

## 龙眼顶芽转录组简单重复序列(SSR)标记信息分析及分子标记开发

郭栋梁, 王静, 韩冬梅, 潘学文, 李建光\*

广东省农业科学院果树研究所, 农业部亚热带果树生物学与遗传资源利用重点实验室, 广州510641

**摘要:** 随着新一代测序技术的发展, 大量的转录组数据和表达序列标签(EST)成为开发简单重复序列(SSR)标记的可利用资源。本研究利用MISA软件筛选龙眼(*Dimocarpus longan*)顶芽转录组数据库序列, 从114 445条龙眼转录组unigene序列中发现11 546个SSR位点, SSR出现频率为10.09%。其中1 975条unigene含有两个或两个以上EST-SSR位点, 占有SSR位点的比例为17.10%, SSR出现的平均距离为7.52 kb。从龙眼转录组SSR核苷酸基序类型来看, 二核苷酸(52.11%)和三核苷酸(46.15%)出现频率最高, 占有核苷酸出现频率的99.26%。在龙眼转录组SSR中二核苷酸重复基元出现频率最高的是AG/CT (4 250个, 占36.81%), 三核苷酸重复基元出现频率最高的是AAG/CTT (1 109个, 占9.61%)。对含SSR位点的9 571条unigene序列进行引物设计, 共设计出了8 347对SSR位点特异引物。随机挑选合成50对EST-SSR引物, 以‘石硖’、‘储良’、‘古山2号’、‘立冬本’等四份龙眼材料的基因组DNA为模板对这批引物进行PCR扩增、筛选, 结果表明, 其中21对引物能产生理想的PCR产物, 有效扩增率为42%; 16对引物扩增条带具有多态性, 占有有效引物的76.2%; 16对多态性引物共扩增获得50个条带, 其中多态性片段21个, 每对引物平均产生1.31个多态性片段。

**关键词:** 龙眼; 转录组; SSR

龙眼(*Dimocarpus longan*)是原产我国的热带、亚热带水果, 距今有2 000多年的栽培历史。我国目前拥有300多个龙眼品种(系), 主要分布在广东、福建、广西、海南、四川、台湾等省和自治区(潘丽梅等2008)。龙眼在长期的自然进化过程中, 受自然环境的优劣淘汰及人为选择机制的影响, 形成了丰富的龙眼资源库(许奇志等2008)。丰富的龙眼资源储备是我们龙眼选育种的基础, 而对龙眼种质资源的综合评价及鉴定则是选育种的首要工作。在龙眼漫长而悠久的栽种历史中, 由于无序的引种及各地方言的影响导致存在同名异物或者同物异名的混乱现象(易干军等2003), 亟需对我国龙眼种质资源进行鉴定区分。但是, 仅仅依据龙眼外观形态学已经不能对龙眼品种进行有效鉴别和区分(许奇志等2008)。

分子标记凭借其易操作、低成本、高分辨率等优点广泛应用在植物种质资源的鉴定与分类(王金星等2011)和遗传多样性(Yonemoto等2006)等相关的研究上。在龙眼种质资源鉴定与分类中主要应用的分子标记包括随机扩增多态性DNA (random amplified polymorphic DNA, RAPD)标记、扩增片段长度多态性(amplified fragment length polymorphism, AFLP)标记、简单重复序列间扩增(inter-simple sequence repeat, ISSR)标记、相关序列扩增多态性(sequence-related amplified polymorphism, SRAP)

等。陈有志(2001)用RAPD分子标记推测了新种质‘龙荔’的父母本, 从而揭示出龙眼和荔枝(*Litchi chinensis*)杂交的潜在可能。许奇志(2008)用RAPD分子标记分析了24份原产于闽、粤、川三省的龙眼资源的遗传多样性。高慧颖等(2007)利用RAPD分子标记研究了龙眼遗传多样性与其地域来源的相关性, 提出泰国和印度尼西亚的部分龙眼基因型可能来源于中国的论断。易干军等(2003)采用AFLP标记研究了46份龙眼资源的遗传进化关系。赵玉辉等(2008)利用SRAP标记鉴定了‘紫喜娘’荔枝和‘石硖’龙眼杂交后代的真伪。曾黎辉(2007)通过ISSR标记将55份龙眼种质材料划分为中国龙眼与泰国龙眼两个大类。陈虎等(2010)利用ISSR技术将37份龙眼材料分为中国、泰国和越南三大类, 分类结果与其品种地理来源相符合。郑姗等(2012)利用5条稳定的多态ISSR引物, 将17份龙眼种质分为4类群, 其中引物py-9能有效区分17份龙眼资源。

简单重复序列(simple sequence repeat, SSR)标记具有多态性高、共显性、易检测、重复性高、数量丰富、对基因组有很好的覆盖性等特点, 因此

收稿 2018-03-02 修定 2018-03-29  
资助 国家自然科学基金青年基金(31301741)和广东省农业科学院院长基金(201309)。

\* 通讯作者(550777196@qq.com)。

广泛应用于植物遗传多样性的研究、新品种鉴定、分子标记辅助育种(marker assisted selection, MAS)、遗传连锁图谱的构建等方面。根据SSR的来源可将其分为基因组SSR和表达序列标签(expressed sequence tag, EST)-SSR, EST-SSR分子标记因其具有较好通用性且开发方法简单、成本低等特点应用更为广泛。

刘峰等(2012)利用EST-SSR分子标记技术构建了12条辣椒(*Capsicum annuum*)连锁群。Chen等(2007)以分离得到的141个EST-SSR标记构建了枸橼(*Poncirus trifoliata*)与甜橙(*Citrus sinensis*)之间的遗传图谱。袁菊红(2011)用SSR标记技术分析了15份石蒜属(*Lycoris* spp.)材料和部分中国水仙(*Narcissus tazetta* var. *chinensis*)的亲缘关系。韩瑞霞等(2012)开发了28对苦荞(*Fagopyrum esculentum*) SSR引物并对起源于我国不同地域的苦荞种质的遗传多样性和亲缘关系进行了分析。李书柯等(2011)利用水稻(*Oryza sativa*)基因组数据,通过SSR分子标记技术分析了62份漳浦野生稻和海南野生稻的遗传多样性。孙清明等(2013)利用自主开发的荔枝EST-SSR标记与SNP标记,鉴定了荔枝新品种‘御金球’的亲本来源为‘糯米糍’和‘焦核怀枝’。目前基于EST数据开发的龙眼SSR标记非常少,限制了龙眼分子标记水平的研究发展。洪仕南等(2015)以‘红核子’龙眼胚性愈伤组织转录组数据为基础进行EST-SSR分子标记研究,完成了EST-SSR-PCR试验体系的优化,并将优化后的EST-SSR标记应用于龙眼的亲缘关系分析。本研究对前期获得的龙眼顶芽转录组测序产生的海量EST序列进行全转录组SSR鉴定及引物设计,揭示龙眼顶芽转录组SSR的分布规律和特性,拟为龙眼遗传图谱构建、系谱分析和MAS育种提供高效的分子标记资源。

## 1 材料与方 法

### 1.1 材料

转录组数据为本课题组于2014年采用Illumina高通量测序获得,测序所选用的材料为四年生‘石硖’龙眼(*Dimocarpus longan* Lour.)每次新梢老熟后的顶芽,随机选择3个不同单株的5个新梢顶芽,分别提取其RNA并等量混和用于文库构建。测序委托北京诺禾致源生物信息科技有限公司采用Illumina HiSeq 2500测序平台完成。随后采用转录组

Trinity组装软件进行序列组装。获得的龙眼顶芽转录组数据已上传到美国国立生物技术信息中心(National Center for Biotechnology Information, NCBI)的Sequence Read Archive (SRA)公用数据库,序列号为SRA403714。

用于SSR引物筛选与多态性分析的龙眼品种顶芽取自广东省农业科学院果树研究所龙眼种质资源圃。

### 1.2 方法

#### 1.2.1 RNA和DNA提取

龙眼顶芽RNA提取参照翟世博等(2016)报道的方法。龙眼顶芽片基因组DNA提取参照洪仕南等(2015)的方法。

#### 1.2.2 ‘石硖’龙眼转录组SSR位点鉴别及SSR引物设计

使用MISA软件(<http://pgrc.ipk-gatersleben.de/misa/misa.html>)对‘石硖’龙眼顶芽转录组中的unigene进行SSR位点搜索,搜索标准为:二核苷酸、三核苷酸、四核苷酸、五核苷酸和六核苷酸最少重复次数分别为6、4、3、3和3次。利用Primer3软件(<http://primer3.sourceforge.net/releases.php>)为含有SSR位点的unigene批量设计引物,引物设计的参数参考龙眼已报道的SSR序列引物:序列长度18~27 bp,扩增产物长度100~350 bp,GC含量40%~60%,退火温度57~63°C,上下游引物的退火温度值相差不大于2°C;尽量避免出现发卡结构、二聚体、错配和引物二聚体等二级结构。

#### 1.2.3 龙眼SSR-PCR扩增及产物检测

PCR反应体系为25  $\mu$ L,其中包括:10 $\times$ PCR Buffer (1.5 mmol $\cdot$ L<sup>-1</sup> Mg<sup>2+</sup>) 2  $\mu$ L, 80 ng DNA模板, 0.4 mmol $\cdot$ L<sup>-1</sup> dNTPs, 0.2  $\mu$ mol $\cdot$ L<sup>-1</sup>引物, 1.5 U *Taq* 酶。扩增反应在TaKaRa PCR Thermal Cycler Dice仪器上进行,扩增程序为:94°C预变性5 min; 94°C变性30 s,退火(根据引物合成单上的 $T_m$ 减2°C) 30 s, 72°C延伸1 min, 35个循环; 72°C延伸10 min。扩增产物采用琼脂糖凝胶分离鉴定。

## 2 实验结果

### 2.1 龙眼转录组中SSR的分布及结构特点

通过对龙眼转录组的114 445条unigene(序列总长约86 800.57 kb)序列进行搜索,发现这些unigene序列中含有11 546个SSR位点,其中1 975条unigene

含有两个或两个以上EST-SSR位点。总体上, SSR发生频率为10.09%, 平均每7.52 kb出现一个SSR。龙眼转录组SSR的类型丰富, 从二核苷酸至六核苷酸重复类型均存在, 其中二核苷酸和三核苷酸重复类型出现频率较多, 分别占总SSR出现频率的52.11%和46.15%; 四核苷酸、五核苷酸和六核苷酸重复类型出现频率较少, 分别占总SSR出现频率的1.48%、0.13%和0.13% (表1)。

龙眼转录组SSR重复单元的重复次数分布在5~24次之间, 其中5~10次重复的SSR位点有11 282个, 占总数的97.71%; 11次重复以上的有264个, 占3.29%。所有龙眼转录组SSR重复单元中, 以6次重复次数的SSR最多, 占30.31%; 5次重复的次之, 占28.09% (表1)。龙眼转录组SSR的长度从10~272 bp不等, 分布在10~15 bp的约有7 146个, 占整个SSR的61.89%; 16~20 bp的有2 828个, 占24.49%; 21~25 bp的有719个, 占6.23%; 长度大于25 bp的有852个, 占7.38%。

## 2.2 龙眼转录组SSR基序重复类型和频率特征

从龙眼转录组SSR核苷酸基序类型来看, 其11 546个SSR位点包含61种重复基元, 二核苷酸至六核苷酸重复基元分别有4、10、19、13、12种; 从分布频率来看, 出现最多的重复基元是AG/CT (4 250个, 占36.81%)和AAG/CCT (1 109个, 占9.61%)。二核苷酸重复基元以AG/CT、AC/GT和AT/AT为主, 占二核苷酸重复基元总数的99.67%。三核苷酸重复基元以AAG/CTT、ACC/GGT和AGG/CCT最多, 分别占三核苷酸重复基元总数的20.81%、16.69%和13.70% (表2)。四核苷酸重复基元以AAAG/CTTT、AAAC/GTTT和AAAT/ATTT最多, 分别占四核苷酸重复基元总数的15.2%、12.87%

和12.28%; 五核苷酸重复基元以AACAC/GTGTT和AATTC/AATTG出现频率最高, 分别占其总数的13.33%和13.33%; 六核苷酸重复基元出现频率最高的是AAGAGG/CCTCTT、ACTCAG/AGTCTG和AGCCGC/CGGCTG, 其发生频率均占其总数的13.33%。

## 2.3 龙眼转录组SSR引物的有效性筛选及多态性验证

为含SSR位点的9 571条unigene序列设计引物, 共设计了8 347对SSR位点特异引物。为了验证其引物的有效性, 随机挑选合成了50对EST-SSR引物, 包括二核苷酸、三核苷酸、四核苷酸、五核苷酸及六核苷酸重复基元的SSR位点(表3)。

以‘石硤’、‘储良’、‘古山2号’、‘立冬本’四份龙眼材料的基因组DNA为模板对这批引物进行PCR扩增、筛选, 结果表明, 其中的21对引物(表3)能产生理想的PCR产物, 有效扩增率为42%。

利用21个有效EST-SSR引物进行扩增、多态性评价, 结果表明, 16对引物扩增条带具有多态性, 占有效引物的76.2%。16对引物共得到50个条带, 其中多态性片段21个, 每对引物平均产生1.31个多态性片段(图1)。

## 3 讨论

SSR分子标记技术凭借其突出的优越性, 被广泛应用于种质鉴定与分析、亲缘关系及遗传多样性分析、构建遗传连锁图谱等研究领域。长久以来, SSR的发展受到基因组数据库和EST-SSR序列缺乏的限制。随着新一代测序技术的飞速发展, 大量的转录组数据和EST序列被提交到公共数据库中, 成为开发SSR标记的可利用资源。

表1 龙眼转录组中EST-SSR类型、数量及分布频率

Table 1 Types, numbers and frequencies of EST-SSRs in longan transcriptome

| 重复类型 | 数量    |       |       |       |     |     |     | 总计     | 分布频率/% |
|------|-------|-------|-------|-------|-----|-----|-----|--------|--------|
|      | 5     | 6     | 7     | 8     | 9   | 10  | >10 |        |        |
| 二核苷酸 | 0     | 2 105 | 1 309 | 956   | 786 | 641 | 260 | 6 017  | 52.11  |
| 三核苷酸 | 3 091 | 1 444 | 709   | 80    | 1   | 0   | 3   | 5 328  | 46.15  |
| 四核苷酸 | 133   | 37    | 1     | 0     | 0   | 0   | 0   | 171    | 1.48   |
| 五核苷酸 | 12    | 1     | 0     | 0     | 2   | 0   | 0   | 15     | 0.13   |
| 六核苷酸 | 7     | 3     | 2     | 1     | 1   | 0   | 1   | 15     | 0.13   |
| 合计   | 3 243 | 3 500 | 2 021 | 1 037 | 790 | 641 | 264 | 11 546 | 100    |

表2 龙眼转录组EST-SSR中二碱基和三碱基重复基元的类型及发生频率

Table 2 Dinucleotide and trinucleotide EST-SSR repeat motifs and their frequencies in ESTs of longan transcriptome

| 重复基元类型  | 发生频率数 | 占总发生频率的比例/% |
|---------|-------|-------------|
| AC/GT   | 872   | 3.90        |
| AG/CT   | 4 250 | 19.01       |
| AT/AT   | 875   | 3.91        |
| CG/CG   | 20    | 0.09        |
| AAC/GTT | 585   | 2.62        |
| AAG/CTT | 1 109 | 4.96        |
| AAT/ATT | 380   | 1.70        |
| ACC/GGT | 889   | 3.98        |
| ACG/CGT | 172   | 0.77        |
| ACT/AGT | 93    | 0.42        |
| AGC/CTG | 453   | 2.03        |
| AGG/CCT | 730   | 3.27        |
| ATC/ATG | 726   | 3.25        |
| CCG/CGG | 191   | 0.85        |

本研究通过搜索项目组前期获得的龙眼顶芽转录组数据库序列,从114 445条龙眼顶芽转录组uni-gene (序列总长约86 800.57 kb)序列中发现了11 546个SSR位点,SSR出现频率为10.09%,其中1 975条uni-gene含有两个或两个以上EST-SSR位点,占有所有SSR位点的比例为17.10%,SSRs出现的平均距离为7.52 kb。这个SSR出现频率结果明显优于洪仕南等(2015)利用‘红核子’龙眼胚性愈伤组织转录组数据获得的EST-SSRs。与已有报道的其他树种的SSR出现比例相比,低于萝卜(*Raphanus sativus*)的23.79%(Wang等2012)、柑橘(*Citrus reticulata*)的21.74%(Jiang等2006)、茶(*Camellia sinensis*)的14.70%(Wu等2013)和腊梅(*Chimonanthus praecox*)的12.35%(李响等2013),但是优于杜仲(*Eucommia ulmoides*)的2.90%(黄海燕等2013)、胡桃(*Juglans regia*)的3.97%(冯怡等2011)和红豆杉(*Taxus chinensis*)的2.07%(李炎林等2014)。各物种SSRs的出现比例与物种间的真实SSR信息差异有关,但是也受到SSR查找软件、SSR长度设定以及EST数据库中的数据量及来源等主观条件的影响。

从龙眼SSR出现频率来看,二核苷酸(52.11%)和三核苷酸(46.15%)出现频率最高,占有所有核苷酸出现频率的99.26%,这个结果与洪仕南等(2015)利用‘红核子’龙眼胚性愈伤组织获得的结果一致。前人在茶树(杨华等2011)、荔枝(孙清明等2013)、

水仙(祁世明2014)等其他植物中的研究结果表明在转录组SSR位点中二核苷基元和三核苷基元是优势重复单元,二核苷基元和三核苷基元数目在总SSRs中的占比最高达98.66%(水仙),最低为69.09%(荔枝),这与我们的研究结果一致。在龙眼SSR二核苷酸重复基元中出现频率最高的是AG/CT,这与在刺梨(*Rosa roxburghii*)、芙蓉李(*Prunus salicina*)、蔷薇科果树杏(*Prunus armeniaca*)、桃(*P. persica*) (Jung等2005)中的研究结果一致。龙眼SSR中最丰富的三核苷酸重复基元为AAG/CTT,这与双子叶植物中三核苷酸主要重复类型是AAG/CTT的观点相符(Morgante等2002)。

本研究中,我们从设计的8 347对SSR位点特异引物中随机挑选合成50对EST-SSR引物进行引物的有效性检测。在合成的50对SSR引物中,有21对引物能够扩增出理想的PCR产物,有效扩增率为42%;16对引物具有多态性,占可扩增引物的76.2%。龙眼转录组SSR多态性比例高于辣椒的36.84%(刘峰等2012)、连翘(*Forsythia suspensa*)的37.5%(王兴春等2015)、红豆杉的38.71%(李炎林等2014)、洋葱(*Allium cepa*)的45%(李满堂等2015)、马铃薯(*Solanum tuberosum*)的47.5%(巩楠等2015)、野三七(*Panax stipuleanatus*)的50%(李翠婷等2014)和甜瓜(*Cucumis melo*)的73.3%(胡建斌和李建吾2009),低于草莓(*Fragaria × ananassa*)的85.71%(董清华等2011)。转录组SSR多态性高低可能受DNA转录序列的保守性、材料数量及材料之间差异程度等多种因素的影响(鄢秀芹等2015;巩楠等2015)。本研究采用琼脂糖凝胶,而未采用常用的聚丙烯酰胺凝胶进行SSR扩增产物的分离,琼脂糖凝胶分辨率较低,导致16对龙眼多态性引物只获得多态性片段21个。本研究结果表明采用龙眼转录组数据进行SSR标记开发是可行的,开发出来的龙眼转录组SSR标记可用于龙眼种质资源遗传多样性分析、分子标记辅助育种、遗传图谱构建以及功能基因的挖掘等方面。

#### 参考文献(References)

- Chen C, Bowman KD, Choi YA, et al (2007). EST-SSR genetic maps for *Citrus sinensis* and *Poncirus trifoliata*. *Tree Genet Genomes*, 4 (1): 1–10
- Chen H, He X, Zhu J, et al (2010). Genetic relationships analysis of 37 longan germplasm resources by using ISSR

表3 50对龙眼EST-SSR引物信息

Table 3 Information of 50 pairs of EST-SSR primers developed from longan

| 引物编号 | 重复单元                  | 引物序列  | 预期产物/bp | 扩增带数 | 多态性带数 | 多态率/% |
|------|-----------------------|---|---------|------|-------|-------|
| 1    | (CCATCA) <sub>5</sub> | 正向: TGGGAGCCGATTCTGGTAGA<br>反向: CCAAGAACAACGGTCCCCAA    | 232     | 0    | 0     | 0     |
| 2    | (AGATG) <sub>5</sub>  | 正向: TGGGACTGTTGTTGGTGCTT<br>反向: CCGGAGTCCAAGATTGCTT     | 186     | 3    | 1     | 33.3  |
| 3    | (AAAT) <sub>5</sub>   | 正向: TGGGAACAGACGAGGGTCAT<br>反向: TGCATTTATGACTGGAAGCCA   | 197     | 2    | 0     | 0     |
| 4    | (AAAG) <sub>5</sub>   | 正向: TGGCCAGTCAGCAATTTGGA<br>反向: ACGAAGCATCATCTTGACAATCC | 215     | 0    | 0     | 0     |
| 5    | (CTTT) <sub>5</sub>   | 正向: TGGGAAACATGGAGTTGGCT<br>反向: TGGATGGTGAGAAGCCAACC    | 253     | 0    | 0     | 0     |
| 6    | (AC) <sub>6</sub>     | 正向: GCAATGCAGTTGTTGGGTGT<br>反向: GCAACGTCTTGCTCTGCTG     | 169     | 2    | 1     | 50    |
| 7    | (AC) <sub>7</sub>     | 正向: GCACGCAGCAGACGTTTATC<br>反向: GCGATTGCAAAGGCTATGGG    | 154     | 0    | 0     | 0     |
| 8    | (AC) <sub>9</sub>     | 正向: GCAGAGAATGCTCCCCCTCC<br>反向: GGGGTGTTGGGGATGGAAT     | 207     | 3    | 2     | 66.7  |
| 9    | (AG) <sub>10</sub>    | 正向: GCAGATTCCACCCACAAGA<br>反向: TCGAAGAGCTGTGCAAGGAG     | 218     | 2    | 0     | 0     |
| 10   | (AG) <sub>6</sub>     | 正向: GCAGCAACAACAACAACAGC<br>反向: CTGCTGTTGCTGTTGCTGAG    | 193     | 0    | 0     | 0     |
| 11   | (AG) <sub>7</sub>     | 正向: GCCATTGTCATTCATGGCGT<br>反向: CGCCATCATCCCTTCTCCTC    | 225     | 0    | 0     | 0     |
| 12   | (AT) <sub>10</sub>    | 正向: GCGAAAGAGAGGGAGAGACG<br>反向: AGGTGCGTAAATCCAAGGCA    | 178     | 2    | 1     | 50    |
| 13   | (AT) <sub>7</sub>     | 正向: GCGGGTTGGTGATGATGTTG<br>反向: CAAAAGTCGACCCAGACCGA    | 166     | 0    | 0     | 0     |
| 14   | (CA) <sub>6</sub>     | 正向: GCTCTGCTTTGGAGGATGGT<br>反向: AGCTGTGCTGATGAGCTCAA    | 195     | 0    | 0     | 0     |
| 15   | (CA) <sub>9</sub>     | 正向: GCTGCAGCTTTCAGATGGA<br>反向: CCCCAACTGCACCCACTAAT     | 203     | 0    | 0     | 0     |
| 16   | (CG) <sub>6</sub>     | 正向: GCTGCGTTAGTCCCAAAGC<br>反向: TGGTGAATCAGCAGCAGAG      | 158     | 3    | 1     | 33.3  |
| 17   | (CT) <sub>10</sub>    | 正向: GCTGGATGCCCTTGTTTTG<br>反向: CGCTTACCTCCCACCTGAAA     | 235     | 0    | 0     | 0     |
| 18   | (CT) <sub>10</sub>    | 正向: GCTGGCAATGTCAAGGTTGG<br>反向: ACTCTGGAAGTCATGGAACCA   | 222     | 0    | 0     | 0     |
| 19   | (CT) <sub>6</sub>     | 正向: GCTGGCTGCATGGTTTTTCA<br>反向: GTGATTTGCTGCTGCTGCT     | 174     | 4    | 1     | 25    |
| 20   | (CT) <sub>7</sub>     | 正向: GGAGGAGGAGGAGGAGGATG<br>反向: ATCATTACTTGCCGCCACCA    | 192     | 0    | 0     | 0     |
| 21   | (GA) <sub>10</sub>    | 正向: GGCGGGTCAATGTGAAATGT<br>反向: ACGTGAGCTCAGACTAGGGT    | 163     | 3    | 1     | 33.3  |
| 22   | (GA) <sub>11</sub>    | 正向: GGCGGGTCTGATTAGTGCTT<br>反向: ACAATCGTCACCTGCATCCA    | 192     | 0    | 0     | 0     |
| 23   | (GA) <sub>6</sub>     | 正向: GGCGTCCATGGACCTCTATG<br>反向: GGGAGAGCCACCTTGTCATT    | 214     | 4    | 2     | 50    |
| 24   | (GA) <sub>8</sub>     | 正向: GGGGTTTTGTGGTAGGACGT<br>反向: CCACTGTAACAGAAGTTGCGC   | 188     | 0    | 0     | 0     |
| 25   | (GA) <sub>9</sub>     | 正向: GGGTAAAGCCATTCAGGTTGC<br>反向: GCCATTGAATCTGCTGATCACT | 134     | 2    | 1     | 50    |

表3 (续)

| 引物编号 | 重复单元               | 引物序列   | 预期产物/bp | 扩增带数 | 多态性带数 | 多态率/% |
|------|--------------------|--|---------|------|-------|-------|
| 26   | (GT) <sub>10</sub> | 正向: GGGTGGTTCTGTCTGTGCT<br>反向: TCTGGTAACTGCTAGCGCTG  | 125     | 0    | 0     | 0     |
| 27   | (GT) <sub>6</sub>  | 正向: GGGTGTATCAGCCGTGGTAG<br>反向: TCGACAGCAGTTCCACCAAG | 210     | 0    | 0     | 0     |
| 28   | (TA) <sub>6</sub>  | 正向: GGTCTGCTGATGCCTGAAGT<br>反向: TCCGACATCAACAAAATGGC | 188     | 5    | 1     | 20    |
| 29   | (TA) <sub>7</sub>  | 正向: GGTGTGGGTTGTGGATGGAT<br>反向: AGGACACTTTAAGGTCCCCT | 173     | 2    | 0     | 0     |
| 30   | (TC) <sub>6</sub>  | 正向: GTCGGCAGTGACAGAACTCA<br>反向: CCAAAGGTGCGTCAACGAAG | 128     | 0    | 0     | 0     |
| 31   | (TC) <sub>7</sub>  | 正向: GTGATGGCCAAGAGGAGCTT<br>反向: AGAGTCTTACCAGCACACGC | 145     | 0    | 0     | 0     |
| 32   | (TG) <sub>6</sub>  | 正向: GTTTCACCACCGCAACTTC<br>反向: GTTGGGAAGGAGGTGATGG   | 162     | 3    | 1     | 33.3  |
| 33   | (AAC) <sub>5</sub> | 正向: TCTCCCTCTCTGTGTGTGC<br>反向: TCTGTCCACCCAAAACCACC  | 214     | 2    | 0     | 0     |
| 34   | (AAG) <sub>5</sub> | 正向: TCGTCATCGTCGTCGTCATC<br>反向: GGTTTACAGGGGAGGCTGTT | 226     | 0    | 0     | 0     |
| 35   | (AAT) <sub>5</sub> | 正向: TCCGCTTCTCTCCTTGCAAG<br>反向: TTTCAAGGCCTTCCCACCC  | 118     | 3    | 2     | 66.7  |
| 36   | (ACA) <sub>5</sub> | 正向: TATCGAGGGATACGCTGGGG<br>反向: GCATGGGCAATTTGTGGACT | 139     | 0    | 0     | 0     |
| 37   | (ACA) <sub>5</sub> | 正向: TATGCACCACAAACCACCA<br>反向: GCAGGATTTAAAGCATGCCGA | 216     | 4    | 1     | 25    |
| 38   | (ACC) <sub>5</sub> | 正向: TCCATTGCATGTCCTCGCTT<br>反向: GTGGCCAATGCACCATCTTG | 143     | 0    | 0     | 0     |
| 39   | (ACG) <sub>5</sub> | 正向: TCCTCCTCCTCAAGACCAG<br>反向: GCACTCAAACACACTCGCAG  | 156     | 0    | 0     | 0     |
| 40   | (ACT) <sub>5</sub> | 正向: TCCTTGTGTCAGTTTGGGT<br>反向: GTCAGAACACCAGCTCCGAT  | 224     | 0    | 0     | 0     |
| 41   | (AGA) <sub>5</sub> | 正向: TACCGAATGCCAAGTGAGG<br>反向: GTGGCGCTTGAATAGTTGCC  | 208     | 3    | 1     | 33.3  |
| 42   | (AGC) <sub>5</sub> | 正向: TCAACGCTAAGACGCTTCTCA<br>反向: CGCATTATCAGCCTCGTTG | 196     | 0    | 0     | 0     |
| 43   | (AGC) <sub>5</sub> | 正向: TCACAACAACACAACTCCCT<br>反向: GAGACCATCAGAGCCGATGG | 182     | 0    | 0     | 0     |
| 44   | (AGG) <sub>5</sub> | 正向: TCTCGTTCATTTGCACACGC<br>反向: TCACCCAAGTGCACAGATCC | 143     | 4    | 2     | 50    |
| 45   | (AGT) <sub>5</sub> | 正向: TGCCACAAGCCTGTTACTGG<br>反向: AGCAGATACCTTTGGCACCC | 116     | 0    | 0     | 0     |
| 46   | (ATA) <sub>5</sub> | 正向: TCAGGAGGAAAGGCAAGCAA<br>反向: CCCCCACAAAAGAGAGGCAA | 132     | 0    | 0     | 0     |
| 47   | (ATC) <sub>5</sub> | 正向: TCGAATGCTGCCTTGTACGT<br>反向: ATTTGGCAGCGAGGGATAGG | 194     | 0    | 0     | 0     |
| 48   | (ATG) <sub>5</sub> | 正向: TCGGTTATGGGGAGGTAGCA<br>反向: AACCCCTGCATTGACACAA  | 178     | 2    | 0     | 0     |
| 49   | (ATT) <sub>6</sub> | 正向: TCTCCGAAGACCGAGAAAGT<br>反向: TGGTGGAGGTGAGGTTTTG  | 156     | 2    | 1     | 50    |
| 50   | (TGG) <sub>5</sub> | 正向: TAGCCGGGAATGTAGGGACA<br>反向: GCGATTGCACTCTCCCTT   | 215     | 0    | 0     | 0     |

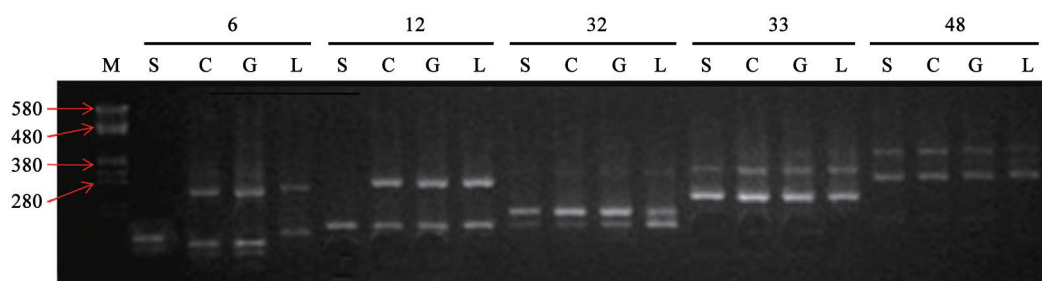


图1 引物6、12、32、33、48在四个龙眼品种中扩增情况

Fig.1 Amplification of primers 6, 12, 32, 33 and 48 showed in four longan species

S: ‘石硖’; C: ‘储良’; G: ‘古山2号’; L: ‘立冬本’。

- markers. *Genom Appl Biol*, 29 (2): 288–292 (in Chinese with English abstract) [陈虎, 何新华, 朱建华等(2010). 37份龙眼种质资源亲缘关系的ISSR分析. *基因组学与应用生物学*, 29 (2): 288–292]
- Chen Y, Liu C (2001). RAPD identification and analysis of longan varieties. *China Fruits*, (4): 28–29 (in Chinese) [陈有志, 刘成明(2001). 龙眼品种RAPD鉴别及分析. *中国果树*, (4): 28–29]
- Dong XH, Wang XC, Zhao MZ, et al (2011). Development of EST-derived SSR markers and their application in strawberry genetic diversity analysis. *Sci Agr Sin*, 44 (17): 3603–3612 (in Chinese with English abstract) [董清华, 王西成, 赵密珍等(2011). 草莓EST-SSR标记开发及在品种遗传多样性分析中的应用. *中国农业科学*, 44 (17): 3603–3612]
- Feng Y, Zhang Z, Zhang S, et al (2011). Development of walnut EST-SSR markers and primer design. *J Anhui Agr Sci*, 39 (36): 22217–22219, 22276 (in Chinese with English abstract) [冯怡, 张智俊, 张舍龙等(2011). 胡桃EST-SSR标记的筛选及引物设计. *安徽农业科学*, 39 (36): 22217–22219, 22276]
- Gao HY, Jiang F, Chen XP, et al (2007). Random amplification polymorphic DNA (RAPD) analysis of the typical longan genotypes from different origins. *J Fujian For Sci Tech*, 34 (1): 67–71 (in Chinese with English abstract) [高慧颖, 姜帆, 陈秀萍等(2007). 不同地域的代表性基因型龙眼RAPD分析. *福建林业科技*, 34 (1): 67–71]
- Gong L, Cheng Y, Gan X, et al (2015). Mining EST-SSRs based on potato transcriptome and analyzing their polymorphism. *Mol Plant Breeding*, 13 (7): 1535–1544 (in Chinese with English abstract) [巩樵, 程永芳, 甘晓燕等(2015). 马铃薯转录组EST-SSR挖掘及其多样性分析. *分子植物育种*, 13 (7): 1535–1544]
- Han RX, Zhang ZW, Wu B, et al (2012). Development of SSR markers and application in analysis of genetic diversity in tartary buckwheat (*Fagopyrum tataricum*). *J Plant Genet Resour*, 13 (5): 759–764 (in Chinese with English abstract) [韩瑞霞, 张宗文, 吴斌等(2012). 苦荞SSR引物开发及其在遗传多样性分析中的应用. *植物遗传资源学报*, 13 (5): 759–764]
- Hong S (2015). Development and application of SSR primers based on the transcriptome in *Dimocarpus longan* Lour. (dissertation). Fuzhou: Fujian Agriculture and Forestry University (in Chinese with English abstract) [洪仕南(2015). 基于龙眼转录组的SSR引物开发与应用研究(学位论文). 福州: 福建农林大学]
- Hu JB, Li JW (2009). Information on EST-SSR loci in melon (*Cucumis melo* L.) and marker exploitation. *Acta Horti Sin*, 36 (4): 513–520 (in Chinese with English abstract) [胡建斌, 李建吾(2009). 甜瓜EST-SSR位点信息及标记开发. *园艺学报*, 36 (4): 513–520]
- Huang H, Du H, Wuyun T, et al (2013). Development of SSR molecular markers based on transcriptome sequencing of *Eucommia ulmoides*. *Sci Silv Sin*, 49 (5): 176–181 (in Chinese with English abstract) [黄海燕, 杜红岩, 乌云塔娜等(2013). 基于杜仲转录组序列的SSR分子标记的开发. *林业科学*, 49 (5): 176–181]
- Jiang D, Zhong GY, Hong QB (2006). Analysis of microsatellites in citrus unigenes. *Acta Genet Sin*, 33 (4): 345–353
- Jung S, Abbott A, Jesudurai C, et al (2005). Frequency, type, distribution and annotation of simple sequence repeats in Rosaceae ESTs. *Funct Integr Genomics*, 5: 136–143
- Li CT, Zhang GH, Ma CH, et al (2014). Analysis on SSR loci information in transcriptome of *Panax vietnamensis* var. *fuscidiscus* and its polymorphism. *Chin Tradit Herb Drugs*, 45 (10): 1468–1472 (in Chinese with English abstract) [李翠婷, 张广辉, 马春花等(2014). 野三七转录组中SSR位点信息分析及其多态性研究. *中草药*, 45 (10): 1468–1472]
- Li MT, Zhang SL, Deng P, et al (2015). Analysis on SSR information in transcriptome of onion and the polymorphism. *Acta Horti Sin*, 42 (6): 1103–1111 (in Chinese with English abstract) [李满堂, 张仕林, 邓鹏等(2015). 洋葱转录组SSR信息分析及其多态性研究. *园艺学报*, 42 (6): 1103–1111]
- Li SK, Jiang C, Wang JY (2011). Genetic diversity of *Oryza*

- rufipogon* Griff. in Zhangpu County of Fujian Province by using SSR markers. *J Plant Genet Resour*, 12 (1): 75–79 (in Chinese with English abstract) [李书柯, 江川, 王金英(2011). 用SSR标记分析福建漳浦野生稻的遗传多样性. *植物遗传资源学报*, 12 (1): 75–79]
- Li X, Yang N, Zhao KG, et al (2013). Development and primer selection of EST-SSR molecular markers based on transcriptome sequencing of *Chimonanthus praecox*. *J Beijing For Univ*, 35 (S1): 25–32 (in Chinese with English abstract) [李响, 杨楠, 赵凯歌等(2013). 腊梅转录组EST-SSR标记开发与引物筛选. *北京林业大学学报*, 35 (S1): 25–32]
- Li YL, Yang XX, Zhang JY, et al (2014). Studies on SSR molecular markers based on transcriptome of *Taxus chinensis* var. *mairei*. *Acta Horti Sin*, 41 (4): 735–745 (in Chinese with English abstract) [李炎林, 杨星星, 张家银等(2014). 南方红豆杉转录组SSR挖掘及分子标记的研究. *园艺学报*, 41 (4): 735–745]
- Liu F, Wang YS, Tian XL, et al (2012). SSR mining in pepper (*Capsicum annuum* L.) transcriptome and the polymorphism analysis. *Acta Horti Sin*, 39 (1): 168–174 (in Chinese with English abstract) [刘峰, 王运生, 田雪亮等(2012). 辣椒转录组SSR挖掘及其多态性分析. *园艺学报*, 39 (1): 168–174]
- Morgante M, Hanafey M, Powell W (2002). Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nat Genet*, 30: 194–200
- Pan LM, Zhu JH, Peng HX, et al (2008). Research advances in relatives of longan in China. *J Guangxi Agr Sci*, 39 (6): 811–814 (in Chinese with English abstract) [潘丽梅, 朱建华, 彭宏祥等(2008). 我国龙眼近缘植物研究进展. *广西农业科学*, 39 (6): 811–814]
- Qi S (2014). Development and application of SSR markers based on BAC end sequences of *Narcissus* (dissertation). Fuzhou: Fujian Agriculture and Forestry University (in Chinese with English abstract) [祁世明(2014). 基于水仙BAC末端序列的SSR标记开发与应用(学位论文). 福州: 福建农林大学]
- Sun Q, Li Y, Xiang X, et al (2013). A novel litchi germplasm (*Litchi chinensis* Sonn.), Yujinqiu, indentified by EST-SSR and SNP analysis. *Mol Plant Breeding*, 11 (3): 403–414 (in Chinese with English abstract) [孙清明, 李永忠, 向旭等(2013). 利用SNP和EST-SSR分子标记鉴定荔枝新种质御金球. *分子植物育种*, 11 (3): 403–414]
- Wang JX, Wang H, Lai JY, et al (2011). Advance in identity technology of germplasm resource in *Juglans*. *Nonwood For Res*, 29 (1): 161–165 (in Chinese with English abstract) [王金星, 王滑, 赖家业等(2011). 核桃属植物种质资源鉴定技术研究进展. *经济林研究*, 29 (1): 161–165]
- Wang S, Wang X, He Q, et al (2012). Transcriptome analysis of the roots at early and late seedling stages using Illumina paired-end sequencing and development of EST-SSR markers in radish. *Plant Cell Rep*, 31: 1437–1447
- Wang X, Tan H, Chen Z, et al (2015). Assembly and characterization of the transcriptome and development of SSR markers in *Forsythia suspensa* based on RNA-Seq technology. *Sci Sin Vitae*, 45 (3): 301–310 (in Chinese with English abstract) [王兴春, 谭河林, 陈钊等(2015). 基于RNA-Seq技术的连翘转录组组装与分析及SSR分子标记的开发. *中国科学: 生命科学*, 45 (3): 301–310]
- Wu H, Chen D, Li J, et al (2013). De novo characterization of leaf transcriptome using 454 sequencing and development of EST-SSR markers in tea (*Camellia sinensis*). *Plant Mol Biol Rep*, 31: 524–538
- Xu Q, Li T, Chen X, et al (2008). RAPD analysis on 24 longan germplasm resources. *J Xiamen Univ-Nat Sci*, 47 (S2): 26–29 (in Chinese with English abstract) [许奇志, 李韬, 陈秀萍等(2008). 24份龙眼种质资源RAPD分析. *厦门大学学报(自然科学版)*, 47 (S2): 26–29]
- Yan XQ, Lu M, An HM (2015). Analysis on SSR information in transcriptome and development of molecular markers in *Rosa roxburghii*. *Acta Horti Sin*, 42 (2): 341–349 (in Chinese with English abstract) [鄢秀芹, 鲁敏, 安华明(2015). 刺梨转录组SSR信息分析及其分子标记开发. *园艺学报*, 42 (2): 341–349]
- Yang H, Chen Q, Wei CL, et al (2011). Analysis on SSR information in *Camellia sinensis* transcriptome. *J Anhui Agr Univ*, 38 (6): 882–886 (in Chinese with English abstract) [杨华, 陈琪, 韦朝领等(2011). 茶树转录组中SSR位点的信息分析. *安徽农业大学学报*, 38 (6): 882–886]
- Yi G, Tai W, Huo H, et al (2003). Studies on the genetic diversity and relationship of longan cultivars by AFLP analysis. *Acta Horti Sin*, 30 (3): 272–276 (in Chinese with English abstract) [易干军, 谭卫萍, 霍合强等(2003). 龙眼品种(系)遗传多样性及亲缘关系的AFLP分析. *园艺学报*, 30 (3): 272–276]
- Yonemoto Y, Chowdhury AK, Kato H, et al (2006). Cultivars identification and their genetic relationships in *Dimocarpus longan* subspecies based on RAPD markers. *Sci Horti*, 109 (2): 147–152
- Yuan JH (2011). Analysis of genetic diversity and inter-specific relationships of *Lycoris* (Amaryllidaceae) from China by SSR-PCR. *Subtrop Plant Sci*, 40 (4): 1–7 (in Chinese with English abstract) [袁菊红(2011). 中国石蒜属遗传多样性与亲缘关系的SSR分析. *亚热带植物科学*, 40 (4): 1–7]
- Zeng LH, Hong ZT, Lin WZ, et al (2009). ISSR analysis of germplasm in longan. *J Fujian Agr For Univ-Nat Sci*, 38 (3): 239–242 (in Chinese with English abstract) [曾黎辉, 洪自同, 林文忠等(2009). 龙眼种质资源的ISSR分析. *福建农林大学学报(自然科学版)*, 38 (3): 239–242]
- Zhai SB, Su LC, Fang YN, et al (2016). Effects of KClO<sub>3</sub> on the expression of *CONSTANS* family genes in the control of flower bud formation of longan. *Plant Physiol J*, 52



- (9): 1429–1437 (in Chinese with English abstract) [翟世博, 苏良辰, 方妍妮等(2016). KClO<sub>3</sub>处理对龙眼花芽形成关键基因*CONSTANS*表达的影响. 植物生理学报, 52 (9): 1429–1437]
- Zhao YH, Hu YL, Guo YS, et al (2008). Inter-generic hybrids obtained from cross between litchi and longan cultivars and their molecular identification. *J Fruit Sci*, 25 (6): 950–952 (in Chinese with English abstract) [赵玉辉, 胡又厘, 郭印山等(2008). 荔枝、龙眼属间远缘杂种的获得及分子鉴定. 果树学报, 25 (6): 950–952]
- Zhen S, Zhang LJ, Zhang XY, et al (2012). DNA fingerprinting of 17 *Dimocarpus longan* germplasms by using ISSR markers. *Fujian J Agr Sci*, 27 (12): 1308–1312 (in Chinese with English abstract) [郑姗, 张立杰, 张小艳等(2012). 龙眼新品种(品系)的ISSR指纹图谱研究初探. 福建农业学报, 27 (12): 1308–1312]

## Analysis on simple sequence repeat (SSR) information in apical transcriptome and development of molecular markers in *Dimocarpus longan*

GUO Dong-Liang, WANG Jing, HAN Dong-Mei, PAN Xue-Wen, LI Jian-Guang\*

*Institute of Fruit Tree Research, Guangdong Academy of Agricultural Sciences; Key Laboratory of South Subtropical Fruit Biology and Genetic Resource Utilization, Ministry of Agriculture, Guangzhou 510641, China*

**Abstract:** With the development of next-generation sequencing technologies, a large number of transcriptome data and EST sequences have become available resources for developing simple sequence repeat (SSR) markers. Transcriptome of *Dimocarpus longan* terminal buds was screened using MISA software. A total of 11 546 SSRs that occurred in 114 445 unigenes was identified and the frequency of these SSRs was 10.09%. 1 975 unigenes contain two and more SSRs, and the frequency of two and more SSRs was 17.10%. The mean distance of SSRs was 7.52 kb in the unigenes. Among longan transcriptome SSRs, dinucleotide and trinucleotide were major repeat types, accounting for 52.11% and 46.15%, respectively. Furthermore, AG/CT and AAG/CTT were the most frequent motifs in dinucleotide and trinucleotide repeats, accounting for 36.81% and 9.61%, respectively. 8 347 pairs of SSR primers were designed for 9 571 SSR loci using Primer 5.0 software and 50 pairs of primers were randomly selected to evaluate their polymorphism across four varieties, including ‘Shixia’, ‘Chuliang’, ‘Gushan No.2’ and ‘Lidongben’. The results show that the PCR products of 21 primers were clear and effective and the effective ratio was 42%. Sixteen pairs of primers were polymorphic, and their polymorphic percentage reached 76.2%. Among 50 amplification bands, 21 polymorphic bands were obtained by using 16 pairs polymorphic primers, and the average polymorphic amplification band of each EST-SSR primer was 1.31.

**Key words:** longan; transcriptome; SSR

Received 2018-03-02 Accepted 2018-03-29

This work was supported by the National Natural Science Foundation of China (31301741), and the President Foundation of the Guangdong Academy of Agricultural Sciences (201309).

\*Corresponding author (550777196@qq.com).