DOI: 10.3724/SP.J.1006.2020.94158

基于黄麻转录组序列 SNP 位点的 CAPS 标记开发与验证

陶爱芬^{1,2} 游梓翊² 徐建堂² 林荔辉² 张立武² 祁建民^{2,*} 方平平^{2,*}

¹福建农林大学金山学院, 福建福州 350002; ²福建农林大学教育部作物遗传育种与综合利用重点实验室 / 福建省分子设计育种实验室, 福建福州 350002

摘要: 黄麻 CAPS 分子标记的开发, 可以为黄麻遗传多样性分析、种质资源鉴定和分子标记辅助选择等研究提供有效方法。本试验以黄麻 179 和爱店野生种为材料, 采用 Illumina HiSeq 4000 测序技术进行转录组测序, 对其 SNP 位点进行分析, 设计与木质素合成基因 4CL、COMT 及转录因子 MYB 相关的 SNP 引物, 在此基础上, 应用 dCAPS Finder2.0 软件开发了 CAPS 标记, 并对其有效性进行了验证。结果表明: (1) 组装后的黄麻 unigene 序列共 72,674 条, 序列总长度为 29,705,997 bp, 检测到的 SNP 位点总数为 67,567 个, 平均每 440 bp 有 1 个 SNP。 (2) 获得了 39 对分别与 4CL、COMT 和 MYB 相关的 SNP 引物, 从中筛选获得 26 对 CAPS 标记引物, 开发成功率为 66.7%, 其中 11 对 CAPS 标记具有多态性, 多态性比例为 43.2%。 (3) 开发的 CAPS 标记能较好地 12 份不同类型的黄麻种质区分开来, 表明 CAPS 是适用于黄麻研究的较理想的分子标记方法, 为黄麻遗传基础研究提供了可靠工具。

关键词: 黄麻; 转录组; SNP 位点; CAPS 标记; 开发

Development and verification of CAPS markers based on SNPs from transcriptome of jute (*Corchorus L.*)

TAO Ai-Fen^{1,2}, YOU Zi-Yi², XU Jian-Tang², LIN Li-Hui², ZHANG Li-Wu², QI Jian-Min^{2,*}, and FANG Ping-Ping^{2,*}

¹ Jinshan College, Fujian Agriculture and Forestry University, Fuzhou 350002, Fujian, China; ² Key Laboratory of Ministry of Education for Genetics, Breeding and Multiple Utilization of Crops / Key Laboratory of Crops by Design, Fujian Agriculture and Forestry University, Fuzhou 350002, Fujian, China

Abstract: Developing CAPS markers can provide a new and effective tool for genetic diversity analysis, germplasm identification and marker-assistant selection of jute. The transcriptional sequence analysis was carried out by Illumina HiSeq 4000 with jute 179 and the wild species in Aidian. Furthermore, the SNP loci were screened and primers related to lignin synthesis genes 4CL, COMT and transcription factor MYB were developed. CAPS markers were screened based on above SNP primers using dCAPS Finder2.0 and their polymorphism was verified. A total of 72,674 unigene sequences were got in jute transcriptome by assembly, with a length of 29,705,997 bp, a number of SNP loci of 67,567, and an average of 1 SNP per 440 bp. Thirty-nine pairs of SNP primers related to 4CL, COMT, and MYB were designed by Oligo8, of which 26 pairs of CAPS primers were screened, with a developing rate of 66.7%. Among the 26 pairs of CAPS markers, 11 pairs were polymorphic, with a polymorphic ratio of 43.2%. The screened CAPS markers could effectively distinguish 12 different accessions of jute germplasm, indicating CAPS is an ideal molecular marker which can be a reliable tool used in genetic research of jute.

Keywords: jute; transcriptome; SNP loci; CAPS markers; development

黄麻属于椴树科(Tiliaceae)黄麻属(*Corchorus L.*) 强、吸水性好等优点。近年来, 黄麻纤维被广泛应用一年生草本植物, 其韧皮纤维具有易降解、透气性 用于建材和可降解地膜等产业^[1]。但与苧麻、亚麻

本研究由国家现代农业产业技术体系建设专项(CARS-19-E06)资助。

This study was supported by the China Agriculture Research System (CARS-19-E06).

* 通信作者(Corresponding authors): 祁建民, E-mail: qijm863@163.com, Tel: 0591-83789483; 方平平, E-mail: fangpp1990@qq.com, Tel: 0591-83789483

第一作者联系方式: E-mail: 281770126@qq.com, Tel: 0591-83789483

Received (收稿日期): 2019-10-23; Accepted (接受日期): 2020-03-24; Published online (网络出版日期): 2020-04-16.

URL: <http://kns.cnki.net/kcms/detail/11.1809.S.20200416.1246.002.html>

等相比,黄麻韧皮纤维细胞壁木质化程度较高、纤维素含量较低,导致纤维粗、硬,品质不佳,使其只适合加工绳索和包装用麻袋等,限制了黄麻在纺织业等工业领域的应用^[2]。目前,国内外研究人员已克隆了包括 4CL 和 COMT 在内的多个木质素生物合成途径中的关键酶基因^[3]。而 MYB (V-myb avian myeloblastosis viral oncogene homolog)转录因子是植物转录因子家族中最大的成员之一,它参与并对植物生长和发育的各个方面产生重大影响,在植物次生代谢调节、激素和环境因子反应、细胞分化、器官形态发生和细胞周期调节中均有重要作用^[4]。若能开发出与黄麻木质素合成和 MYB 转录因子相关的标记,则将为黄麻种质资源鉴定、育种亲本选配和遗传学研究等提供重要方法。

生物的遗传因子受到多个因素的影响,并在基因组水平上产生一定的波动。而分子标记技术是直接反映 DNA 分子水平多态性的技术,与其他标记相比,具有无表型效应、分布广、多态性高等优点,并且不受环境的影响^[5-6]。近年来取得长足发展的转录组测序技术,具有通量高、分辨率高、不受限制等优点^[7-8]。利用该技术可以获得大量潜在的 SNP (Single Nucleotide Polymorphism)和 SSR (Simple Sequence Repeat)标记,并用于各种分子生物学研究^[9-10]。与其他标记方法相比,SNP 标记遗传稳定性高、位点分布广泛且富有代表性,因此已成为当前研究最多的分子标记之一^[11-12]。虽然 SNP 的检测方法众多,但大部分方法均有其不足,例如操作繁琐或者成本过高,从而限制了其在植物遗传育种领域

的应用^[13-14]。其中较经济并且适合常规分子生物学实验的方法,是把 SNP 多态转化成酶切扩增多态性 (Cleaved Amplified Polymorphic Sequences, CAPS)标记或衍生 CAPS (Derived Cleaved Amplified Polymorphic Sequences, dCAPS)标记^[15]。王彩芬等^[16]和司文洁等^[17]研究表明, CAPS 标记技术具有多态性高、共显性、所需 DNA 量少、操作简便、结果稳定可靠等优点。CAPS 分子标记技术建立以来,已经在水稻、小麦、玉米、大豆、番茄、马铃薯等作物中得到广泛应用^[18-21]。

目前,分子标记技术已广泛应用于黄麻遗传多样性分析、起源与演化、遗传连锁图谱构建等研究上^[22-26]。关于黄麻分子标记的开发已取得一些进展,例如 SSR 标记^[27-29]、SNP 标记^[30-31]都被开发利用,但是关于黄麻 CAPS 标记的开发,未见有相关文献报道,而与黄麻木质素合成基因和 MYB 转录因子相关的 CAPS 标记的开发与验证,更是未见报道。鉴于 CAPS 分子标记的优点及其在黄麻上开发应用的滞后性,开展该研究是十分必要的。

1 材料与方法

1.1 材料

1.1.1 植物材料 用于转录组测序的黄麻材料为黄麻 179 和爱店野生种。2017 年 4 月种植于福建农林大学田间科技园,采用盆栽栽培,每盆 10 株,3 次重复,出苗后 30 d,真叶长到 7~8 片时,取茎和叶混合进行转录组测序。用于 CAPS 多态性引物筛选验证的 12 份黄麻种质资源的名称和类型见表 1。所有

表 1 用于 CAPS 引物多态性验证的黄麻品种名称和类型
Table 1 Name and type of jute accessions used in the validation of CAPS markers

编号 Number	名称 Name	品种类型 Type of accessions
1	梅峰 4 号 Meifeng 4	圆果种 <i>Corchorus capsularis</i>
2	越南圆果 Yuenanyuanguo	圆果种 <i>Corchorus capsularis</i>
3	广东独尾麻 Guangdonguweima	圆果种 <i>Corchorus capsularis</i>
4	中黄麻 1 号 Zhonghuangma 1	圆果种 <i>Corchorus capsularis</i>
5	日本大分青皮 Japanese Dafenqingpi	圆果种 <i>Corchorus capsularis</i>
6	福农 1 号 Funong 1	长果种 <i>Corchorus olitoris</i>
7	马里野生种 Maliyeshengzhong	长果种 <i>Corchorus olitoris</i>
8	广西长果 Guangxichangguo	长果种 <i>Corchorus olitoris</i>
9	日本 5 号 Japanese 5	长果种 <i>Corchorus olitoris</i>
10	翠绿 Cuilyu	长果种 <i>Corchorus olitoris</i>
11	龙溪长果 Longxichangguo	长果种 <i>Corchorus olitoris</i>
12	台湾加利麻 Taiwanjialima	圆果种 <i>Corchorus capsularis</i>

材料均为盆栽, 于 2017 年 8 月份播种, 每盆 10 株, 置自然条件下生长, 长至 5~6 片真叶时, 取幼嫩叶片混合后提取 DNA。

1.1.2 试剂及仪器 DNA 提取试剂盒、限制性内切酶等购自北京百泰克生物技术有限公司, 其他常用试剂为实验室分析纯常规试剂; PCR 仪为德国 Eppendorf 公司的 MC-gradient; 电泳仪购自六一仪器厂(北京), 型号为 DYY-6。

1.2 方法

1.2.1 黄麻转录组测序和 SNP 位点鉴别 基于 Illumina HiSeq 4000 测序技术平台, 构建黄麻转录组文库并进行建库测序, 获得黄麻转录组数据。在转录组测序的基础上, 采用 GATK 针对 RNA-Seq 的 SNP 识别流程, 进行 SNP 多态性位点识别^[32], 识别标准是, (1) 35 bp 范围内连续出现的单碱基错配不超过 3 个; (2) 经过序列深度标准化的 SNP 质量值大于 2.0。

1.2.2 与 4CL、COMT 和 MYB 转录因子相关的黄麻 SNP 引物设计 在黄麻转录组序列中, 寻找与木质素合成酶基因 4CL 和 COMT 以及 MYB 转录因子相关的 unigene, 并采用 Oligo 8 设计 SNP 引物, 主要参数为: (1) 引物长度在 18~25 bp 之间; (2) PCR 预期扩增产物大小在 100~500 bp 之间; (3) 退火温度(T_m) 在 53~58 之间, 上、下游引物的 T_m 值相差不大于 4 ; (4) GC 含量在 40%~60% 之间。

1.2.3 CAPS 标记开发及内切酶的选择 采用 dCAPS Finder 2.0 进行 CAPS 引物开发和内切酶的选择, 主要步骤为, (1) 在黄麻 unigene 序列中寻找与 CAPS 标记的 PCR 扩增体系中 SNP 引物相对应的 unigene 片段; (2) 选取其 SNP 位点前后各 30 bp 长度的碱基序列; (3) 将 SNP 位点原序列和突变序列分别输入 dCAPS Finder 2.0 的对话框中, 寻找可以对位点进行切割的限制酶, 其对应的标记即为基于 SNP 的 CAPS 标记。

1.2.4 黄麻 DNA 的提取 采用北京百泰克生物公司的 BioTeke 试剂盒, 按说明书提取 DNA。

1.2.5 CAPS 标记的 PCR 扩增体系构建 采用束永俊等^[14]的 CAPS 标记的 PCR 扩增程序, 反应体系为 20 μ L, 包含 DNA 2 μ L、上下游引物各 1 μ L、PRR MIX 10 μ L、无菌水 6 μ L。PCR 程序为 95 预变性 3 min; 95 变性 30 s, 53 退火 30 s, 74 延伸 45 s, 共 9 个循环; 95 变性 30 s, 53 退火 45 s, 72 延伸 45 s, 共 14 个循环; 最后 72 延伸 10 min, 4 保存。

1.2.6 限制性核酸内切酶的酶切反应 按照内切酶的使用说明书构建限制性酶切反应体系, 反应体系可分为省时酶和非省时酶 2 种。

1.2.7 琼脂糖凝胶电泳 用 1.2% 的琼脂糖凝胶检测扩增和酶切产物, 电泳电压为 100 V, 电流为 100 mA, 电泳时间为 40 min, 电泳完成后用培清 600 型凝胶成像仪观察条带并拍照记录。

1.2.8 数据统计分析 按照扩增和酶切条带的有无计数, 当某一条带出现时赋值为“1”, 不存在时赋值为“0”, 从而把图形资料转换成数据资料。参照 Nei 等^[33]的方法求得品种之间的遗传距离, 然后用 DPS 软件中的类平均聚类法(UPGMA)对 12 份黄麻种质进行聚类分析。

2 结果与分析

2.1 黄麻转录组中 SNP 位点的数量与分布

黄麻转录组经组装后共获得 72,674 条 unigene 序列, 检测到的 SNP 位点总数为 67,567 个, 序列总长度为 29,705,997 bp。在黄麻 unigene 序列中, 平均每 440 bp 会有一个 SNP 位点。SNP 位点分布密度如表 2 和图 1 所示。由表 2 可知, 含 0~1 个 SNP 位点的 unigene 序列最多, 有 66,525 条, 占总数的 91.53%; 含 2 个以上 SNP 位点的 unigene 序列数目急剧下降, 其中含 1~2 个 SNP 位点的 unigene 序列有 2092 条, 比例为 2.88%; 含 2~3 个 SNP 位点的 unigene 序列有 1277 条, 占 1.76%; 含 3 个以上 SNP 位点的 unigene 序列总数亦较少, 有 2780 条, 比例为 3.81%。表明绝大多数黄麻 unigene 上含有单个 SNP 位点, 含有 2 个及以上 SNP 位点的 unigene 数量不多, 总计约 8% 左右, 其中含 5~8 个 SNP 位点的 unigene 数量尤其少, 比例非常低。

2.2 与 4CL、COMT 及 MYB 转录因子相关的 CAPS 标记开发

在对与黄麻木质素合成基因 4CL、COMT 及 MYB 转录因子相关的 unigene 分析的基础上, 共获得 39 对与上述基因及转录因子相关的 SNP 引物(表 3)。其中和 4CL 相关的 SNP 引物最少, 仅有 6 对, 和 COMT 相关的有 15 对, 数量居中, 而与 MYB 转录因子相关的最多, 有 18 对。在此基础上, 采用 dCAPS Finder 2.0 软件开发基于 SNP 位点的 CAPS 标记, 共获得 26 对有酶切位点的 CAPS 标记(表 4), 开发比率为 66.7%, 这 26 对 CAPS 标记分别具有能被 *Ssp* I、*Cla* I 和 *Mse* I 等 17 种内切酶识别的位点。

表 2 黄麻 *unigene* 中 SNP 位点的分布密度表
Table 2 Distribution density of SNP loci in the *unigenes* of jute

SNP 位点的数目 Number of SNP loci	对应 <i>unigene</i> 序列的数目 Corresponding number of <i>unigenes</i>	所占比例 Ratio of SNP (%)
0-1	66,525	91.53
1-2	2092	2.88
2-3	1277	1.76
3-4	948	1.30
4-5	735	1.01
5-6	258	0.35
6-7	234	0.32
7-8	174	0.24
>8	431	0.59

图 1 黄麻 *unigene* 中 SNP 位点分布的趋势图
Fig. 1 Trend of SNP loci distribution in the *unigenes* of jute
横坐标为 SNP 的密度, 即每 kb 基因序列上的 SNP 数目; 纵坐标为拥有对应密度的 *unigene* 数目。
The abscissa shows the density of the SNPs, which means the number of SNPs per kb gene sequence in jute; the ordinate shows the number of *unigenes* with corresponding density of SNPs.

由表 4 可知, 在 26 对 CAPS 引物中, 仅有 3 对与黄麻木质素合成基因 4CL 相关, 比例为 11.5%, 其对应的 *unigene* 序列仅有一条, 即 *unigene_17597*, 能识别位点的内切酶种类也仅有一种, 为 *Ssp* I; 有 9 对 CAPS 引物与 COMT 相关, 占 34.6%, 数量居中, 它们分别与 *unigene_04800*、*unigene_09002* 和 *unigene_08713* 等 3 条 *unigene* 序列相对应, 能识别变异位点的内切酶也分别有 3 种, 即 *Cla* I、*Bsi* YI 和 *Ava* III; 而与 MYB 转录因子相关的 CAPS 标记最多, 有 14 对, 占总数的 53.8%, 其对应的 *unigene* 序列有 14 条, 各不相同, 除 MYBCAPS8 和 MYBCAPS13 对应的内切酶均为 *Sec* I 之外, 其他内切酶种类也均有所不同。表明在 3 种 CAPS 标记中, 与 MYB 转录因子相关的 CAPS 标记具有多样化的、可以被多种内切酶识别的 SNP 变异位点。

表 3 39 对 SNP 引物的编号、名称及序列
Table 3 Name and sequence of 39 pairs of SNP primers

引物名称 Primer name	正向引物序列 Forward primer sequence (5'-3')	反向引物序列 Reverse primer sequence (5'-3')
4CL SNP1-1	GTGCTTTTCATCGGTGGTTT	TGTCGTAAAAC TGCTGCTGG
4CL SNP1-2	CGATTTTCATCTCGGCATTT	GCCACCCATAGTTTGGCTA
4CL SNP2-1	AGCTCCCATGAAAGCAAAGA	AGCTCCCATGAAAGCAAAGA
4CL SNP2-2	AGCTCCCATGAAAGCAAAGA	CGATTTTCATCTCGGCATTT
4CL SNP3-1	GAGCTTAGCGCGAGAAGTGT	TGTCGTAAAAC TGCTGCTGG
4CL SNP3-2	TCTCGGCATTTCTAACCACC	GCCACCCATAGTTTGGCTA
COMT SNP1F1	CACCGTTTTCCGTTGTTCTT	GTTGGAAAGCTAGCCAAACG
COMT SNP1F2	ACACTCCATCTTGGTTTGGC	GCTACCCATGGCGTTAAAAA
COMT SNP1F3	CCAGGTTTTCAACTGCAA	GCATCTCCATTGGGAACACT
COMT SNP1F4	TGAAGATGACAATAAAGCGAAAA	GTGGACAATCAGTGTGTGGC
COMT SNP1F5	ATGCAAGAAAGCCGAGTTTG	GCAGGATTCCAAGGTTTCAA
COMT SNP2F1	AAGGTGATAGAAACACGCCG	GTTGGAAAGCTAGCCAAACG
COMT SNP2F2	ACTCCATCTTGGTTTGGCAC	GCTACCCATGGCGTTAAAAA

(续表 3)

引物名称	正向引物序列	反向引物序列
Primer name	Forward primer sequence (5'–3')	Reverse primer sequence (5'–3')
COMT SNP2F3	GCTTAACCGGTTGTTGATG	GCATCTCCATTGGGAACACT
COMT SNP2F4	TGAAGATGACAATAAAGCGAAAA	GCATCTCCATTGGGAACACT
COMT SNP2F5	ATGCAAGAAAGCCGAGTTTG	TTGATTGCATCATGTTGGCT
COMT SNP3F1	AATTTCACCGTTTTCGTTG	GTTGGAAAGCTAGCCAAACG
COMT SNP3F2	ACTCCATCTTGTTTGGCAC	AAAACTGCGGTGGAGATTG
COMT SNP3F3	GCTTAACCGGTTGTTGATG	CCATTGTCTGGAATGGCTTT
COMT SNP3F4	TGAAGATGACAATAAAGCGAAAA	TGGACAATCAGTGTGTGGCT
COMT SNP3F5	ATGCAAGAAAGCCGAGTTTG	CATAATCCGGGTGAAAAAGA
MYB SNP F1	TTGCAATGGATTTGTGCAAT	GATTCCCAGCAAATGGAAGA
MYB SNP F2	TGCCTTGCTCTCTCACTAAGC	TTGAGGCCACATGTTCTAATTG
MYB SNP F3	CCTTGCTATCATTTGCCATT	TTTATGGCCCTCAAACTGG
MYB SNP F4	CAAAAGCTCATCCTCTCCG	CAAACTGCTTGGCTCATCAA
MYB SNP F5	AAGTGTTCAAAGAGAAGCAGCA	TTATAGATGGCGATGGAGGC
MYB SNP F6	AAGCAGAAGCAAAATCCCAA	AAGTAGCCATGGAGGTGTGG
MYB SNP F7	GGCCATGAGTTTCAACGACT	GACGGGAAGAGAAAACCCTT
MYB SNP F8	CTGCGTTTGTAACCCAGAT	GCTTCCTCTTCTGCTTCTCG
MYB SNP F9	CGACTCTTTCGGGACTCAAG	TCGTCGGCGTTTAAGAAGTT
MYB SNP F10	TTACACCACCGTAACCGACA	AGATGTGGATCGGATCAAGG
MYB SNP F11	AAGGAAGCCATGGAAGGACT	ATGCTTCTACCAATGCCAGG
MYB SNP F12	TCGATCATCAGCAACCAAAA	GCATTGAATTTCCGTGGTT
MYB SNP F13	CCTGATCTTAAGCGTGGCTC	TTTGGTGACGGAGTTGATGA
MYB SNP F14	GGCCATTGGAAAAC TCAAAA	TGGATGAAGAGCCTTTCACA
MYB SNP F15	ATCTTTGCCCAAAAATGCTG	GCAATTACCTGCTTCCCAAA
MYB SNP F16	GAGAAGGGAAATGGCATCAA	CTTTGGTGAACTGGAAGGA
MYB SNP F17	TTCTTCAGCTCTCCAACGCT	GGCTTGTGGTATTGTGAGCA
MYB SNP F18	CCCTTTCTAGCATATGGGCA	AACATCCCTCCATATTGATGTT

表 4 CAPS 引物名称、对应的 unigene 编号及内切酶名称
Table 4 Name, corresponding unigene code, and endonuclease name of CAPS primers

序号	引物名称	unigene 编号	内切酶的名称
Code	Primer name	unigene code	Endonuclease name
1	4CL CAPS1-1	unigene_17597	<i>Ssp</i> I
2	4CL CAPS2-1	unigene_17597	<i>Ssp</i> I
3	4CL CAPS3-1	unigene_17597	<i>Ssp</i> I
4	COMT CAPS1-1	unigene_04800	<i>Cla</i> I
5	COMT CAPS1-3	unigene_09002	<i>Bsi</i> YI
6	COMT CAPS1-5	unigene_08713	<i>Ava</i> III
7	COMT CAPS2-1	unigene_04800	<i>Cla</i> I
8	COMT CAPS2-3	unigene_09002	<i>Bsi</i> YI
9	COMT CAPS2-5	unigene_08713	<i>Ava</i> III
10	COMT CAPS3-1	unigene_04800	<i>Cla</i> I
11	COMT CAPS3-3	unigene_09002	<i>Bsi</i> YI
12	COMT CAPS3-5	unigene_08713	<i>Ava</i> III
13	MYB CAPS1	unigene_47261	<i>Psi</i> I
14	MYB CAPS2	unigene_17018	<i>Mse</i> I

(续表 4)

序号 Code	引物名称 Primer name	unigene 编号 unigene code	内切酶的名称 Endonuclease name
15	MYB CAPS3	unigene_53709	<i>Mbo</i> II
16	MYB CAPS5	unigene_64377	<i>Mnl</i> I
17	MYB CAPS6	unigene_10508	<i>Mbo</i> I
18	MYB CAPS7	unigene_21631	<i>Mfe</i> I
19	MYB CAPS8	unigene_31305	<i>Sec</i> I
20	MYB CAPS9	unigene_16840	<i>Hpy</i> 99 I
21	MYB CAPS11	unigene_02399	<i>Alu</i> I
22	MYB CAPS12	unigene_02162	<i>Asu</i> I
23	MYB CAPS13	unigene_46122	<i>Sec</i> I
24	MYB CAPS14	unigene_10460	<i>Tsp</i> RI
25	MYB CAPS17	unigene_59879	<i>Rsa</i> I
26	MYB CAPS18	unigene_49512	<i>Mae</i> I

加粗字体表示有多态性的引物。
The black font indicates the primers with polymorphism.

2.3 多态性 CAPS 引物的筛选

对上述 26 对 CAPS 引物进行 PCR 扩增和酶切处理表明, 有 11 对 CAPS 引物的 PCR 产物可以被酶切开来, 得到多态性条带(表 4 中黑体字标注, 扩增结果见图 2), 占总数的 42.31%, 而另外 15 对引物的 PCR 产物无法被酶切开, 虽然有清晰的条带但没有多态性。在多态性 CAPS 引物中, 与木质素合成酶基因 COMT 相关的引物有 6 对(54.5%), 与 MYB 转录因子相关的引物有 5 对(45.5%), 二者数量相当, 而未获得和 4CL 相关的多态性 CAPS 标记引物。

2.4 多态性 CAPS 标记的验证

以 12 个黄麻品种为材料, 对所筛选的 11 对 CAPS 标记的有效性进行验证。聚类分析表明(图 3), 在遗传相似系数为 0.5 处画线时, 12 个黄麻品种被区分为 2 个不同的类群, 第 I 个类群包括 7 个品种, 除

“日本 5 号”外, 其他 6 个品种均为圆果种黄麻; 第 II 个类群包括 5 个黄麻品种, 均为长果种黄麻。由此可见, 所用的黄麻材料大致有按 2 个不同栽培类型, 即长果种和圆果种黄麻区分的趋势。来自日本的圆果种黄麻“日本大分青皮”和长果种黄麻“日本 5 号”被聚在一起, 推测这 2 个品种虽然属于不同的栽培类型, 但其木质素合成相关的基因以及 MYB 转录因子均可能有很高的相似性, 因而聚在同一个小类群中。同时进一步细分发现, 在第 I 个类群中, “台湾加利麻”和“龙溪长果”又各自成一类, 没有和其他品种的黄麻聚在一起, 推测这 2 个品种的木质素合成相关的基因以及 MYB 转录因子较为特殊, 与其他品种有所区别。综上所述, 所筛选出来的 11 对 CAPS 多态性标记, 可以将 12 份不同类型的黄麻品种较好地区别开来, 验证了所开发的 CAPS 标记的有效性和可行性。

图 2 COMT CAPS1-3 引物扩增和酶切结果
Fig. 2 Amplified and enzyme digestion result of CAPS primer COMT CAPS1-3
M: marker; 1~12 表示试验所用的 12 份黄麻种质资源的代号。
M: marker; 1~12: the code of 12 jute accessions used in this study.

图 3 12 份黄麻品种基于 CAPS 标记的聚类图

Fig. 3 Dendrogram of 12 jute accessions based on CAPS markers

3 讨论

3.1 黄麻转录组中 SNP 位点的分布

SNP 是由单碱基突变引起的 DNA 序列多态性, 具有数量丰富、分布广泛、覆盖度高等特点^[34]。黄麻转录组序列中共检测到 67,567 个 SNP 位点, 平均每 440 bp 出现一个 SNP 位点, 与其他作物相比较, 出现的频率较低。前人研究表明, 水稻基因组中约 268 bp 存在 1 个 SNP^[35], 玉米基因组中每 60~120 bp 就存在 1 个 SNP^[36], 大豆基因组中每 185~266 bp 有 1 个 SNP^[37], 而茶树基因组中平均每 172 bp 有 1 个 SNP^[38]。SNP 出现的频率, 第一, 与物种本身基因组的差异有关, 不同物种不同。第二, 存在于非编码序列中的 SNP 会提高 SNP 出现的频率^[39], 而黄麻中存在于非编码序列的 SNP 位点较少, 导致其出现的频率降低。第三, 试验样本大小及测序长度也会影响 SNP 的频率^[40], 今后应进一步增加黄麻转录组的测序长度, 以提高 SNP 的频率。

3.2 所开发 CAPS 标记的多态性

本试验的 CAPS 标记是在前期黄麻转录组序列中 SNP 位点的基础上开发的。SNP 标记因为在基因组中分布广泛、遗传稳定等特点, 在植物遗传育种研究方面有巨大的优势, 但是由于检测手段的限制, 阻碍了 SNP 标记的广泛应用。而 CAPS 方法无需昂贵的设备、操作简单, 可以实现中通量的 SNP 检测, 因而可以大大促进 SNP 标记在植物遗传育种中的应用^[38]。本研究以 39 对 SNP 引物为基础, 获得了 26 对 CAPS 标记, 开发成功的比率为 66.7%, 与柑橘 (65%) 和水稻 (67.8%) 的成功率相似^[17,41], 但低于茶

树和大豆 SNP 快速转化为 CAPS 方法的比例 (81.8% 和 86.21%), 高于大豆传统方法转化的比例 (52.51%)^[14]。造成这种差异的原因, 首先与酶切位点选择和分析的标准有关, 严格的酶切位点选择和分析标准, 可以提高转化的成功率^[38]; 其次, 由于某些 SNP 并未引起限制性内切酶识别位点的改变, 不能直接使用内切酶进行检测; 再次, 若在同一扩增产物中存在多个与 SNP 所在酶切识别位点相同的酶切位点, 直接使用内切酶检测难以分辨^[19]。另外, 多态性 CAPS 引物所占的比率为 42.3%, 也比较低。原因可能是本研究开发的 CAPS 引物是和 4CL、COMT 等木质素合成酶基因及 MYB 转录因子相关的, 针对性和目的性较强, 而所用的供试材料在这些位点仅有部分多态性, 从而导致多态性引物的比率降低。也有可能是物种本身的特异性引起的, 潜宗伟等研究亦表明, 分子标记引物的多态性比率因物种不同而有所差异^[42]。

3.3 所开发 CAPS 标记的有效性

所开发的 CAPS 标记引物可以将供试的 12 份黄麻种质较好地区别开来, 供试材料有大致按栽培类型, 即长果种黄麻和圆果种黄麻聚类的趋势。同时, 来自日本的 2 个品种聚在一起, 体现了按种质地理来源聚类的特点。另外, 基于 CAPS 标记的聚类分析也能鉴别出个别特异的种质类型, 如台湾加利麻。以上结果均表明, 所开发的 CAPS 标记在黄麻遗传多样性分析上是有效和可行的, 且可用于黄麻种质资源鉴定等研究, 为黄麻杂交育种亲本选配提供理论依据。综上所述, 本试验所开发的黄麻 CAPS 标记是黄麻遗传育种研究中的一项理想的分子标记方法, 可为黄麻分子标记辅助育种和性状改良提供有效方法。

4 结论

对黄麻转录组序列中的 SNP 位点进行了分析, 发现平均每 440 bp 出现一个 SNP 位点, 其分布频率低于其他作物, 推测与物种之间的基因组差异性有关, 而黄麻中存在于非编码序列的 SNP 位点较少也是影响因素之一。同时, 设计了与黄麻木质素合成基因 4CL、COMT 及转录因子 MYB 相关的 39 对 SNP 引物, 并以此为基础开发了 26 对 CAPS 标记, 开发成功率为 66.7%, 其中 11 对 CAPS 标记具有多态性, 多态性比例为 43.2%。开发的 CAPS 标记能较好地 将 12 份不同类型的黄麻种质按栽培类型区分开来,

同时亦能鉴别出特异种质, 表明所开发的 CAPS 是适用于黄麻种质资源鉴定的较理想的分子标记方法, 可为黄麻杂交育种亲本选配提供理论依据, 同时也为黄麻分子标记辅助选择育种提供了有效方法。

References

- [1] 熊和平. 麻类作物育种学. 北京: 中国农业科学技术出版社, 2008. pp 156–157.
Xiong H P. Breeding Science of Bast and Leaf Fiber Crops. Beijing: Chinese Agricultural Science and Technology Press, 2008. pp 156–157 (in Chinese).
- [2] 兰红艳, 张延辉. 麻类纤维的性能及其应用. 上海毛麻科技, 2009, (3): 1–5.
Lan H Y, Zhang Y H. The properties and applications of bast fibers. *Shanghai Hemp Technol*, 2009, (3): 1–5 (in Chinese with English abstract).
- [3] 孔华, 郭安平, 郭运玲, 刘恩平, 贺立卡. 木质素生物合成及转基因调控研究进展. 热带农业工程, 2009, 33(5): 47–53.
Kong H, Guo A P, Guo Y L, Liu E P, He L K. Advances in study of lignin biosynthesis and its genetic manipulation. *Trop Agric Engineer*, 2009, 33(5): 47–53 (in Chinese with English abstract).
- [4] Li J, Han G, Sun C, Sui N. Research advances of MYB transcription factors in plant stress resistance and breeding. *Plant Signal Behav*, 2019, 14: 1613131. doi: 10.1080/15592324.2019.1613131.
- [5] 李利斌, 牛佳玉, 王永强, 杨宗辉, 杨桂兰, 侯丽霞, 刘淑梅, 曹齐卫, 孟昭娟. 分子标记技术在黄麻农艺性状基因定位上的应用. 植物生理学报, 2019, 55: 1–13.
Li L B, Niu J Y, Wang Y Q, Yang Z H, Yang G L, Hou L X, Liu S M, Cao Q W, Meng Z J. Application of molecular markers in gene localization of cucumber (*Cucumis sativus*) agronomic. *J Plant Physiol*, 2019, 55: 1–13 (in Chinese with English abstract).
- [6] 张明飞, 于卓, 于肖夏, 李景伟, 马艳红, 吴国芳, 王颖. 四倍体马铃薯 SRAP 分子遗传连锁图谱的构建. 草业学报, 2019, 28(9): 190–199.
Zhang M F, Yu Z, Yu X X, Li J W, Ma Y H, Wu G F, Wang Y. Construction of a genetic linkage map for tetraploid potato based on SRAP markers. *Acta Pratac Sin*, 2019, 28(9): 190–199 (in Chinese with English abstract).
- [7] Yang X, Zhang Y, Cheng Y F, Chen X H. Transcriptome analysis reveals multiple signal network contributing to the *Verticillium* wilt resistance in eggplant. *Sci Hortic*, 2019, 256. doi: 10.1016/j.scienta.2019.108576.
- [8] Zhang Z H, Cao B L, Li N, Chen Z J, Xu K. Comparative transcriptome analysis of the regulation of ABA signaling genes in different rootstock grafted tomato seedlings under drought stress. *Environ Exp Bot*, 2019, 166. doi: 10.1016/j.envexpbot.2019.103814.
- [9] 贾昌路, 张瑶, 朱玲, 张锐. 转录组测序技术在生物测序中的应用研究进展. 分子植物育种, 2015, 13: 2388–2394.
Jia C L, Zhang Y, Zhu L, Zhang R. Application progress of transcriptome sequencing technology in biological. *Mol Plant Breed*, 2015, 13: 2388–2394 (in Chinese with English abstract).
- [10] 王尧龙, 黄璐琦, 袁媛, 查良平. 药用植物转录组研究进展. 中国中药杂志, 2015, 11: 2055–2061.
Wang Y L, Huang L Q, Yuan Y, Zha L P. Research advances on analysis of medicinal plants transcriptome. *Chin J Chin Mater Med*, 2015, 11: 2055–2061 (in Chinese with English abstract).
- [11] 肖炳光, 邱杰, 曹培健, 桂毅杰, 卢秀萍, 李永平, 樊龙江. 利用基因组约法开发烟草 SNP 标记及遗传作图. 作物学报, 2014, 40: 397–404.
Xiao B G, Qiu J, Cao P J, Gui Y J, Lu X P, Li Y P, Fan L J. Development and genetic mapping of SNP markers via genome complexity reduction in tobacco. *Acta Agron Sin*, 2014, 40: 397–404 (in Chinese with English abstract).
- [12] Tao A F, Huang L, Wu G F, Reza K A, Qi J M, Xu J T, Fang P P, Lin L H, Zhang L W, Lin P Q. High-density genetic map construction and QTLs identification for plant height in white jute (*Corchorus capsularis* L.) using specific locus amplified fragment (SLAF) sequencing. *BMC Genomics*, 2017, 18: 355.
- [13] Hayden M J, Tabone T, Mather D E. Development and assessment of simple PCR markers for SNP genotyping in barley. *Theor Appl Genet*, 2009, 119: 939–951.
- [14] 束永俊, 李勇, 柏锡, 才华, 纪巍, 朱延明. 基于基因重测序信息的大豆基因靶向 CAPS 标记开发. 作物学报, 2009, 35: 2015–2021.
Shu Y J, Li Y, Bai X, Cai H, Ji W, Zhu Y M. Development of soybean gene-driven functional CAPS markers based on gene re-sequencing. *Acta Agron Sin*, 2009, 35: 2015–2021 (in Chinese with English abstract).
- [15] 杨广阔, 陈子强, 陈在杰, 苏军, 陈松彪, 王锋. 基于二代测序数据开发以 93-11 为亲本的水稻 SNP-dCAPS 标记的研究实例. 分子植物育种, 2014, 12: 1288–1295.
Yang G K, Chen Z Q, Chen Z J, Su J, Chen S B, Wang F. Developing rice SNP-dCAPS markers based on next generation re-sequencing data of 93-11 as parental line, a case study. *Mol Plant Breed*, 2014, 12: 1288–1295 (in Chinese with English abstract).
- [16] 王彩芬, 刘冬成, 马晓玲, 安永平, 张文银, 马静. 水稻耐盐基因 SKC1 特异性 CAPS 标记的开发与验证. 分子植物育种, 2015, 13: 2437–2440.
Wang C F, Liu D C, Ma X L, An Y P, Zhang W Y, Ma J. Development and validation of the specific CAPS markers of rice salt tolerant gene. *Mol Plant Breed*, 2015, 13: 2437–2440 (in Chinese with English abstract).
- [17] 司文洁, 吴林楠, 郭利建, 周梦蝶, 刘香利, 马猛, 赵惠贤. 小麦粒重相关基因 TaCYP78A5 功能标记开发及验证. 作物学报, 2019, 12: 1905–1911.
Si W J, Wu L N, Guo L J, Zhou M D, Liu X L, Ma M, Zhao H X. Development and validation of the functional marker of grain weight-related gene TaCYP78A5 in wheat (*Triticum aestivum* L.). *Acta Agron Sin*, 2019, 12: 1905–1911 (in Chinese with English abstract).
- [18] 雷天刚, 何永睿, 彭爱红, 许兰珍, 刘小丰, 姚利晓, 邹修平, 江东, 陈善春. 柑橘 CAPS 标记和 AS-PCR 引物的开发. 园艺学报, 2012, 39: 1027–1034.
Lei T G, He Y R, Peng A H, Xu L Z, Liu X F, Yao L X, Zou X P, Jiang D, Chen S C. Development of CAPS markers and allele-specific PCR primers in citrus. *Acta Hortic Sin*, 2012, 39: 1027–1034 (in Chinese with English abstract).

- [19] 阚帅帅. 玉米主效抗丝黑穗病候选基因预测及 dCAPS 标记开发. 东北农业大学硕士学位论文, 黑龙江哈尔滨, 2012.
Kan S S. Major Resistance Candidate Gene Prediction and dCAPS Markers Development of Head Smut in Maize. MS Thesis of Northeast Agriculture University, Harbin, Heilongjiang, China, 2012 (in Chinese with English abstract).
- [20] 朱文文, 徐建飞, 李广存, 段绍光, 刘杰, 卞春松, 庞万福, Walter D J, 金黎平. 马铃薯块茎形状基因 CAPS 标记的开发与验证. 作物学报, 2015, 41: 1529–1536.
Zhu W W, Xu J F, Li G C, Duan S G, Liu J, Bian C S, Pang W F, Walter D J, Jin L P. Development and verification of a CAPS marker linked to tuber shape gene in potato. *Acta Agron Sin*, 2015, 41: 1529–1536 (in Chinese with English abstract).
- [21] 尚小红, 周慧文, 严华兵, 肖亮, 陈新. 木薯块根肉质颜色基因 CAPS 标记的开发与验证. 分子植物育种, 2018, 16: 873–879.
Shang X H, Zhou H W, Yan H B, Xiao L, Chen X. Development and verification of a CAPS marker for color gene in cassava tuber. *Mol Plant Breed*, 2018, 16: 873–879 (in Chinese with English abstract).
- [22] 陶爱芬, 徐建堂, 祁建民, 吴建梅, 粟建光, 林培清, 李爱青, 李木兰, 方平平, 林荔辉. SRAP 和 ISSR 及两种方法结合在分析黄麻属起源与演化上的比较. 作物学报, 2011, 37: 2277–2284.
Tao A F, Xu J T, Qi J M, Wu J M, Su J G, Lin P Q, Li A Q, Li M L, Fang P P, Lin L H. Method comparison of using SRAP and ISSR and combination of both in origin and evolution of jute. *Acta Agron Sin*, 2011, 37: 2277–2284 (in Chinese with English abstract).
- [23] 陶爱芬, 祁建民, 李木兰, 方平平, 林荔辉, 徐建堂. SRAP 结合 ISSR 方法分析黄麻属的起源与演化. 中国农业科学, 2012, 45: 16–25.
Tao A F, Qi J M, Li M L, Fang P P, Lin L H, Xu J T. Origin and evolution of jute analysed by SRAP and ISSR methods. *Sci Agric Sin*, 2012, 45: 16–25 (in Chinese with English abstract).
- [24] 巫桂芬, 徐鲜均, 徐建堂, 陶爱芬, 张立武, 魏丽真, 潘漠, 方平平, 林荔辉, 祁建民. 利用 SRAP、ISSR、SSR 标记绘制黄麻基因源分子指纹图谱. 作物学报, 2015, 41: 367–377.
Wu G F, Xu X J, Xu J T, Tao A F, Zhang L W, Wei L Z, Pan M, Fang P P, Lin L H, Qi J M. Construction of molecular fingerprinting map in gene pool of jute with SRAP, ISSR, and SSR markers. *Acta Agron Sin*, 2015, 41: 367–377 (in Chinese with English abstract).
- [25] Kundu A, Chakraborty A, Mandal N A, Das D, Karmakar P G, Singh N K, Sarkar D. A restriction-site-associated DNA (RAD) linkage map, comparative genomics and identification of QTL for histological fibre content coincident with those for retted bast fibre yield and its major components in jute (*Corchorus olitorius* L., Malvaceae s. l.). *Mol Breed*, 2015, 35: 19.
- [26] Topdar N, Kundu A, Sinha M K, Sarkar D, Das M, Banerjee S, Kar C S, Satya P, Balyan H S, Mahapatra B S. A complete genetic linkage map and QTL analyses for bast fibre quality traits, yield and yield components in jute (*Corchorus olitorius* L.). *Cytol Genet*, 2013, 47: 129–137.
- [27] 张立武, 袁民航, 何雄威, 刘星, 方平平, 林荔辉, 陶爱芬, 徐建堂, 祁建民. GenBank 数据库中黄麻 EST-SSR 标记的开发及其通用性评价. 作物学报, 2014, 40: 1213–1219.
Zhang L W, Yuan M H, He X W, Liu X, Fang P P, Lin L H, Tao A F, Xu J T, Qi J M. Development and universality evaluation of EST-SSR markers from GenBank in jute. *Acta Agron Sin*, 2014, 40: 1213–1219 (in Chinese with English abstract).
- [28] Das M, Banerjee S, Dhariwal R, Vyas S, Mir R R, Topdar N, Kundu A, Khurana J P, Tyagi A K, Sarkar D, Sinha M K, Balyan H S, Gupta P K. Development of SSR markers and construction of a linkage map in jute. *J Genet*, 2012, 91: 21–31.
- [29] 陶爱芬, 陈婉婷, 祁建民, 徐建堂, 林荔辉, 林培清, 张立武, 方平平. 基于转录组测序的黄麻 SSR 分子标记开发与初步验证. 中国农业大学学报, 2018, 23(6): 24–33.
Tao A F, Chen W T, Qi J M, Xu J T, Lin L H, Lin P Q, Zhang L W, Fang P P. Development and identification of SSR molecular markers based on transcriptome sequences of jute (*Corchorus capsularis* L.). *J China Agric Univ*, 2018, 23(6): 24–33 (in Chinese with English abstract).
- [30] Biswas C, Dey P, Karmakar P G, Satpathy S. Discovery of large-scale SNP markers and construction of linkage map in a RIL population of jute (*Corchorus capsularis*). *Mol Breed*, 2015, 35: 119.
- [31] 陶爱芬, 谢丽丽, 祁建民, 方平平, 林荔辉, 徐建堂, 张立武, 林培清. 与黄麻炭疽病抗性相关的 SSR 分子标记筛选及新型 SNP 标记开发. 福建农林大学学报(自然科学版), 2017, 46(1): 27–33.
Tao A F, Xie L L, Qi J M, Fang P P, Lin L H, Xu J T, Zhang L W, Lin P Q. Screening of SSR markers and developing of newly SNP markers associated with anthracnose resistance in *Corchorus capsularis*. *J Fujian Agric For Univ* (Nat Sci Edn), 2017, 46(1): 27–33 (in Chinese with English abstract).
- [32] Zhao Y, Wang K, Wang W L, Yin T T, Dong W Q, Xu C J. A high-throughput SNP discovery strategy for RNA-seq data. *BMC Genomics*, 2019, 20: 160. doi: 10.1186/s12864-019-5533-4.
- [33] Nei M, Li W H. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA*, 1979, 76: 5269–5273.
- [34] 王振玉, 李威, 周晓箭, 裴小雨, 刘艳改, 周克海, 张文生, 孟清芹, 王海风, 葛勇, 李莹, 刘俊芳, 马雄风, 杨代刚. 棉花单核苷酸多态性标记研究进展. 棉花学报, 2016, 28: 399–406.
Wang Z Y, Li W, Zhou X J, Pei X Y, Liu Y G, Zhou K H, Zhang W S, Meng Q Q, Wang H F, Ge Y, Li Y, Liu J F, Ma X F, Yang D G. Review of single nucleotide polymorphism markers in cotton. *Cotton Sci*, 2016, 28: 399–406 (in Chinese with English abstract).
- [35] 郭海洋. 基于 HRM 技术的水稻全基因组 SNP 标记开发及其初步应用. 华南农业大学硕士学位论文, 广东广州, 2016.
Guo H Y. Development of Rice SNP Marker Based on HRM Technology and Its Application in Gene Location. MS Thesis of South China Agricultural University, Guangzhou, Guangdong, China, 2016 (in Chinese with English abstract).
- [36] Ching A, Caldwell K S, Jung M, Dolan M, Smith O S, Tingey S, Morgante M, Rafalskiet A J. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genet*, 2002, 3: 19.
- [37] Lam H M, Xu X, Liu X. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat Genet*, 2010, 42: 1053–1059.

- [38] 张成才. 茶树 SNP 标记的开发与应用. 中国农业科学院硕士学位论文, 北京, 2012.
Zhang C C. The Development and Applications of SNP Markers in Tea Plants (*Camellia sinensis*). MS Thesis of Chinese Academy of Agricultural Sciences, Beijing, China, 2012 (in Chinese with English abstract).
- [39] Lopez C, Piegu B, Cooke R, Delseny M, Tohme J, Verdier V. Using cDNA and genomic sequences as tools to develop SNP strategies in cassava (*Manihot esculenta* Crantz). *Theor Appl Genet*, 2005, 110: 425–431.
- [40] Alencar S, Silva J O, Togawa R. SNP discovery in apple cultivars using next generation sequencing. *BMC Proce*, 2011, 5: 42.
- [41] Thakur S, Singh P K, Rathour R, Variar M, Prashanthi S K, Gopalakrishnan S, Singh A K, Singh U D, Chand D, Singh N K, Sharma T R. Genotyping and development of single-nucleotide polymorphism (SNP) markers associated with blast resistance genes in rice using Golden Gate assay. *Mol Breed*, 2014, 34: 1449–1463.
- [42] 潜宗伟, 陈海丽, 崔彦玲. 菠菜转录组 SSR 位点分析及其分子标记的开发. 农业生物技术学报, 2016, 24: 1688–1697.
Qian Z W, Chen H L, Cui Y L. Analysis of the SSR Loci and development of molecular markers in *Spinacia oleracea* transcriptome. *Chin J Agric Biotechnol*, 2016, 24: 1688–1697 (in Chinese with English abstract).