研究报告 Research Reports

# 基于转录组数据的印度谷螟微卫星位点分析

唐培安1, 陶冶心1, 薛 吴1, 袁明龙2

(1. 南京财经大学食品科学与工程学院/江苏省现代粮食流通与安全协同创新中心,南京 210023;2. 草地农业生态系统国家重点实验室/兰州大学草地农业科技学院,兰州 730020)

摘要 基于印度谷螟转录组测序数据筛选其功能微卫星(EST-SSR)分子标记,设计该虫的微卫星(SSR)引物并验证其适用性。用 MicroSAtellite 微卫星搜索软件查找印度谷螟转录组中微卫星的数量、重复次数以及所有微卫星的位置信息,采用 Primer Premier 5 软件设计 SSR 引物,并进行 PCR 验证。结果表明含 EST-SSR 的序列 3 173 个, 占总搜索序列的 8.52%,平均每 13 kb 中就出现 1 个 EST-SSR。共发现 192 种微卫星类型,其中重复单元为单碱 基、二碱基和三碱基的比例较高,依次是 28.21%、39.84%、25.43%。除单碱基重复单元外,出现频率最高的是 AT/AT,有 593 次。在 66 对印度谷螟 EST-SSR 引物中,有 28 对 PCR 扩增成功。这说明基于印度谷螟转录组数据开 发微卫星标记是可行的,本研究获得的印度谷螟 EST-SSR 位点为今后开展该虫的种群遗传学研究提供了基础数据。 关键词 转录组; 高通量测序; 螟蛾科; 微卫星; 分子标记 中图分类号; Q 968.1 文献标识码; A DOI; 10.3969/j. issn. 0529-1542. 2017. 03.007

# Analysis of microsatellite loci in *Plodia interpunctella* based on transcriptome dataset

Tang Peian<sup>1</sup>, Tao Yexin<sup>1</sup>, Xue Hao<sup>1</sup>, Yuan Minglong<sup>2</sup>

(1. College of Food Science and Engineering/Collaborative Innovation Center for Modern Grain Circulation and Safety, Nanjing University of Finance and Economics, Nanjing 210023, China; 2. State Key Laboratory of Grassland Agro-Ecosystems/College of Pastoral Agricultural Science and Technology, Lanzhou University, Lanzhou 730020, China)

Abstract This study is aimed to identify gene microsatellite (EST-SSR) loci from the transcriptome database of *Plodia interpunctella*. Based on these SSR sequences, the selected SSR primer pairs were validated. The number, repetition time and location information of all microsatellites obtained with the microsatellite search tool Micro-SAtellite were analyzed. Primer Premier 5 was used to design *P. interpunctella* SSR primers, and then these primer pairs were verified by PCR. A total of 3 173 SSRs were identified in 37 246 Unigenes, with one SSR per 13 kb. The majority of microsatellite loci consisted of mono-, di- and tri-nucleotide motifs (28.21%, 39.84% and 25. 43%, respectively). However, different types of repeat SSRs had considerably different distribution. Except mono-nucleotide, AT/AT was the most frequent repeat motif (593 repeats) among all SSR motifs. Among the 66 designed primer pairs, 28 pairs were amplified successfully. Our study shows that it is feasible to develop microsatellite markers based on *P. interpunctella* transcriptome. The EST-SSRs of *P. interpunctella* obtained in this study are helpful for population genetics studies of this important stored-product pest.

Key words transcriptome; high-throughput sequencing; Pyralidae; microsatellite; molecular marker

微卫星(microsatellite)又称 SSR(simple sequence repeat,简单重复序列),是指以 1~6个核苷酸为重复单元串联组成的重复序列<sup>[1]</sup>,1981年,微卫 星首次被发现,它们大都以 2~3个碱基重复类型为主, 主要分布于原核生物和真核生物的基因组中<sup>[2-3]</sup>。微 卫星具有高多态性、多等位性、共显性、高可重复性、数 量丰富和对基因组有很好覆盖性等特征<sup>[4-5]</sup>。因此,该 标记已被广泛应用于生物遗传研究中<sup>[6]</sup>。

**收稿日期**: 2016-06-22 修订日期: 2016-08-02

基金项目: 粮食公益性行业科研专项(201413007-2,201513002-5-3);国家重点研发计划专项(2016YFD0401004-04);江苏省高校优势学 科建设工程 联系方式 E-mail:tangpeian@163.com

印度谷螟 Plodia interpunctella 属鳞翅目 Lepidoptera 螟蛾科 Pyralidae,是世界性分布的重大储粮害虫。 该虫在我国南北方广泛分布[7],依据2004-2005年第6 次全国储粮昆虫调查,印度谷螟被列为危害严重日分 布广泛的主要鳞翅目储粮害虫之一<sup>[8]</sup>。目前,关于印度 谷螟 SSR 的研究还未见报道,在 GenBank 中也没有发 现登录的印度谷螟 SSR 序列,不能满足其种群遗传结 构研究的需要。因此,开发印度谷螟的微卫星位点,是 有效开展其种群遗传结构和基因交流、系统发育与分 子进化等方面研究的重要前提和首要工作,同时对于 揭示其危害规律和综合防治等都具有重要意义。传统 开发 SSR 标记的方法需要构建文库、筛选 SSR 克隆、测 序等复杂且工作量大的步骤[9-11]。随着新一代高通量 测序技术的成长,尤其是转录组测序的飞速发展,降低 了微卫星分子标记的开发成本,提高了开发效率,为非 模式物种 SSR 位点的大批量开发提供了一种经济、高 效的途径[12-13]。从转录组数据中获得的与功能基因相 关SSR 位点又叫表达序列标签 SSR(expressed sequence tag SSR,EST-SSR)。目前,从昆虫转录组数据库中筛 选功能 EST-SSR 标记已有许多成功的报道<sup>[14-17]</sup>。本 研究基于已测印度谷螟转录组数据,挖掘其 EST-SSR 位点,分析这些微卫星位点的类型和分布特点,探讨鉴 定到的 EST-SSR 位点的适用性,以期获得可供该虫种 群遗传学研究所需的核分子标记,也为印度谷螟近缘 物种微卫星位点的发掘提供可借鉴的方法。

# 1 材料与方法

# 1.1 材料及转录组数据来源

印度谷螟采自江苏吴江,以大米粉为饲料,在温 度为( $30\pm1$ )℃、湿度为 75%±5%、24 h 无光照条 件的恒温培养箱中培养 10 代以上,选取 4 龄幼虫作 为试验材料。采用德国 Qiagen 公司的 RNeasy Plus Universal Mini Kit 试剂盒提取 RNA,由华大科技进 行高通量转录组测序(转录组数据的 NCBI 登录号: SRP060836)。通过 Illumina Hiseq2000 平台测序,总 计产出 5.7 GB 数据,共组装 Unigene 37 246 个,总长 49 682 542 bp,平均长度 1 334 bp,N50 达到 2 307 bp, 过滤后质量不低于 20 的碱基比例(Q20,测序错误率 <1%)为 98. 31%,表明测序结果良好。

# 1.2 微卫星序列筛选

以组装出来的 Unigene 作为参考序列,使用 Perl 操作平台下的 SSR 分析软件 MicroSAtellite (MISA) (http://pgrc.ipkgatersleben.de/misa/)在 Unigene 中搜 索 SSR 位点。筛选条件为:单碱基重复次数≥12,二碱 基重复次数≥6,三碱基重复次数≥5,四、五、六碱基重 复次数≥4,并且两个相邻微卫星重复单元之间的最小 长度设为100 bp。

# 1.3 印度谷螟 SSR 引物设计及 PCR 扩增

基于 SSR 重复单元两翼序列设计引物,采用 Primer 5.0 对印度谷螟 EST-SSR 设计了 66 对引 物。引物筛选条件如下:①引物不含有重复序列;② 将获得的引物比对到 Unigene 序列,引物 5′端允许 有 3 个碱基的错配,而 3′端允许有 1 个碱基的错配; ③去除比对到非目标 Unigene 的引物,保证引物与 Unigene 唯一匹配。④使用 ssr\_finder 校验 SSR,使 用产物序列来寻找 SSR,检验结果是否与 MISA 结 果相同,并筛选出相同的 SSR 产物。设计好的引物 由金斯瑞生物技术公司(江苏南京)合成。

采用 TIANamp Genomic DNA Kit(北京天根) 试剂盒提取单头印度谷螟 4 龄幼虫的基因组 DNA。 PCR 扩增体系:基因组 DNA 1  $\mu$ L,上游引物 2  $\mu$ L,下 游引物 2  $\mu$ L,灭菌水 7.5  $\mu$ L,ExTaq Mix 12.5  $\mu$ L。 PCR 反应条件:94℃预变性 3 min,再进行 30 个循 环,每个循环包括 94℃变性 30 s,Tm 复性 30 s, 72 ℃延伸 1 min;最后 72 ℃延伸 10 min。PCR 产 物用 2%的琼脂糖凝胶电泳检测。

# 2 结果与分析

从构建的印度谷螟转录组数据库中,获得 EST-SSR 序列 3 173 个,占总搜索基因的 8.52%,平均每 13 380 bp 中就出现 1 个 SSR 位点。

#### 2.1 微卫星数量及分布特征

单碱基、二碱基、三碱基、四碱基、五碱基和六碱 基重复的 EST-SSR 数分别为 895、1 264、807、100、 28、79,依次占总 SSR 个数的 28.21%、39.84%、25.43%、 3.15%、0.88%、2.49%(表 1)。可见,在印度谷螟 基因数据库中二碱基重复序列占据优势,其次是单 碱基和三碱基重复序列。此外,复合型 SSR 有 72 个,占 SSR 总数的 2.27%。在单碱基、二碱基、三碱 基、四碱基、五碱基、六碱基 SSR 中出现频率最多的 分别是 A/T、AT/AT、CCG/CGG、AAAT/ATTT、ACG-CGC/GCGCGT(图 1),在各自重复单元类型中所占比 例依次为 98.55%、46.91%、36.31%、33.00%、17.86%、 12.66%。在五碱基和六碱基中各类重复基元分布 较为平均,这是由于此类重复较少的原因。

单碱基重复次数集中在 12~15 次, 二碱基重复 次数集中在 6~8 次, 三碱基重复次数集中在 5~8 次, 四碱基重复次数集中于 5~6 次, 五碱基和六碱 基则以重复 4 次为主。且随着碱基重复单元的重复 次数增加, 其数量均呈下降趋势。

%

表1 印度谷螟转录组中不同微卫星类型分布												
Table 1      Distribution of different SSR types in the Plodia interpunctella transcriptome												
重复类型				重复社	欠数 Nu	mber of r	epeats				总数	比例/%
Repeat type	4	5	6	7	8	9	10	11	12	> 12	Total	Ratio
单碱基 Mononucleotide	0	0	0	0	0	0	0	0	459	436	895	28.21
二碱基 Dinucleotide	0	0	631	270	132	81	68	61	21	0	$1 \ 264$	39.84
三碱基 Trinucleotide	0	556	160	71	20	0	0	0	0	0	807	25.43
四碱基 Tetranucleotide	0	70	30	0	0	0	0	0	0	0	100	3.15
五碱基 Pentanucleotide	24	4	0	0	0	0	0	0	0	0	28	0.88
六碱基 Hexanucleotide	79	0	0	0	0	0	0	0	0	0	79	2.49
总数 Total	103	630	821	341	152	81	68	61	480	436	3 173	
比例/% Ratio	2.77	16.97	22.11	9.18	4.09	2.18	1.83	1.64	12.93	11.74		



图 1 印度谷螟转录组不同碱基类型微卫星位点分布



# 2.2 微卫星长度分布特征

印度谷螟转录组中发现的 3 173 个 EST-SSR 长 度从12~192个碱基不等,平均长度为16.24个碱基。 序列长度 12~15 个碱基的有 2 175 个,占总 SSR 个 数的 68.5%,其中单碱基、二碱基和三碱基重复单元 所占比例分别为 35.54%、39.72%、24.74%,分布较 为平均。序列长度 16~20 个碱基有 560 个,所占比例 为17.6%,其中二碱基重复单元比例最高(48.04%),其 次是三碱基重复单元(27.50%)。序列长度 21~25 个碱基的有 297 个,占总 EST-SSR 的 9.3%,其中 三碱基重复单元占 29.97%,其次是二碱基重复单 元(27.27%),再次是六碱基重复单元(24.58%),其 他重复单元数量较少。序列长度大于 25 的数量较 少,仅占总 EST-SSR 的 4.4%,且全部是复合型(图 2)。



①: SSR序列长度12~15个碱基; ②: SSR序列长度16~20个碱基; ③: SSR 序列长度21~25个碱基; ④: SSR序列长度大于25个碱基 (1): SSR of 12-15 nt in length; (2): SSR of 16-20 nt in length; (3): SSR of 21-25 nt in length; (4): SSR of longer than 25 nt in length

印度谷螟转录组中 EST-SSR 的长度分布 图 2 Fig. 2 Length distribution of EST-SSRs in the Plodia interpunctella transcriptome

#### 2.3 微卫星在转录组编码区分布特征

印度谷螟转录组数据中,有 23 310 个 Unigene 序列功能注释成功。在 23 310 个注释成功的 Unigene 序列中发现 EST-SSR 位点 2 226 个,总长度 13 182 617 bp,其中编码区 858 个 EST-SSR 位点。编 码区 EST-SSR 中二碱基重复 319 个,三碱基重复 378 个,在非编码区二碱基重复 517 个,三碱基重复 291 个。说明在编码区以三碱基重复为主,非编码 区以二碱基重复为主。

#### 2.4 微卫星引物设计与 PCR 扩增

根据印度谷螟可用于引物设计的 EST-SSR 序 列共随机设计了 66 对引物,部分引物信息见表 2。 以单头印度谷螟 4 龄幼虫 DNA 为模板,进行 PCR 扩增,结果表明 66 对引物中的 31 对引物可以扩增 得到条带(图 3)。其中两对引物扩增产物与预期相 差较大,一对引物产生非特异性条带,其余 28 对引 物与预期相符。

	Table 2	Information for 31 primer pairs of EST-SSRs in Plodia interpunctella								
SSR 名称 SSR name	重复单元(重复数) Repeat unit (Number of repeats)	上游引物 Forward primer	下游引物 Reverse primer	产物大小/bp Product size						
SSR1	CT(12)	GGTGTATGGGATTAGAACCCTTT	TATCAGAGTGCGCTCATAGAACA	147						
SSR2	TA(10)	ATATTCGTAAGCTTGCCGTTTTT	ATGTGGATTGATTTTCATTGGAC	157						
SSR3	AC(11)	TTGGTAGACTTTGCGTGAAAGAG	CAATACATGCGAAAGTTTGTGTG	101						
SSR4	TCGCAC(4)	CTTGCGCATCACCAGTCTC	CAAGGACAAGCTAGACAAGTTGC	124						
SSR5	TGC(8)	CTACGGTCGACACAGACTCG	AGTTACTACTACGCCGCCAGTTA	137						
SSR6	TA(10)	AAGTGTTGCCATGCTAACGTAAT	TCTGGCAACTTTTATATCAACCA	145						
SSR7	AT(11)	TCAGAATTACATTTCAAGTTTTTAAGGG	CAGTTTCCCATACTTACCACCAA	158						
SSR8	AT(10)	GAGTTTCAAATTGCACCCAATAG	ATCAGTTTTTGAAGTGAATGCGT	160						
SSR9	CTG(8)	ACTGGTACCACTGCTGCG	GTAACTGGGGCAACTGGAAC	138						
SSR10	CCCGCT(4)	GTCAAAGTCCAAGTCCGTGTC	GATTTAGATTTCGAGCGGGAG	135						
SSR11	CATA(6)	ATCTAAAAACCATATCGTTCCCG	AAAAATCATTTGCTGAACTGTTG	147						
SSR12	ATAAC(5)	CCATCTTCGAATAGAAGTTTGGA	GTCCGCAATGATGAAGATGTAAC	160						
SSR13	GCTGT(4)	CAGTGCTACAGCCTCACGTC	TCGTATTTTTTCTATTATGTCGAGGG	159						
SSR14	TC(10)	AGCGCAAATGATAGAGAAGGAAT	CGAATGCTCGAAATTTCTTATGAC	151						
SSR15	TC(10)	GTTTTTACATTTGCATCCTCAGC	GATGTCTCTAACGAGCTCCTTTTC	136						
SSR16	AT(10)	AGCCACAAGGGAGATTTCTTATC	CATGATTGAGAAGCAAGATGTTTT	150						
SSR17	TC(11)	ACGCGTAGTCTTAGCTTGTGTTC	TTTAACGAGAATGTAGCCGGTAG	129						
SSR18	TTATA(4)	TCAAGCAAGTTTTCATCCAGTTT	TGTCACTTGTCATCTCCCAGTAA	127						
SSR19	CA(11)	CACTACATCCACGAAATACACCC	TGCCAGAACATATGAGAATCAAA	159						
SSR20	TA(10)	GCATGTTTGCGTACAACCAGTAT	TGCTTTGCTCGTGTATGAGAAT	154						
SSR21	AC(10)	GCGTACATGTTTAGTGCACGAGT	AGTCGTGAGCAATAAGACAGCTT	149						
SSR22	AT(10)	ATACGTTTGAGGGGTTACTCCAT	GAGTTGTTGCTTATCAATGCGAT	151						
SSR23	TA(10)	ACATTGCGATTATTTGGATGTGT	CAACCTCACTAACGGCATACATT	150						
SSR24	AG(11)	TTCTAGTATGAGCGAAACGAAAGA	GGTAATGAATGTGAGTTGTTTATGTCA	108						
SSR25	AATAA(4)	AGGTAGAACTGATAGGTAGCCGC	TAATTTCCTGGTTGGATTTGGAT	141						
SSR26	AT(10)	AATGCCATACGATAGTTTTTGTG	AAAATGTTCCAAATAAGGAAGGC	147						
SSR27	AT(10)	TCAGGAAAGTGAATCAGTGTGAA	TCGAGATAAATTCACGAAAATGA	116						
SSR28	TC(10)	TGTCAATTTGGAAATCACCTCAT	GCGAATTGTGATCCTATTGTTGT	117						
SSR29	ATAAC(5)	CCATCTTCGAATAGAAGTTTGGA	GTCCGCAATGATGAAGATGTAA	160						
SSR30	ACA(8)	GAAGAAGCACATGTTAACGCATA	CTGCATCTGCTGGTGTTGTT	115						
SSR31	CAA(8)	GAGTTCCCAACAATAAGGTTTCC	GCCATAAGAGGAGGAATTTGATT	124						

#### 表 2 印度谷螟 31 对 EST-SSR 引物信息





M: 分子量标准500 bp Marker; 1~31泳道的SSR依次为表2中SSR顺序 M: Molecular weight standard 500 bp Marker; SSRs in lane 1-31 were in the same order as in table 2

图 3 印度谷螟 31 个 EST-SSR 位点的扩增电泳图

Fig. 3 Agarose gel showing the amplification of 31 EST-SSR loci in the Plodia interpunctella transcriptome

### 3 讨论

本研究通过生物信息学方法从印度谷螟转录组 数据库 37 246 条 Unigenes 中共获得 EST-SSR 位点 3 173个,其出现频率(8.52%)高于大多数的昆虫,如 扶桑绵粉蚧 Phenacoccus solenopsis(6.33%)<sup>[18]</sup>,烟粉 虱 Bemisia tabaci(5.07%)<sup>[19]</sup>,橘小实蝇 Bactrocera dorsalis(4.23%)<sup>[11]</sup>,黄粉虫 Tenebrio molitor(1.67%)<sup>[20]</sup>, 云南切梢小蠹 Tomicus yunnanensis(1.29%)<sup>[21]</sup>和 葱地种蝇 Delia antiqua(1.12%)<sup>[22]</sup>等;而目前仅有 少数昆虫的 EST-SSR 位点出现频率高于印度谷螟, 如黑翅土白蚁 Odontotermes formosanus(9.98%)<sup>[23]</sup>。 导致这种差异的原因,可能与 EST-SSR 筛选的标准 不同以及用于测序的 RNA 质量相关。当然,最本 质的原因应是物种本身的差异所致。

本研究表明,印度谷螟转录组中微卫星的种类较为丰富,1~6核苷酸重复类型都有出现,主要重复类型为1~3核苷酸重复,占SSR总数的93.48%。在单碱基重复基元中,A/T是占优势的重复基元,这与已研究的大多数昆虫类似,如褐飞虱 Nilaparvata lugens<sup>[24]</sup>、黏虫 Mythimna separata<sup>[15]</sup>和扶桑绵粉蚧 P. solenopsis<sup>[18]</sup>等。研究发现,基于转录组数据库发掘的昆虫微卫星以三碱基重复 SSR 出现的频率最高,如二点委夜蛾 Athetis lepigone<sup>[25]</sup>、灰飞虱 Laodel phax striatellus<sup>[26]</sup>、橘小实蝇 B. dorsalis<sup>[11]</sup>、黄粉虫 T. molitor<sup>[20]</sup>和温带臭虫 Cimex lectularius<sup>[16]</sup>等。普遍的观点认为,三碱基重复序列占据主导地位,是由于三联体密码子的选择作用

所致<sup>[27]</sup>。因为密码子的基本单位就是三碱基,其他 类型的重复单元会引起阅读框的改变,增加错配或 者移码突变的机会。本研究发现,印度谷螟编码区 中三碱基重复也为优势重复单元,支持该论点。然 而,在印度谷螟的所有 EST-SSR 中最优势重复单元 为二碱基重复,这与丽蝇蛹集金小蜂 Nasonia vitripennis、N. giraulti 和 N. longicornis<sup>[28]</sup>、黑腹果 蝇 Drosophila melanogaster<sup>[29]</sup>和细梢小卷蛾 Rhyacionia leptotubula<sup>[17]</sup>等类似。以往对昆虫 EST-SSR 的研究中,以 GC/GC 为核心基元的两碱基重 复 SSR 的存在非常稀少,几乎为零。但是在印度谷 螟二碱基重复单元中 GC/GC 含量占了 1/3 左右, 仅仅低于重复单元 AT/AT,这与同属于鳞翅目的 黏虫<sup>[15]</sup>、二点委夜蛾<sup>[25]</sup>和细梢小卷蛾<sup>[17]</sup>相似。因 此,我们推断 GC/GC 重复基元可能与鳞翅目的某 些特殊功能有关系。此外,研究表明,SSR 的多态性 与其长度以及核心基元的重复次数的多少具有相关 性,即长度越长、核心基元的重复次数越多,其多态 性越高<sup>[30]</sup>。印度谷螟的 EST-SSR 长度集中于 12~ 15个碱基占据了70%,而且还有12%(多为二、三 碱基重复单元)长度大于 20 个碱基。因此,印度谷 螟可能具有较高的遗传多态性,这也是该虫高度适 应环境而成为世界性害虫的遗传基础。

众所周知,鳞翅目昆虫 SSR 开发难度较大,由 于其 SSR 的多态性较低、存在微卫星家族现象并有 大量的无效等位基因<sup>[31]</sup>。为得到质量较高且具有 潜在多态性的引物,在筛选引物时应尽量选择侧翼 长度大于 30 bp 的位点,且相似度低于 70%,核心重 植物保护

复单元重复次数二碱基在 10 次以上, 三碱基应在 8 次以上, 可降低家族现象的影响并且避免非特异性 扩增<sup>[11]</sup>。本文利用转录组数据, 设计得到的 66 对 EST-SSR 引物, 28 对扩增得到预期大小的条带, 占 合成引物的 42.42%。尽管这些位点的多态性还有 待进一步检验, 但其数量足够开展印度谷螟种群遗 传学等相关研究。导致 EST-SSR 位点扩增失败有 诸多原因: 扩增得到条带远大于预期, 可能由于基因 组 DNA 中含有内含子而在转录后未出现, 干扰了 扩增过程; 而其中一对引物产生非特异性条带可能 是由于 EST-SSR 位点位于同源基因序列上; 扩增失 败无条带, 可能是因为 EST-SSR 在基因组中的表达 丰度较低而导致产物浓度很低, 未能检测到。

目前,SSR研究主要应用于遗传学分析。对于 非模式昆虫的研究其难点在于 SSR 引物的筛洗,因 为在昆虫等节肢动物基因组内 SSR 的含量较少,且 存在家族现象,这加大了研究难度。传统微卫星筛 选方法费时费力,且费用高昂。转录组测序技术的 发展解决了这一难题,从转录组测序结果中可以直 接筛选出质量合格的 SSR,经过验证其在相近物种 间具有较高的通用性。随着转录组测序技术的进一 步发展,测序通量加大,成本降低,这为昆虫 SSR 的 大规模开发提供了可能,目 SSR 所在基因的表达水 平也可以监测到,为深入研究 SSR 在基因调控和修 饰中所产生的作用提供了技术保障<sup>[32]</sup>。但基于转 录组测序获得的 EST-SSR 位点,其核心重复次数相 较于基因组的 SSR 少,这是因为在组装过程中重复 次数多的 SSR 会发生断裂或者丢失,这影响了 EST-SSR 位点的多态性。另外, cDNA 中缺少内含 子或者内含子位于 SSR 引物之间,这就导致部分 EST-SSR 位点的扩增失败。另一方面,基于转录组 测序数据发掘 EST-SSR 位点具有快速、方便、低廉 以及数据量大等优点,已成为当前发掘昆虫微卫星 位点的重要策略。

综上所述,本研究基于印度谷螟转录组数据筛 选获得了大量 EST-SSR 位点,并对其数量分布、相 关特性及适用性进行了分析。结果表明,基于转录 组数据发掘印度谷螟微卫星位点是高效可行的,也 是发掘非模式生物 EST-SSR 位点的一种经济、快 速、有效的方法。本研究获得的印度谷螟 EST-SSR 位点,为今后开展该虫的遗传多样性及种群遗传结 构研究,阐明其种群遗传分化及抗药性基因流动规律,以及制定有效防控策略等提供了基础数据。

#### 参考文献

- [1] 孙荆涛,杨现明,葛成,等. 微卫星分子标记在昆虫分子生态学研 究上的应用[J]. 南京农业大学学报,2012,35(5):103-112.
- [2] Spritz R A. Duplication/deletion polymorphism 5'-to the human β globin gene [J]. Nucleic Acids Research, 1981, 9(19): 5037 -5048.
- [3] Selkoe K A, Toonen R J. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers[J]. Ecology Letters, 2006,9(5): 615-629.
- [4] Sandhu K S, Karaoglu H, Zhang P, et al. Simple sequence repeat markers support the presence of a single genotype of *Puccinia psidii* in Australia[J]. Plant Pathology, 2016, 65 (7): 1084 - 1094.
- [5] Madesis P, Ganopoulos I, Tsaftaris A. Microsatellites: evolution and contribution [M]//Walker J M. Methods in molecular biology, USA: Humana Press, 2013: 1-13.
- [6] Hadonou A M, Sargent D J, Wilson F, et al. Development of microsatellite markers in *Fragaria*, their use in genetic diversity analysis, and their potential for genetic linkage mapping [J]. Genome, 2004, 47(3): 429 - 438.
- [7] 简富明.印度谷螟生物学特性初步研究[J].西南农业学报, 1993,6(3):80-84.
- [8] 严晓平,宋永成,沈兆鹏,等.中国储粮昆虫 2005 年最新名录 [J].粮食储藏, 2006,35(2):3-9.
- [9] Andrés J A, Bogdanowicz S M. Isolating microsatellite loci: looking back, looking ahead [J]. Molecular Methods for Evolutionary Genetics, 2011, 772: 211 – 232.
- [10] Zane L, Bargelloni L, Patarnello T. Strategies for microsatellite isolation: a review [J]. Molecular Ecology, 2002, 11(1): 1-16.
- [11] 魏丹丹,石俊霞,张夏瑄,等. 基于转录组数据的桔小实蝇微卫星 位点信息分析[J]. 应用生态学报,2014,25(6):1799-1805.
- [12] Schoebel C N, Brodbeck S, Buehler D, et al. Lessons learned from microsatellite development for nonmodel organisms using 454 pyrosequencing [J]. Journal of Evolutionary Biology, 2013, 26 (3): 600-611.
- [13] Fernandez-Silva I, Whitney J, Wainwright B, et al. Microsatellites for next-generation ecologists: a post-sequencing bioinformatics pipeline [J]. PLoS ONE, 2013, 8 (2): e55990.
- [14] Comeault A A, Sommers M, Schwander T, et al. *De novo* characterization of the *Timema cristinae* transcriptome facilitates marker discovery and inference of genetic divergence [J]. Molecular Ecology Resources, 2012, 12 (3): 549-561.
- [15] 胡艳华,李敏,张虎芳,等.粘虫转录组中SSR位点的信息分析
  [J].山西农业大学学报:自然科学版,2015,35(5):484-489.
  (下转75页)

科)[J]. 动物分类学报, 2012, 37(4): 894-898.

- [4] 师国荣,师增胜. 旱地菊芋优质高产栽培技术[J]. 农业科技与 信息, 2010(1): 18.
- [5] 于静,田喜颖. 吉林省菊芋栽培技术[J]. 特种经济动植物, 2015(10): 38-39.
- [6] 任玉敏. 多用途植物菊芋的栽培及利用探讨[J]. 现代农村科 技, 2015(21): 14-15.
- [7] 张云慧,程登发.突发性暴发性害虫监测预警研究进展[J]. 植物保护,2013,39(5):55-61.
- [8] 吴孔明,陆宴辉,王振营. 我国农业害虫综合防治研究现状与 展望[J]. 昆虫知识, 2009,46(6): 831-836.
- [9] 蔡小娜,韩旭,沈佐锐,等. 基于蛾翅翅脉特征的夜蛾昆虫数字 化分类研究(鳞翅目:夜蛾科)[J]. 环境昆虫学报,2016,38
   (2):348-353.
- [10] 杨红珍,沈佐锐,李湘涛. 昆虫自动鉴定技术研究与展望[J]. 四川动物,2011,30(5): 834-838.
- [11] 陶丹,王峥荣,李光林,等. 基于解模糊算法的蚕蛹图像恢复及 雌雄识别[J]. 农业工程学报, 2016,32(16): 168-174.

- [12] 于新文,沈佐锐,高灵旺,等. 昆虫图像几何形状特征的提取技术研究[J]. 中国农业大学学报,2003,8(3):47-50.
- [13] Chown S L, Gaston K J. Body size variation in insects: a macroecological perspective [J]. Biological Reviews, 2010, 85(1):139 - 169.
- [14] 周隆,文礼章,王少丽,等. 不同寄主种群 B 型烟粉虱体内营养 和成虫体长差异比较[J]. 植物保护, 2011,37(4):87-90.
- [15] 赵汗青,沈佐锐,于新文. 数学形态学在昆虫分类学上的应用研究. []. 在总科阶元上的应用研究[J]. 昆虫学报,2003,46(2):201-208.
- [16] 赵汗青,沈佐锐,于新文. 数学形态学在昆虫分类学上的应用 研究. I. 在目级阶元上的应用研究[J]. 昆虫学报, 2003,46 (1):45-50.
- [17] 沈佐锐,赵汗青,于新文. 数学形态学在昆虫分类学上的应用研究. Ⅲ. 在科阶元上的应用研究[J]. 昆虫学报, 2003,46(3): 339-344.
- [18] 蔡小娜,黄大庄,沈佐锐,等.利用翅的数学形态特征对蛾类昆 虫进行分类鉴定的系统研究Ⅰ——在总科级阶元上的应用 [J].中国农业大学学报,2013,18(4):96-104.

(责任编辑:田 喆)

(上接 48 页)

- [16] Bai X, Mamidala P, Rajarapu S P, et al. Transcriptomics of the bed bug (*Cimex lectularius*)[J]. PLoS ONE, 2011, 6 (1): e16336.
- [17] Zhu J Y, Li Y H, Yang S, et al. *De novo* assembly and characterization of the global transcriptome for *Rhyacionia leptotubula* using illumina paired-end sequencing[J]. PLoS ONE, 2013, 8 (11): e81096.
- [18] 罗梅,张鹤,宾淑英,等. 基于转录组数据高通量发掘扶桑绵 粉蚧微卫星引物[J]. 昆虫学报,2014,57(4):395-400.
- [19] Xie W, Meng Q S, Wu Q J, et al. Pyrosequencing the *Bemisia tabaci* transcriptome reveals a highly diverse bacterial community and a robust system for insecticide resistance [J]. PLoS ONE, 2012, 7 (4); e35181.
- [20] 朱家颖,吴国星,杨斌. 基于转录组数据高通量发掘黄粉甲微 卫星引物[J]. 昆虫学报,2013,56(7):724-728.
- [21] 袁远,张丽芳,吴国星,等. 云南切梢小蠹微卫星的高通量发掘 [J]. 环境昆虫学报,2014,36(2):166-170.
- [22] Zhang Yujuan, Hao Youjin, Si Fengling, et al. The *de novo* transcriptome and its analysis in the worldwide vegetable pest, *Delia antiqua* (Diptera: Anthomyiidae)[J]. G3: Genes | Genomes | Genetics, 2014, 4 (5): 851-859.
- [23] Huang Q Y, Sun P D, Zhou X G, et al. Characterization of head transcriptome and analysis of gene expression involved in caste differentiation and aggression in *Odontotermes formosanus* (Shiraki)[J]. PLoS ONE, 2012, 7 (11): e50383.
- [24] 刘玉娣, 侯茂林. 褐飞虱 EST 资源的微卫星信息分析[J]. 昆 虫学报, 2010, 53(3): 239-247.

- [25] Li L T, Zhu Y B, Ma J F, et al. An analysis of the Athetis lepigone transcriptome from four developmental stages [J]. PLoS ONE, 2013, 8 (9): e73911.
- [26] Zhang F G, Guo H Y, Zheng H J, et al. Massively parallel pyrosequencing-based transcriptome analyses of small brown planthopper (*Laodelphax striatellus*), a vector insect transmitting rice stripe virus (RSV) [J]. BMC Genomics, 2010, 11 (1): 303.
- [27] 潘海涛, 汪俊君, 王盈盈, 等. 小麦 EST-SSR 标记的开发和遗传作图[J]. 中国农业科学, 2010, 43 (3): 452-461.
- [28] Pannebakker B A, Niehuis O, Hedley A, et al. The distribution of microsatellites in the Nasonia parasitoid wasp genome
  [J]. Insect Molecular Biology, 2010, 19(S1): 91-98.
- [29] Schug M D, Wetterstrand K A, Gaudette M S, et al. The distribution and frequency of microsatellite loci in *Drosophila mel*anogaster [J]. Molecular Ecology, 1998, 7(1): 57 - 70.
- [30] Meglécz E, Nève G, Biffin E, et al. Breakdown of phylogenetic signal: a survey of microsatellite densities in 454 shotgun sequences from 154 non model eukaryote species[J]. PLoS ONE, 2012, 7 (7): e40861.
- [31] Zhang D X. Lepidopteran microsatellite DNA: redundant but promising [J]. Trends in Ecology & Evolution, 2004, 19 (10): 507 – 509.
- [32] Jhanwar S, Priya P, Garg R, et al. Transcriptome sequencing of wild chickpea as a rich resource for marker development [J]. Plant Biotechnology Journal, 2012, 10 (6): 690 - 702.