

星天牛转录组 SSR 位点特征分析*

韩小红** 王伊凡 卢赐鼎 林浩宇 是雨霏 何欢 张飞萍 梁光红***

(福建农林大学林学院, 生态公益林重大有害生物防控福建省高校重点实验室, 福州 350002)

摘要 【目的】为了获得星天牛 *Anoplophora chinensis* 的 SSR 位点信息并开发其 SSR 分子标记技术, 进一步为其遗传多样性以及综合治理提供理论依据。【方法】利用 MISA 软件, 对星天牛转录组数据进行简单重复序列 (SSR) 位点筛选与分析; 使用 Primer 3 软件设计引物, 采用 PCR 扩增以及电泳检测, 筛选 SSR 引物, 开发星天牛 SSR 分子标记技术。【结果】在 9 325 条 unigene 序列中共挖掘到 2 360 个 SSR 位点, 出现频率为 25.31%, 涉及 SSR 位点序列 1 758 条, 发生频率为 18.85%。星天牛转录组中 SSR 的主要重复类型为单碱基重复, 其次是三碱基重复, 分别占总数的 79.03%、12.54%。在核苷酸重复类型中, A/T 基元种类数目最多, 所占比例高达 99.30%。SSR 长度为 10-11 bp 的占比最高, 为 56.10%; 重复次数为 10 次的数量最多, SSR 位点数为 1 188 (50.34%)。重复次数和长度的分析结果对 SSR 位点的多态性获得了初步验证。在随机挑选序列设计的 60 对引物中, 53 对扩增产物达到预期大小, 候选引物可用率高达 88%, 可在今后的研究中利用。【结论】本文对星天牛 SSR 位点的信息分析以及引物的设计与验证将有助于星天牛基因挖掘、种群遗传结构、遗传多样性、进化关系和综合治理的研究。

关键词 星天牛; 转录组; 微卫星; 重复类型; 重复基元; 引物

Characteristics of the SSR loci in the *Anoplophora chinensis* transcriptome

HAN Xiao-Hong** WANG Yi-Fan LU Ci-Ding LIN Hao-Yu SHI Yu-Fei
HE Huan ZHANG Fei-Ping LIANG Guang-Hong***

(College of Forestry, Fujian Agriculture and Forestry University, Key Laboratory of Integrated Pest Management
in Ecological Forests, Fujian Agriculture and Forestry University, Fuzhou 350002, China)

Abstract 【Objectives】SSR (Simple Sequence Repeat) markers are a very useful method for species identification, molecular linkage mapping, and genetic diversity studies. In order to develop molecular markers for *Anoplophora chinensis*, SSR markers were identified based on transcriptome sequence. This study provides a basis for studies of genetic diversity and the integrated control of *A. chinensis*. 【Methods】The software MISA was used to explore all microsatellite loci in the transcriptome database of *A. chinensis*. Primers were designed using the software Primer Premier 3.0, and screened using PCR amplification and electrophoresis. 【Results】A total of 2 360 SSRs were identified through software analysis; a frequency of 25.31%. These had 1758 sequences, accounting for 18.85% of the total number of sequences. Most repeat types were mononucleotide motifs (79.03%) followed by trinucleotide motifs (12.54%). The dominant repeat type was A/T (99.30%). The proportion of SSR length to 10-11 bp was highest (56.10%), the maximum number of repeats was 10 and the number of SSR loci was 1188 (50.34%). These results indicate the feasibility of using polymorphism analysis to develop SSR markers for *A. chinensis*. A total of 60 primer pairs were randomly selected and verified by PCR amplification, 53 of which were able to amplify the expected products; a primer availability rate as high as 88% which could be used in future studies. 【Conclusion】

*资助项目 Supported projects: 国家重点研发项目 (2018YFC1200400); 国家重点研发项目 (2017YFD0600105); 国家自然科学基金 (31870641); 福建省高校产学研合作项目 (2018N5101); 福建农林大学优秀研究生学位论文资助基金资助项目 (闽农林大研[2019]2号)

**第一作者 First author, E-mail: hxhdax@163.com

***通讯作者 Corresponding author, E-mail: fjlhg@126.com

收稿日期 Received: 2019-04-09; 接受日期 Accepted: 2019-08-07

The information analysis of SSR loci and the design and validation of primers will be helpful for gene mining, and for research on population genetic structure, genetic diversity, evolutionary relationships and the integrated control of *A. chinensis*.

Key words *Anoplophora chinensis*; transcriptome; SSR; repeat type; motif type; primer

星天牛 *Anoplophora chinensis*, 属鞘翅目 Coleoptera 天牛科 Cerambycidae, 欧洲和地中海植物保护组织 (EPPO) 将其列为 A2 类检疫性有害生物名单。星天牛在国内多地均有分布, 是危害我国杨树、柳树、榆树、法国梧桐、紫薇、悬铃木、苦楝、木麻黄、枣树、板栗、柑桔等多种树木的一种重要蛀干害虫, 由于长期适应不同生境和寄主, 继而形成不同的地理种群, 严重制约当地林业以及对外贸易的发展 (魏建荣等, 2011)。星天牛的防治主要采用化学杀虫剂, 取得了较好的防治效果 (黄金水等, 2001), 但长期使用化学杀虫剂可能导致其抗药性增强, 并有利于其种群进一步分化, 从而改变其种群的遗传结构 (Ran *et al.*, 2009)。但到目前为止, 由于缺少有效的分子标记, 导致星天牛的抗药性与其种群遗传结构的相关研究尚未进行, 也限制了基因流研究技术在其抗药性监测、综合治理等方面发挥关键作用。

SSR (Simple sequence repeat, SSR) 分子标记, 又称微卫星或简单序列重复, 通常包含 1-6 个碱基, 为重复基元组成的一段 DNA 序列 (Kalia *et al.*, 2011; 刘昭阳等, 2016), 具有很好的多态性、共显性以及数量丰富等特点, 已在生物遗传研究中得到广泛应用 (Varshney *et al.*, 2005; Deepti *et al.*, 2017)。然而, 关于星天牛 SSR 研究的目前尚无相关报道, 研究并获得其 SSR 位点信息和位点特征, 将是进一步研究星天牛种群遗传结构研究的重要基础和技术手段。传统的 SSR 分子标记技术试验过程复杂、工作量大且效率较低等缺陷 (常玉梅等, 2005; Zane *et al.*, 2010), 特别在研究基因序列未知的新物种或生物信息极为有限的非模式物种时, 存在较大难度 (Andrés and Bogdanowicz, 2011; 唐培安等, 2017)。近年来高通量测序不断发展, 使用转录组技术开发 SSR 分子标记则提供了一种较为经济、方便且高效的途径, 除上述优点外, 还可大

批量的开发其引物, 故为许多新物种及非模式昆虫的系统发生、遗传进化及其遗传多样性等研究领域提供了更多的选择 (Bai *et al.*, 2011; 杨帆等, 2014)。已有许多研究在获得转录组数据的基础上, 进行 SSR 位点的分析并设计引物, 比如对桔小实蝇 *Bactrocera dorsalis* (魏丹丹等, 2014)、印度谷螟 *Plodia interpunctella* (唐培安等, 2017)、小菜蛾 *Plutella xylostella* (柯富士, 2015)、甜菜夜蛾 *Spodoptera exigua* (Pascual *et al.*, 2012) 等昆虫在基于转录组数据的研究基础上, 成功筛选其 SSR 位点, 获得该物种的 SSR 位点特征, 并且通过此方法成功设计、验证并筛选出可用引物, 大大提高了试验效率、节约了研究成本, 降低了研究难度, 更多相关研究也都表明基于高通量测序数据开发 SSR 标记具有可靠的优势 (Untergasser *et al.*, 2012; Li *et al.*, 2013; Duan *et al.*, 2017; 李微等, 2017; 熊翠玲等, 2017)。

对于星天牛微卫星位点的研究还处于空白状态, 无法满足对其进行更深入研究的需要, 本研究利用生物信息学方法, 基于前期转录组测序结果对星天牛进行 SSR 位点的挖掘, 对其 SSR 位点及特征进行统计与分析, 并对其进行引物设计及初步验证, 为星天牛的多态性和遗传多样性研究提供依据。

1 材料与方法

1.1 试虫的 RNA 提取及转录组测序

供试的 12 头星天牛成虫采集于福建省福州市, 经饥饿处理后, 采用 TRIzol 法提取星天牛总 RNA, 将提取的 RNA 进行电泳检测; 使用 Nanodrop 分光光度计 (IMPLEN, CA, USA) 检测 OD₂₆₀/OD₂₈₀ 的值, RNA 的完整性利用 Agilent 2100 (Agilent Technologies, CA, USA) 进行检测。将检测合格的样品送往北京百迈客公司进行转录组 Illumina 高通量测序。

1.2 转录组 SSR 位点的搜索

以组装获得的大于 1 000 bp 的 unigene 作为参考序列, 使用 SSR 分析软件 MISA (MicroSAteellite) 在 unigene 中筛选 SSR 位点, 对筛选的 SSR 统计分析。筛选标准为: 单核苷酸最少重复次数为 10 次, 二核苷酸、三核苷酸、四核苷酸、五核苷酸、六核苷酸依次为 6、5、5、5、5 次, 复合型 SSR 间隔不超过 100 bp。

1.3 SSR 引物设计、DNA 提取与 PCR 扩增

基于星天牛 SSR 数据, 使用 Primer 3 软件 (Schoebel *et al.*, 2013; 时小东等, 2016; 孟翔等, 2017) 设计引物 (福州擎科公司合成), 使用试剂盒 (OMEGA Insect DNA Kit) 提取单头星天牛 DNA, 提取后进行 PCR 扩增。提取的反应总体积为 20 μ L: 包含 DNA 模板 1 μ L, 2 $\times$ EasyTaq PCR SuperMixTip (全式金生物) 10 μ L, Forward primer 0.4 μ L, Reverse primer 0.4 μ L, ddH₂O 8.2 μ L 补足至 20 μ L。然后采用 T100 PCR 扩增仪 (美国伯乐公司) 进行扩增, 扩增条件为: 94 $^{\circ}$ C 预变性 5 min; 94 $^{\circ}$ C 变性 30 s, 60 $^{\circ}$ C 退火 30 s, 延伸 72 $^{\circ}$ C 30 s, 35 个循环, 72 $^{\circ}$ C 延伸 10 min, 反应结束后, 取样品 5 μ L 进行凝胶电泳 (1%琼脂糖) 检测。

2 结果与分析

2.1 RNA 提取质量及测序结果

经各项检测, 其 RNA 提取质量良好, 没有发生降解与污染情况, 纯度较高, 且样品总体完整性较高, 满足后续测序质量要求。全部测序数据的 reads 的质量参数 Q20 为 94.5%, 表明测序质量好, 结果可靠, 达到后续生物信息学分析的质量要求。从组装后得到 32 247 条 unigenes 中取大于 1 000 bp 的 unigenes 共 9 325 条 (占全部 Unigenes 的 28.92%), 作为 SSR 位点挖掘的基础数据。

2.2 SSR 的数量

利用 MISA 软件对星天牛转录组进行 SSR 位点搜索, 在 9 325 条 unigenes 中搜索筛选获得

2 360 个 SSR 位点, 其出现频率为 25.31%, 平均每 9.16 kb 有一个 SSR 位点。2 360 个 SSR 位点散布在 1 758 条 unigene 上, 发生频率是 18.85%, 其中, 包含 1 个以上 SSR 的序列数目为 427 条, 88 条 unigene 具有复合型 SSR。

2.3 SSR 基元分布特征

在星天牛转录组 SSR 的所有核苷酸重复类型中, 单核苷酸 (p1) 重复基元的 SSR 占比最高 (79.03%)。其次是三核苷酸 (p3) 重复, 占 12.54% (表 1)。在已获得的 82 种重复基元中, 三核苷酸类型最多 (52 种), 其次是四核苷酸 (p4) 15 种、二核苷酸 (p2) 11 种、单核苷酸 4 种。其中, 在三核苷酸重复的 52 个基元类别中, SSR 位点数范围在 1-18 之间, 最多的是 TTA, 其次是 TAA (17)、GAA (17)、ATA (16)、AAG (16)、ATT (15); 在其基元类型中 AAT/ATT 的 SSR 位点数目是 91 个, 占总 SSR 位点的 3.86%, 高于其他三核苷酸重复基元类型, AAG/CTT、AGC/CTG 依次占比为 2.84%、1.53%, 其他三核苷酸基元类型占比不足 1% (图 1)。在核苷酸重复类型中, A/T 重复基元类别最丰富, 所占比例高达 99.30%; AT/AT、AAT/ATT、AAAT/ATTT 分别在二核苷酸、三核苷酸、四核苷酸中出现频率最多, 它们在其各自重复基元类型中分别占比 70.79%、30.74%、57.14%。表明在星天牛转录组 SSR 位点中富含 A、T 碱基。

对不同类型的 SSR 进行密度分布统计, p1 型密度最大, 为 80.07 个 $\cdot$ Mb⁻¹, p2 型密度大小为 7.40 个 $\cdot$ Mb⁻¹、p3 型密度大小为 13.09 个 $\cdot$ Mb⁻¹、p4 型密度大小为 0.88 个 $\cdot$ Mb⁻¹、c 型 (混合 SSR, 即包含至少两个完美 SSR, 且之间距离小于 100 bp, 在 unigene 上的位置后者>前者) 密度大小为 3.61 个 $\cdot$ Mb⁻¹、c*型 (混合 SSR, 即包含至少两个完美 SSR, 且之间距离小于 100 bp, 在 unigene 上的位置后者 $\leq$ 前者) 密度大小为 0.046 个 $\cdot$ Mb⁻¹ (图 2)。

2.4 SSR 基元长度与重复次数

星天牛 SSR 数据中, 其重复长度在 10-199

表 1 星天牛转录组 SSR 重复基元类型及优势重复基元的组成数量

Table 1 Constitute of SSR repeat motif types and dominant repeat motif in *Anoplophora chinensis* transcriptome

重复类型 Repeat type	数目 Number	发生频率 Frequency (%)	优势重复基元 Dominant repeat motif
单核苷酸 (p1) Mononucleotide	1 865	79.03	A/T (99.30%)
二核苷酸 (p2) Dinucleotide	178	7.54	AT/AT (70.79%)
三核苷酸 (p3) Trinucleotide	296	12.54	AAT/ATT (30.74%)
四核苷酸 (p4) Tetranucleotide	21	0.89	AAAT/ATTT (57.14%)

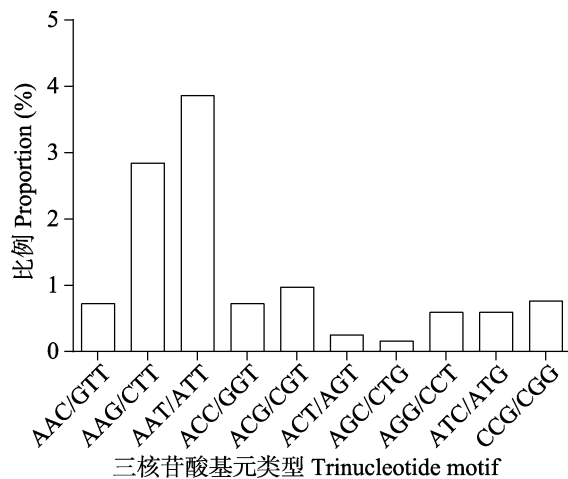


图 1 SSR 中三碱基基元类别比例

Fig. 1 Proportion of trinucleotide motif in SSR

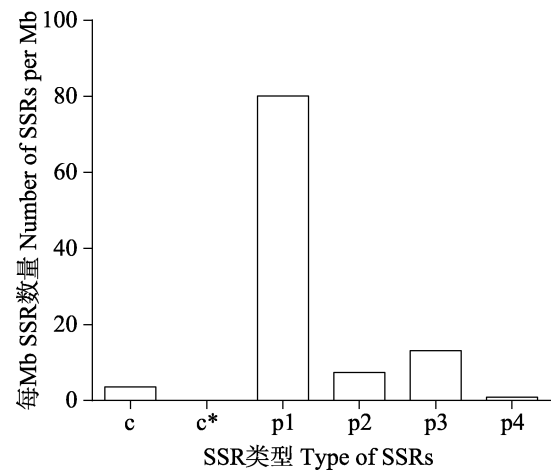


图 2 SSR 密度分布图

Fig. 2 Distribution of SSR density

表 2 星天牛 SSR 基元长度及其比例

Table 2 The length and proportion of SSR in *Anoplophora chinensis*

分布 Distribution	重复基元长度 Motif length (bp)												总计 Total
	10	11	12	13	14	15	16	17	18	19	20	>20	
数量 Number	1 002	321	252	63	87	324	44	0	41	0	25	201	2 360
比例 Proportion (%)	42.46	13.60	10.68	2.67	3.69	13.73	1.86	0	1.73	0	1.06	8.52	100

个碱基不等, 平均长度为 14.13 个碱基。以 10 bp 为筛选条件在星天牛转录组 SSR 数据中进行筛选, SSR 长度主要集中于 10-24 bp (表 2)。长度为 10-11 bp 的 SSR 数量分别为 1 323 个, 其比例为 56.10%; 序列长度 12-24 个碱基有 881 个, 所占比例为 37.33%; 序列长度大于 24 的数量较少, 仅占 0.47%, 且多数为复合型 SSR。

重复次数的统计结果表明, 不同核苷酸的重复基元的重复次数主要分布范围为 5-20 之间 (图 3)。单核苷酸重复次数主要分布范围是 10-16,

为 1 181 个, 占数的 50.04%; 二核苷酸、三核苷酸、四核苷酸重复次数分别主要分布于 5-13、5-10、5, 分别占总 SSR 的 3.94%、8.31%、0.42%。当基元的重复次数逐渐增大, 其所占比例在下降, 重复次数为 5-15 次的最多, 占比达到 98.99%。总体发现为: 低重复次数最多 (70.81%, 5-10 次); 一般重复次数居中 (29.15%, 11-20 次), 包括 11-15 次 (28.18%), 16-20 次 (0.97%); 较高重复次数比例最小 (0.04%/>20 次) (图 4)。在所有碱基重复类型中, 二碱基、三碱基 SSR

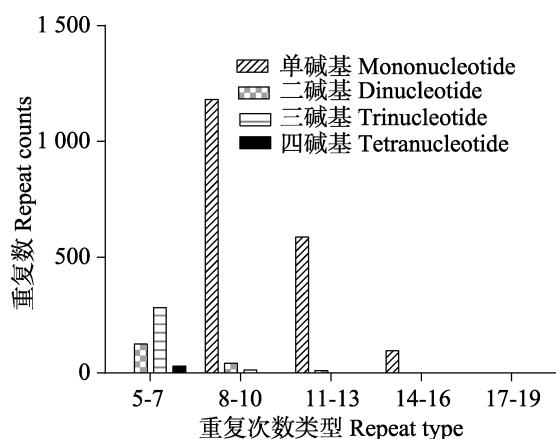


图3 星天牛转录组不同碱基类型的 SSR 重复次数分布
Fig. 3 Distribution of SSR repeats among different motif types in *Anoplophora chinensis* transcriptome

重复数目最为接近。二碱基类型较少, 占总数的 7.54%, 其中, AT/AT 数量最多, 占比最大; 三碱基重复占总数的 12.54%, AAT/ATT 的数量为 91, 为最高比例 (图 5)。重复 6 次的二碱基的 SSR 数量最大, 共有 103 个; 三碱基重复 5 次 SSR 数目最多 (209 个)。在重复次数从 10 开始, 二碱基出现 12 次, 三碱基类型出现次数为 0。在二碱基重复类型中, 其重复次数分布于 6-11 次的较多, 三碱基主要为 5-9 次。

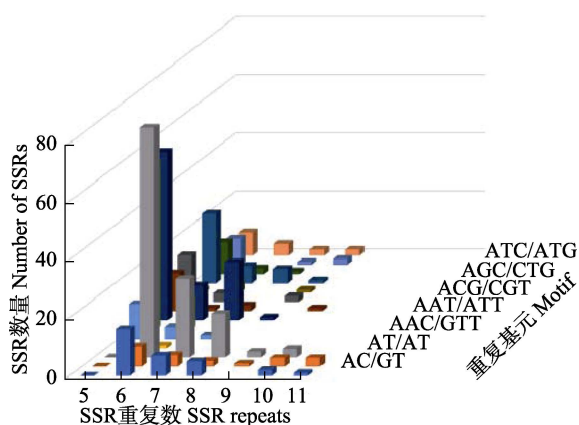


图5 星天牛二核苷酸、三核苷酸重复基元
SSR 分布频率及重复数

Fig. 5 Frequency distribution dinucleotide and trinucleotide SSRs by motif and repeats number of *Anoplophora chinensis*

2.5 SSR 引物与 PCR 扩增分析

对所设计的 60 对引物进行初步验证分析

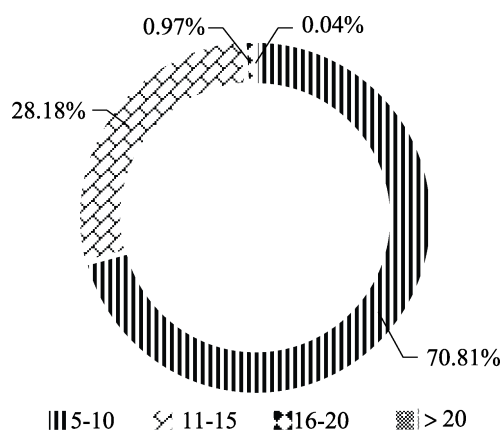


图4 星天牛转录组中 SSR 位点的重复次数分布
Fig. 4 Distribution of the number of repeats of SSR loci in the transcriptome of *Anoplophora chinensis*

(表 3), 结果表明 60 对引物中有 57 对引物可以扩增得到条带, 3 对引物扩增不出条带, 57 对中有 1 对引物有非特异性条带的产生, 3 对引物扩增产物与预期相差较大, 其余 53 对引物与预期相符, 其产物大小在 100-300 bp 之间, 但这些位点的多态性高低有待于进一步的实验进行检验 (图 6)。

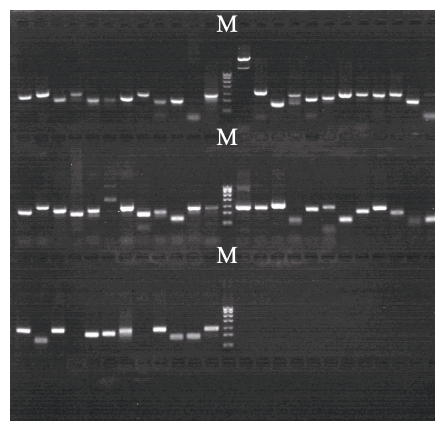


图6 1-60 号引物的扩增
Fig. 6 Amplification of primer 1-60

Marker 从上至下依次为: 700、600、500、400、300、200、100 bp。从左至右依次为: 引物 1-12 及 Marker; Marker 及引物 13-24; 引物 25-36 及 Marker; Marke 及引物 37-48; 引物 49-60 及 Marker。

Marker from top to bottom is: 700, 600, 500, 400, 300, 200, 100 bp. Sequentially from left to right: primers 1-12 and marker; marker and primers 13-24; primers 25-36 and marker; marker and primers 37-48; primers 49-60 and marker.

表 3 53 对 SSR 引物信息
Table 3 Information of 53 pairs of SSR primers

引物 编号 Primer No.	转录组 序列 Unigene	上游引物 Forward primer (5'-3')	下游引物 Reverse primer (5'-3')	产物大小 Size (bp)	退火温度 Melting temperature (°C)
1	c10014	TCATCTCCCGAATAAGGCAC	AGAAAAACACAAACCGGACG	236	60
2	c10042	GCAGTCTTTCGATGATTTGGA	ACGCATAGACATTCCGCTTC	271	60
3	c10070	ACTTTCCCTCGTGAAAACA	TTAATGTGGACAGCACCACG	183	60
4	c10096	TGCAGTAAACATGATAAGGCAA	TCTGGATCCGCTATTCCAAG	248	59
5	c10111	AACAATTAATCTGATCATTAACCATGT	GCTCTTGTTGCCAAGTGCAT	182	60
6	c10148	TGCCTTTTCAAAGAAAAATGAAA	CGCTCCTTAAATGACAAAACC	173	59
7	c10167	CAGTTACAGCGATGTGGTGG	CGTAATTGTGAGAGCCTAGCG	207	60
8	c10193	TGGCGAAATCGTACCAAAGT	TTCTTGTCATGCTGTTGCGT	256	60
9	c10263	GGGTGATTCTTAAGTTGGGTC	TCAGACTGGCGACTTTTGAA	176	59
10	c10286	AGACGGTGATGGACCAGTTC	GCAGAGGCCCATATTCTTCA	194	60
12	c10549	AATGATAAAACCGCGGAGTG	GTTGTGGGCCAAAAACCA	276	60
14	c10684	TGCGTTGCTGTACTCTCACC	CTCATAAACAGCAACAGCCG	259	60
15	c10758	CAAGTCCATGTTCCCTCCGAT	CAGAGCTTTAACGATAGCACCA	165	59
16	c10961	TTGAGTTGTAGAAATCACAGCGA	TGCATTTATACCACAGTCCCTTT	266	59
17	c11045	AAATGATAGCGCGGTATTG	ATCTTAAGCCACTTCCCCGT	209	59
18	c11159	AAACATGCACGTGATCTTGAA	CCTCGTGGTTTATTTGCTCA	227	59
19	c11237	GACAGCTTTTGCAAGGAAA	ATCTTGAAACGCAACGATCC	267	60
20	c11375	AATGGGGAAATTGCAAGAAA	GCAGCAAACAAATTGCTCAT	260	59
21	c14260	GAGGGAGGGTGTGTATGGA	ACCTTGCGGAAATACGTGTC	260	60
22	c14526	AAAGCCCCTATTCCAAAAACA	AAAGGAGGCGTTAACAAATTGA	267	60
23	c14546	AAGCATCGCAGGGTTCTTTA	CGTCACCAACTACTTCCGGT	180	60
25	c14797	GCCCAGTAGCTCAAACCTCG	ACCTCTCAGGTTTCTCCGT	201	60
26	c14911	TGTGAAAAACATCGATTCTCAAG	CATTTCCCCAAGTTTGCATT	272	59
27	c15230	ATCTTGTCGTTACCCGTTT	GTATGTGTGCTGCAGTGCG	182	59
28	c15268	GGGTCCCCTAAAGTTGTCGT	ATGGCAATGGGAAACCAATA	189	60
29	c15348	CGGAGGTTTCCCATCTGTTA	TAACATTTGCCGACTGCTG	234	59
31	c16368	CACCTTGCTGTAATCCATCG	CCAGCAGATGGCTTTTGTCT	280	60
32	c16412	CACCGAGAGCCTATTACATTCC	GGCGAAAGGTGTTTGGATA	186	59
33	c16415	CTGGGTGGAACAAGGGAGTA	TGTTATTTCAGAAATGTGAAAC CTC	211	59
34	c16494	TGACTTGGCCCTTTTAGGAA	AAGAAACAATATGGTTATTTTAA TGC	144	58
35	c16602	CCATTCCAAGTACCCATGCT	ATCCGGTAATCTGTCCGATG	269	59
36	c16730	TTTGTTCCAGTACAAAATTCACA	CACCTTGGGGTGTGCTAGAT	263	58
37	c17044	TAAAACATTGAGCGCGACAG	GCCTTAGTAGCCGGTTTCCT	264	60

续表 3 (Table 3 continued)

引物 编号 Primer No.	转录组 序列 Unigene	上游引物 Forward primer (5'-3')	下游引物 Reverse primer (5'-3')	产物大小 Size (bp)	退火温度 Melting temperature (°C)
38	c17233	TTTTCTTATTTGACCGCCC	GGTATCGAACCAATCAAATCAA	260	59
39	c17477	ATCACTTCCTCCACACTGGG	GATCCACTTGAACAGGCCAT	249	59
40	c17485	CTTCCTGCGGTTTGATTGTT	TGCATAATGTCAAAATAGATTGTT TAT	122	57
41	c17528	CGCCCCCTCAAACAAACATA	TGTAGCTTGGGTAAAACTGCG	246	60
42	c17835	AGGTCTGCCGTTCTTTTTGA	CCCTTTTGTATAAACGGCAA	272	59
43	c17864	GCTTCCAACGAACGTTTCTC	GCAAGGAGCACACATTTCAA	132	59
44	c17889	TATCGTGCTGGTACTGCTGC	AACAGAACGCTATGCCGAAC	213	60
45	c17918	GTACGCTTCTCTGCCATTCC	GTCGATTACGAGGGTGGAGA	261	60
46	c17934	AATTTGGAGTGGGGTAACCA	TGGGGTACAGTGTATTTTAGCG	198	59
48	c19931	GTGCTGCGACTACTGACGAG	GAGTCCAGCGACGACTTTTT	104	59
49	c19949	GCAGCATAAGCAGGGAGAAG	GGTACGCACCGAAAGGATTA	252	60
50	c19956	AACGTCAGTGTGACGAAAA	CTGACAGCACGTAAACATTGG	150	59
51	c19992	CCATCTCTTGCGGTGTTCTT	AAGCCTCGATTGTTTCGCTA	265	60
53	c20158	CCGAGGCGTGCCTATTATT	GTCGTCGTTAGTCCGCAGTT	217	60
54	c20159	TCAAGCCGAGACACAACAAC	CCACGGAGTCGTACACAAGA	229	59
55	c20180	GTCGGTTAACCAGTTTCGGA	GACAGACCCGGAGAAATTGA	260	59
57	c20213	GGATACCAGCAAACCTGACC	ATTTGTACCGCAAAGGCAAA	279	59
58	c20384	TTGTTTTCAAGTGTTCCTCC	AAAGCAAGCACGTAACGGAT	174	59
59	c20400	TCGTTGGAGGAATAAGTGGG	CCCTCACCTTGGACATGAGT	178	59
60	c20417	TGCTTGCAATCAAAGTCTGG	GGGAGGGATTACCGTTTGTT	279	60

3 讨论

研究发现, SSR 位点出现频率常在不同物种间产生的差异, 可能与数据库大小、搜索筛选 SSR 位点序列标准、RNA 质量、物种特异性等有关, 可作为评价昆虫物种 SSR 的多态性重要参考。本研究从星天牛转录组数据中筛选获得 2 360 个 SSR 位点, 出现频率为 25.31%。其出现频率高于已报道过的多种昆虫, 如荔枝蒂蛀虫 *Conopomorpha sinensis* 15.25% (孟翔等, 2017)、黑翅土白蚁 *Odontotermes formosanus* 9.98% (Huang *et al.*, 2012)、印度谷螟 *P. interpunctella* 8.52% (唐培安等, 2017)、扶桑绵粉蚧 *Phenacoccus solenopsis* 6.33% (罗梅等, 2014)、

烟粉虱 *Bemisia tabaci* 5.07% (Xie *et al.*, 2012)、橘小实蝇 *B. dorsalis* 4.23% (魏丹丹等, 2014)、黄粉虫 *Tenebrio molitor* 1.67% (朱家颖等, 2013)、粘虫 *Mythimna separata* 1.93% (胡艳华等, 2015)、云南切梢小蠹 *Tomicus yunnanensis* 1.29% (袁远等, 2014) 等, 仅低于大垫尖翅 *Epacromius coerulipes* 44.17% (金永玲等, 2015), 表明星天牛转录组中 SSR 的多态性潜能较高。

同时, 较长的重复基序突变频率较低, 物种进化水平相对较低; 较短的重复基序存在数量越多, 则该物种的进化水平相对较高 (孟翔等, 2017)。本研究筛选得到的星天牛 SSR 主要为低级基元 (单碱基、二碱基、三碱基重复), 单核苷酸重复最多 (79.03%), 然后三核苷酸、二核

苷酸重复位于其后,这与大多数已报道的研究结果一致(Wang *et al.*, 2012; 时小东等, 2016)。在单碱基重复基元中,优势重复基元为 A/T,与多种昆虫相似,如黏虫 *M. separata* (胡艳华等, 2015)、扶桑绵粉蚧 *P. solenopsis* (罗梅等, 2014)、印度谷螟 *P. interpunctella* (唐培安等, 2017)、褐飞虱 *Nilaparvata lugens* (刘玉娣和侯茂林, 2010) 等。在 SSR 重复类型中,三核苷酸重复类型最多,占比 63.42%,其次是四核苷酸重复(18.29%)。这可能与三联体密码子有关,mRNA 分子上的 3 个碱基决定 1 个氨基酸,绝大多数 3 碱基重复型多位于编码区,直接与基因表达相关(Gao *et al.*, 2003; 柯富士, 2015; 唐培安等, 2017; 郭欢等, 2018)。在星天牛 SSR 重复类型中没有发现五碱基、六碱基重复,可能与 SSR 位点筛选的标准有关,也可能是物种本身所致,也在某种程度上表明星天牛进化水平相对较高。

除 SSR 位点出现频率与重复基序会对其多态性有影响外,SSR 长度以及重复次数等因素也会对 SSR 分子标记多态性造成重要影响(罗梅等, 2014; 时小东等, 2016)。当 SSR 长度低于 12 bp 时,其多态性较低;在 12-20 bp 左右时,多态性中等;达到 20 bp 后,多态性较为丰富(Temnykh *et al.*, 2001)。而在星天牛转录组 SSR 数据中,其长度所集中的范围表现出其具有中上水平的多态性。另一方面,重复次数与微卫星分子标记的多态性具有正相关性(Gao *et al.*, 2003; 时小东等, 2016),星天牛转录组 SSR 位点不仅在二碱基、三碱基类型中重复次数较多、跨度较大,而且其长度也各不相同,在理论上表明其二核苷酸 SSR 较其它类型核苷酸具更高的多态性,可对其进行精准开发和引物设计,具有重要应用前景。

在 60 对初步设计的潜在引物中,53 对引物符合预期产物大小,与前人研究相比(邓欣等, 2008),本研究可批量设计引物,节省了研究时间与成本,虽然目前无法确定其多态性的高低,但其可用于后续研究的候补引物的可用性较高(88%),少数引物与预期扩增效果不符合,具有非特异性,可能与扩增片段中插入内含子有

关,也有可能与 SSR 位于同源基因序列有关(孟翔等, 2017)。未来还可对与预期相符的引物进行深入挖掘和验证,补充更丰富的虫源 SSR 数据,以进一步验证并提高其多态性和通用性,以便更好的开发 SSR 分子标记。

本研究利用星天牛转录组数据,筛选获得其 SSR 位点信息,对星天牛 SSR 位点的分布、类型及其特征进行了统计与分析,并设计其引物,候选引物可用性较高,为其进一步挖掘潜在 SSR 标记及其遗传多样性研究奠定了基础。

参考文献 (References)

- Andrés JA, Bogdanowicz SM, 2011. Isolating microsatellite loci: Looking back, looking ahead. *Methods Mol. Biol.*, doi: 10.1007/978-1-61779-228-1_12.
- Bai X, Mamidala P, Rajarapu SP, Yuan ML, 2011. Transcriptomics of the bed bug (*Cimex lectularius*). *PLoS ONE*, 6(1): e16336.
- Chang YM, Li SW, Liang LQ, Sun XW. 2005. Strategies for microsatellites isolation. *China Biotechnology*, (s1): 210-214. [常玉梅, 李绍戊, 梁利群, 孙效文, 2005. 微卫星标记的制备策略. 中国生物工程杂志, (s1): 210-214.]
- Deepti N, Swati S, Ramakrishna G, Archana, S, Singh NK, Kishor G, 2017. De novo assembly and characterization of *Cajanus scarabaeoides* (L.) thouras transcriptome by paired-end sequencing. *Frontiers in Molecular Biosciences*, doi: 10.3389/fmolb.2017.00048.
- Deng X, Chen XB, Long SH, Wang XC, Gao Y, He DF, Wang J, Wang YF, 2008. Microsatellite molecular marker enrichment by magnetic beads in flax. *Acta Agronomica Sinica*, 34(12): 2099-2105. [邓欣, 陈信波, 龙松华, 王孝纯, 高原, 何东锋, 王进, 王玉富, 2008. 用磁珠富集法分离亚麻基因组微卫星分子标记. 作物学报, 34(12): 2099-2105.]
- Duan X, Wang K, Su S, Tian R, Li Y, Chen M, 2017. De novo transcriptome analysis and microsatellite marker development for population genetic study of a serious insect pest, *Rhopalosiphum padi* (L.) (Hemiptera: Aphididae). *PLoS ONE*, 12(2): e0172513.
- Guo H, Wang G, Zhang ST, Huang M, 2018. Development of SSR primers for *Simulium* (*Eusimulium*) *angustipes* (Diptera: Simuliidae) based on RNA-seq dataset. *Acta Entomologica Sinica*, 61(7): 815-824. [郭欢, 王刚, 张树田, 黄敏, 2018. 基于 RNA-seq 数据的窄足真蚋 SSR 分子标记开发. 昆虫学报,

- 61(7): 815–824.]
- Gao L, Tang J, Li H, Jia J, 2003. Analysis of microsatellites in major crops assessed by computational and experimental approaches. *Molecular Breeding*, 12(3): 245–261.
- Huang JS, He XY, Ye JX, Huang YQ, Gao ML, 2001. Studies on the control of *Anoplophora chinensis* (F.) by alluring adult with *Melia azedarach* L. *Scientia Silvae Sinicae*, 37 (4): 58–64. [黄金水, 何学友, 叶剑雄, 黄衍庆, 高美玲, 2001. 苦楝引诱防治星天牛研究. 林业科学, 37(4): 58–64.]
- Huang QY, Sun PD, Zhou XG, Lei CL, 2012. Characterization of head transcriptome and analysis of gene expression involved in caste differentiation and aggression in *Odontotermes formosanus* (Shiraki). *PLoS ONE*, 7(11): e50383.
- Hu YH, Li M, Zhang HF, Li SC, Wang Q, Zhao HL, 2015. The information analysis of SSR loci in the *Mythimna separate* (Walker) transcriptome. *Journal of Shanxi Agricultural University (Natural Science Edition)*, 35(5): 484–489. [胡艳华, 李敏, 张虎芳, 李生才, 王青, 赵惠玲, 2015. 粘虫转录组中 SSR 位点的信息分析. 山西农业大学学报(自然科学版), 35(5): 484–489.]
- Jin YL, Cong B, Wang LY, Zhang HY, Dong H, 2015. An analysis of the transcriptome of *Epacromius coerulipes* (Orthoptera: Acrididae). *Acta Entomologica Sinica*, 58(8): 817–825. [金永玲, 丛斌, 王丽艳, 张海燕, 董辉, 2015. 大垫尖翅蝗转录组分析. 昆虫学报, 58(8): 817–825.]
- Ke FS, 2015. Characterization and isolation of microsatellites from the *Plutella xylostella* transcriptome. Master dissertation. Fuzhou: Fujian Agriculture and Forestry University. [柯富士, 2015. 小菜蛾转录组微卫星序列的分析和多态性位点的开发. 硕士学位论文. 福州: 福建农林大学.]
- Kalia RK, Rai MK, Kalia S, Singh R, Dhawan AK, 2011. Microsatellite markers: An overview of the recent progress in plants. *Euphytica*, 177(3): 309–334.
- Li LT, Zhu YB, Ma JF, Li ZY, Dong ZP, 2013. An analysis of the *Athetis lepigone* transcriptome from four developmental stages. *PLoS ONE*, 8(9): e73911.
- Luo M, Zhang H, Bin SY, Lin JT, 2014. High-throughput discovery of SSR genetic markers in the mealybug, *Phenacoccus solenopsis* (Hemiptera: Pseudococcidae), from its transcriptome database. *Acta Entomologica Sinica*, 57(4): 395–400. [罗梅, 张鹤, 宾淑英, 林进添, 2014. 基于转录组数据高通量发掘扶桑绵粉蚧微卫星引物. 昆虫学报, 57(4): 395–400.]
- Li W, Zhang L, Cheng YX, Luo LZ, Jiang XF, 2017. High-throughput discovery of microsatellite markers based on transcriptome sequencing in the oriental armyworm, *Mythimna separata* (Walker). *Journal of Plant Protection*, 44(3): 377–384. [李微, 张蕾, 程云霞, 罗礼智, 江幸福, 2017. 应用转录组测序高通量发掘东方粘虫 SSR 标记. 植物保护学报, 44(3): 377–384.]
- Liu YD, Hou ML, 2010. Analysis of microsatellite information in EST resource of *Nilaparvata lugens* (Homoptera: Delphacidae). *Acta Entomologica Sinica*, 53(3): 239–247. [刘玉娣, 侯茂林, 2010. 褐飞虱 EST 资源的微卫星信息分析. 昆虫学报, 53(3): 239–247.]
- Liu ZY, Li YR, Pan ZY, Guo B, Luo YQ, Tao J, 2016. Phenotypic diversity and genetic variation of Asian long-horned beetle, *Anoplophora glabripennis* (Motschulsky). *Chinese Journal of Applied Entomology*, 53(5): 1045–1057. [刘昭阳, 李玉蓉, 潘忠玉, 郭冰, 骆有庆, 宗世祥, 陶静, 2016. 光肩星天牛两型种群表型多样性分析. 应用昆虫学报, 53(5): 1045–1057.]
- Meng X, Hu JJ, Li YH, Yang GC, 2017. Analysis of SSR loci in transcriptome database of *Conopomorpha sinensis* Bradley (Lepidoptera: Gracilariidae). *Journal of Environmental Entomology*, 39(6): 1219–1224. [孟翔, 胡俊杰, 李艳华, 欧阳革成, 2017. 基于转录组数据的荔枝蒂蛀虫 SSR 位点信息分析. 环境昆虫学报, 39(6): 1219–1224.]
- Pascual L, Jakubowska AK, Blanca JM, Calzares J, Juan F, Gloeckner G, 2012. The transcriptome of *Spodoptera exigua* larvae exposed to different types of microbes. *Insect Biochemistry & Molecular Biology*, 42(8): 557–570.
- Ran C, Chen Y, Wang JJ, 2009. Susceptibility and carboxylesterase activity of five field populations of *Panonychus citri* (McGregor) (Acari: Tetranychidae) to four acaricides. *International Journal of Acarology*, 35(2): 115–121.
- Schoebel CN, Brodbeck S, Buehler D, Cornejo C, Gajurel J, Hartikainen H, 2013. Lessons learned from microsatellite development for nonmodel organisms using 454 pyrosequencing. *Journal of Evolutionary Biology*, 26(3): 600–611.
- Shi XD, Zhu XH, Sheng YZ, Zhuang GQ, Chen F, 2016. Development of SSR markers based on transcriptome sequence of *Phoebe zhennan*. *Scientia Silvae Sinicae*, 52(11): 71–78. [时小东, 朱学慧, 盛玉珍, 庄国庆, 陈放, 2016. 基于转录组序列的楠木 SSR 分子标记开发. 林业科学, 52(11): 71–78.]

- Tang PA, Tao YX, Xue H, Yuan ML, 2017. Analysis of microsatellite loci in *Plodia interpunctella* based on transcriptome dataset. *Plant Protection*, 43(3): 43–48. [唐培安, 陶冶心, 薛昊, 袁明龙, 2017. 基于转录组数据的印度谷螟微卫星位点分析. *植物保护*, 43(3): 43–48.]
- Temnykh S, Declerck G, Lukashova A, 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposon associations, and genetic marker potential. *Genome Res.*, 11(8): 1441–1452.
- Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, 2012. Primer3--new capabilities and interfaces. *Nucleic Acids Research*, 40(15): e115.
- Varshney RK, Graner A, Sorrells ME, 2005. Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology*, 23(1): 48–55.
- Wang B, Ekblom R, Castoe TA, Jones EP, Jacob H, 2012. Transcriptome sequencing of black grouse (*Lyrurus tetrix*) for immune gene discovery and microsatellite development. *Open Biology*, 2(4): 120054.
- Wei DD, Shi JX, Zhang XX, Chen SC, Wei D, Wang JJ, 2014. Analysis of microsatellite loci from *Bactrocera dorsalis* based on transcriptome dataset. *Chinese Journal of Applied Ecology*, 25(6): 1799–1805. [魏丹丹, 石俊霞, 张夏瑄, 陈世春, 魏冬, 王进军, 2014. 基于转录组数据的桔小实蝇微卫星位点信息分析. *应用生态学报*, 25(6): 1799–1805.]
- Wei JR, Zhao WX, Zhang YA, 2011. Research progress on *Anoplophora chinensis*. *Plant Quarantine*, 25(5): 81–85. [魏建荣, 赵文霞, 张永安, 2011. 星天牛研究进展. *植物检疫*, 25(5): 81–85.]
- Xiong CL, Zhang L, Fu ZM, Wang HQ, Hou ZX, Tong XY, 2017. Large-scale development of SSR primers for *Apis cerana* larvae based on its RNA-seq datasets. *Journal of Environmental Entomology*, 39(1): 68–74. [熊翠玲, 张璐, 付中民, 王鸿权, 侯志贤, 童新宇, 2017. 基于 RNA-seq 数据大规模开发中华蜜蜂幼虫的 SSR 分子标记. *环境昆虫学报*, 39(1): 68–74.]
- Xie W, Meng QS, Wu QJ, Wang SL, Yang X, Yang NN, Li RM, Jiao XG, Pan HP, Liu BM, Su Q, Xu BY, Hu SN, Zhou XG, Zhang YJ, 2012. Pyrosequencing the *Bemisia tabaci* transcriptome reveals a highly diverse bacterial community and a robust system for insecticide resistance. *PLoS ONE*, 7(4): e35181.
- Yang F, Huang LH, Zhang AB, 2014. High-throughput transcriptome sequencing technology and its applications in Lepidoptera. *Acta Entomologica Sinica*, 57(8): 991–1000. [杨帆, 黄立华, 张爱兵, 2014. 高通量转录组测序技术及其在鳞翅目昆虫上的应用. *昆虫学报*, 57(8): 991–1000.]
- Yuan Y, Zhang LF, Wu GX, Zhu JY, 2014. High-throughput discovery microsatellites in *Tomicus yunnanensis* (Coleoptera: Scolytinae). *Journal of Environmental Entomology*, 36(2): 166–170. [袁远, 张丽芳, 吴国星, 朱家颖, 2014. 云南切梢小蠹微卫星的高通量发掘. *环境昆虫学报*, 36(2): 166–170.]
- Zane L, Patarnello T, Ludwig A, Fontana F, Congiu L, 2010. Isolation and characterization of microsatellites in the Adriatic sturgeon (*Acipenser naccarii*). *Molecular Ecology Resources*, 2(4): 586–588.
- Zhu JY, Wu GX, Yang B, 2013. High-throughput discovery of SSR genetic markers in the yellow mealworm beetle, *Tenebrio molitor* (Coleoptera: Tenebrionidae), from its transcriptome database. *Acta Entomologica Sinica*, 56(7): 724–728. [朱家颖, 吴国星, 杨斌, 2013. 基于转录组数据高通量发掘黄粉甲微卫星引物. *昆虫学报*, 56(7): 724–728.]