3 三相重复微卫星

本研究表明, 棉铃虫和松毛虫基因组中三相重复微卫星的比例明显高于用相同或类似方法构建文库的其它物种, 而且棉铃虫中三相重复微卫星在最终得到的单拷贝多态性位点中所占的比例要远高于在文库/基因组中所占的比例, 即三相重复微卫星的优化率很高(表 2)。三相重复微卫星在棉铃虫中的较高优化率暗示三相重复微卫星位点是多拷贝位点的可能性远低于其它类型的微卫星位点, 因此, 在以筛选微卫星 DNA 标记为目的的鳞翅目昆虫基因组 DNA 文库构建过程中, 应注意使用三相重复微卫星探针。值得指出的是, 因三相重复微卫星位点的等位基因间长度差别较大, 能够显著降低基因型分析的难度(例如, “影子带”的干扰较轻), 对后续的数据读取、校对和分析十分有利。因而, 如能筛选出尽可能多的三相重复微卫星, 将整体上对实验研究非常有利。所以, 如果经过进一步研究能证明相对偏高的三相重复微卫星丰度和优化率是鳞翅目昆虫微卫星 DNA 的一个共同进化特征, 则会在很大程度上弥补这些物种微卫星分离过程中因其它类型的微卫星丰度低、多拷贝位点比例高所带来的异常困难。

3.4 无效等位基因与微卫星侧翼序列的突变

通常把在基因组中存在却无法通过 PCR 扩增得到产物的等位基因称作无效等位基因, 它们的出现往往与引物结合位点的突变有关。Keyghobadi et al. (1999) 在构建蝴蝶的微卫星文库时发现, 蝴蝶基因组微卫星的侧翼序列有很高的突变率, 使得特异性引物设计和 PCR 扩增的难度加大。在他们所获得的四个微卫星位点中, 无效等位基因频频出现。对其中一个位点的测序分析结果显示, 该位点的侧翼区存在一个 50 bp 的插入或缺失, 一个 1 bp 的插入或缺失以及 5 个点突变。我们所得到的棉铃虫微卫星位点的优化率偏低的结果除与微卫星中较高丰度的多拷贝位点有关外, 与引物侧翼序列的突变速率也有很大关系。我们优化出的 10 个棉铃虫的微卫星位点在种群中都有无效等位基因存在, 其中有两个位点表现尤为突出, 可达 50%, 另有四个位点在 3%~5% 之间。在微卫星位点的优化过程中, PCR 反应的难易程度受引物结合位点的结

构、引物特性、引物与模板 DNA 的退火温度等影响很大, 若引物结合位点的 5' 末端存在突变则会造成引物的 3' 末端最后一个碱基不能与模板结合, 致使引物不能延伸下去, 从而导致 PCR 反应失败。因而 PCR 反应条件优化困难以及无效等位基因的出现都与微卫星侧翼序列的突变有关。理论上对于在 PCR 反应过程中存在较多无效等位基因的单拷贝位点、无产物的位点以及反应条件不理想的位点, 可以考虑重新设计引物加以解决。但我们在其中一个位点微卫星两侧重新设计的引物并不能解决这一问题, 因而推断, 该位点的侧翼序列中可能存在相当多的突变。对一些位点的等位基因的序列分析结果已初步证明了这一推测。

3.5 微卫星 DNA 家族产生的可能机制

我们在棉铃虫和松毛虫中观察到侧翼序列高度相似的微卫星 DNA 家族, 其比率分别为 30.7% 和 13.3%, 而且在棉铃虫的一组微卫星 DNA 家族的 3' 端侧翼序列中存在一段多聚 T/A 序列(图 1)。微卫星 DNA 家族和其它多拷贝微卫星位点已在不同物种中观测到, 例如果蝇(Wilder and Holllocher, 2001), 人(Nadir et al., 1996), 大麦(Ramsay et al., 1999)等。而且还观察到, 这些多拷贝微卫星位点在基因组中常常与一些转座子等可移动遗传因子相关联, 特别是逆转录转座子和逆转录序列(Retrosequences)(Ramsay et al., 1999; Wilder and Holllocher, 2001)。因此, 人们推测基因组中多拷贝微卫星 DNA 的形成和扩散很可能涉及逆转录转座过程(Retrotransposition)或逆转录侵染(Retrofection)(Nadir et al., 1996), 而 3' 端侧翼序列中多聚 T/A 序列的存在是其必然产物(Graur and Li, 2000)。由此我们推测, 棉铃虫一些微卫星 DNA 家族的产生和在基因组中的扩散很可能与逆转录转座过程或逆转录侵染有关, 当然这需要实验进行深入验证。

图 1 所示的 12 个克隆中微卫星序列在重复单位的拷贝数、重复序列的结构以及侧翼序列均有差异, 但可以看出侧翼序列的变异程度远小于微卫星重复序列的变异程度, 这些特征表明: 第一, 该微卫星 DNA 家族在形成以后已经历了较长时间的分化, 各成员间在侧翼序列和微卫星重复单元上均已积累了变异; 第二, 在进化过程中, 微卫星重复单位拷贝数的突变速率远高于碱基组成、特别是侧翼序列的突变速率。因而, 虽然起源于侧翼序列突变的非同源相似在微卫星的进化过程中时有发生, 但在一定

时间尺度内它对种群进化分析的影响并不大，微卫星单元自身极高的突变速率仍使微卫星DNA标记具有其它分子标记所无法替代的应用价值。

本研究表明，微卫星标记系统的建立是一个复杂过程，对鳞翅目昆虫则更是如此。由于其基因组中微卫星DNA丰度较低，特别是大量多拷贝微卫星位点的存在，通过构建基因组DNA文库所筛选出的大多数阳性克隆实际上都不适于引物设计，而且在能够设计引物的位点中，大部分也因不适合于作为多态性遗传标记而被淘汰。因此，绝大多数阳性克隆实际上是不能被用于种群遗传多态性研究的，研究人员对此有必要慎重对待，避免因使用不适用于种群遗传分析的微卫星位点而导致产生隐含错误的数据，从而得出错误的结论。

参考文献 (References)

- Anthony N, Gelembiuk G, Raterman D, Nice C, French-Constant R, 2001. Isolation and characterization of microsatellite markers from the endangered Karner blue butterfly *Lycaeides melissa samuelis* (Lepidoptera). *Hereditas* 134: 271–273.
- Blouin MS, Parsons M, Lacaille V, Lotz S, 1996. Use of microsatellite loci to classify individuals by relatedness. *Molecular Ecology* 5: 393–401.
- Bogdanowicz SM, Mastro VC, Prasher DC, Harrison RG, 1997. Microsatellite DNA variation among Asian and North American gypsy moths (Lepidoptera: Lymantriidae). *Annals of Entomological Society of America* 90: 768–775.
- Bowcock AM, Ruiz-Linares A, Tomfohred J, Minch K, Kidd JR, Cavalli-Sforza LL, 1994. High resolution of human evolutionary tree with polymorphic microsatellites. *Nature* 368: 455–457.
- Caldas A, Hawthorne DJ, Barbosa P, 2002. Isolation and characterization of microsatellite markers from *Zale galbanata* (Lepidoptera: Noctuidae) and amplification in other members of the genus. *Molecular Ecology Notes* 2: 296–297.
- Cassel A, 2002. Characterization of microsatellite loci in *Coenonympha hero* (Lepidoptera: Nymphalidae). *Molecular Ecology Notes* 2: 566–568.
- Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Mil-lasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J, 1996. Comprehensive genetic map of the human genome based on 5, 264 microsatellites. *Nature* 380: 152–154.
- Favre L, Balloux F, Goudet J, Perrin N, 1997. Female-biased dispersal in the monogamous mammal *Crocidura russula*: evidence from field data and microsatellite patterns. *Proceedings of the Royal Society London Series B* 264: 127–132.
- Flanagan NS, Blum MJ, Davison A, Alamo M, Albarrán R, Faulhaber K, Peterson E, McMillan WO, 2002. Characterization of microsatellite loci in neotropical *Heliconius* butterflies. *Molecular Ecology Notes* 2: 398–401.
- Graur D, Li WH, 2000. Fundamentals of Molecular Evolution. 2nd ed. Sunderland: Sinauer Associates Inc., 323–366.
- Harper GL, Piyapattanakorn S, Goilson D, Maclean N, 2000. Isolation of microsatellite markers from the Adonis blue butterfly *Lysandra bellargus*. *Molecular Ecology* 9: 1919–1952.
- Ji YJ, Zhang DX, Hewitt GM, Kang L, Li DM, 2003. Polymorphic microsatellite loci for the cotton bollworm *Helicoverpa armigera* (Lepidoptera: Noctuidae) and some remarks on their isolation. *Molecular Ecology Notes* 3: 102–104.
- Keyghobadi N, Roland J, Stroback C, 1999. Influence of landscape on the population genetic structure of the alpine butterfly *Parnassius smintheus* (Papilionidae). *Molecular Ecology* 8: 1481–1495.
- Koshio C, Tomishima M, Shimizu K, Kim HS, Takenaka O, 2002. Microsatellites in the gypsy moth *Lymantria dispar* L. (Lepidoptera: Lymantriidae). *Applied Entomology and Zoology* 37: 309–312.
- Meglécz E, Pecsenye K, Varga Z, Solignac M, 1998. Comparison of differentiation pattern at allozyme and microsatellite loci in *Parnassius mnemosyne* (Lepidoptera) populations. *Hereditas* 128: 95–103.
- Meglécz E, Solignac M, 1998. Microsatellite loci for *Parnassius mnemosyne* (Lepidoptera). *Hereditas* 128: 179–180.
- Nadir E, Margalit H, Gallily T, BenSasson SA, 1996. Microsatellite spreading in the human genome: Evolutionary mechanisms and structural implications. *Proc. Natl. Acad. Sci. USA* 93: 6470–6475.
- Nède G, Meglécz E, 2000. Microsatellite frequencies in different taxa. *Trends in Ecology & Evolution* 15: 376–377.
- Palo J, Varvio S-L, Hanski I, Vainola R, 1995. Developing microsatellite markers for insect population structure: complex variation in a checker spot butterfly. *Hereditas* 123: 295–300.
- Queller DC, Strassmann JE, Hughes CR, 1993. Microsatellites and kinship. *Trends in Ecology & Evolution* 8: 285–288.
- Ramsay L, Macaulay M, Cardle L, Morgante M, Degli Ivanissevich S, Maestri E, Powell W, Waugh R, 1999. Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. *Plant Journal* 17: 415–425.
- Reddy KD, Abraham EG, Nagaraju J, 1999. Microsatellite in the silk-worm *Bombyx mori*: abundance, polymorphism, and strain characterization. *Genome* 42: 1057–1065.
- Wilder J, Hollocher H, 2001. Mobile elements and the genesis of microsatellites in dipterans. *Molecular Biology & Evolution* 18: 384–392.
- Williams BL, Brawn JD, Paige KN, 2002. Highly polymorphic microsatellite loci for *Speyeria idalia* (Lepidoptera: Nymphalidae). *Molecular Ecology Notes* 2: 87–88.
- Zhang DX, Hewitt GM, 1998. Isolation of animal cellular total DNA. In: Karp A, Isaac PG, Ingram DS ed. *Molecular Tools for Screening Biodiversity: Plants and Animals*. London: Chapman & Hall, 5–9.
- Zhang DX, Hewitt GM, 2003. Nuclear DNA analysis in genetic studies of populations: practice, problems and prospects. *Molecular Ecology* 12: 563–584.
- Zhang DX, Smith JA, Rico C, Hewitt GM, 2001. Isolation and characterization of 10 microsatellite loci in poor cod *Trisopterus minutus* (L.). *Molecular Ecology Notes* 1: 50–52.
- Zhang DX, Yan LN, Ji YJ, Kang L, Hewitt GM, Huang ZS, 2003. Isolation, characterization and cross-species amplification of eight microsatellite DNA loci in the migratory locust *Locusta migratoria*. *Molecular Ecology Notes* 3: 483–486.