

植物甜蛋白 mabinlin II cDNA 的 克隆与序列分析*

胡新文 郭建春 郑学勤

(中国热带农业科学院热带作物生物技术国家重点实验室, 中国海南儋州, 571737)

摘要 根据已由作者克隆的 mabinlin II B 亚基 185 bp 的 cDNA 片段, 设计合成系列引物. 从马槟榔 (*Capparis masakai* L. vl.) 种子提取总 RNA 并纯化 mRNA, 经反转录合成 cDNA 第一链. 采用 RACE 方法, 快捷克隆 mabinlin II 全长 cDNA. 经序列测定与分析, 该 cDNA 为 583 bp, 其中编码序列长 465 bp, 共编码 155 个氨基酸.

关键词 马槟榔, 植物甜蛋白, cDNA 克隆, 序列测定与分析

分类号 Q343, Q785

Cloning and Sequencing of cDNA Encoding Sweet-tasting Plant Protein Mabinlin II

HU Xinwen GUO Jianchun ZHENG Xueqin

(National Key Biotechnology Laboratory for Tropical Crops,
CATAS, Danzhou, Hainan, 571737, PRC)

Abstract A series of primers were synthesised according to the 185 bp of cDNA encoding B subunit of mabinlin II cloned by the author based on its amino acid sequence. mRNA was extracted from the seeds of *Capparis masakai* and purified. The full-length cDNA of mabinlin II was cloned by rapid amplification of cDNA ends (RACE). The result shows that the amplified DNA are 583 bp and its coding sequence is 465 bp and encodes 155 amino acids.

Key words *Capparis masakai* L. vl., sweet-tasting plant protein, cDNA cloning sequence and sequence analysis

马槟榔 (*Capparis masakai* L. vl.) 仅我国热带和亚热带地区有分布, 其种子含甜蛋白 mabinlin, 在重量基础上比较, 其甜度约为蔗糖的 400 倍^[1]. 在 mabinlin^[2] 5 种同系物中, 以

* 本研究为国家自然科学基金资助项目

收稿日期: 1998-05-18. 作者: 胡新文, 35岁, 副研究员, 博士, 硕士生导师; 郭建春, 女, 33岁, 副研究员, 正在捷克马萨里克大学攻读博士学位; 郑学勤, 男, 69岁, 教授, 博士生导师

mabinlin 具最佳热稳定性,即在 80 °C 温浴 48 h 不改变其甜味活性^[2]. mabinlin 由 A、B 二条肽链组成,分别含 33 和 72 个氨基酸,分子量 10.8 kD^[3]. 作为一种新型的甜味剂, mabinlin 具有高甜度、无毒、非热源性等优点,因而具有很高的经济价值. 但马槟榔为高大木质藤本植物,数量十分稀少,致使对它的利用受到很大的限制.

基因工程技术的迅速发展,开辟了利用植物甜蛋白资源的崭新途径. 作者依据 mabinlin 氨基酸序列成功地克隆了其 B 亚基 185 bp 的 cDNA 序列^[4]. 在此基础上,本文又报道了 mabinlin 全长 cDNA 序列. 为进一步开展该基因的分子发酵和植物改良研究提供了充分的基础.

1 材料与方法

1.1 材料

马槟榔 种子采自云南省屏边县.

限制性内切酶及连接酶为 Biolabs 产品, DNA 合成与测序试剂盒为 ABI 公司产品, cDNA 合成试剂盒为 Promega 公司产品, PCR 试剂盒为华美公司产品.

1.2 方法

1.2.1 马槟榔种子总 RNA 提取 参考文献[5]方法进行.

1.2.2 mRNA 的纯化 参考文献[6]方法进行.

1.2.3 引物设计

```

AGACAATGCTGCAACCAAGCTGCGTCAAGTGGACAGACCTTGTGTTTGC      48
CCTGTCCTCAGACAAGCTGCCCAGCAGGTGCTCCAACGACAAAATAATC      96
CAGGCTCCACAGCACTGAGGCGTCTCTTCGATGCCGCAAGAAATTTG      144
CCCAACATCTGCAACATACCGAACATCGGAACTTGCCCAT      185
  
```

图 1 依据 mabinlin II B 亚基 185 bp cDNA 片段设计引物
划线部分依次为 primer 1, primer 3 和 primer 2

Fig. 1 Making designs for primers based on 185 bp of cDNA sequence encoding B subunit of mabinlin
The underlined sequences are orderly primer 1, primer 3 and primer 2

另外,还设计合成了(dT)₁₇接头(primer 4)及接头引物(primer 5):

Primer 4: 5 GACCAACGCGTATCGATGT CGACTTT TTT TTT TTT TTT TTT TTT TTT TTT TTT (CAG)3

Primer 5: 5 GACCAACGCGTATCGATGT CGAC 3

依据 mabinlin 的 3' cDNA 末端与 5' cDNA 末端设计合成 2 个引物,以扩增 mabinlin cDNA 全序列:

primer 6: 5 CCCTAGCAATGGCGAAGCTC 3

primer 7: 5 TAGGAGTAGTGCTAGTT CGAC 3

1.2.4 3' cDNA 末端的扩增

取 1 μg mRNA 于 15.5 μl 水中,65 °C 温浴 3 min,冰浴冷却. 加入 5 X 第一链合成缓冲液 5 μl, 2.5 mmol/L dNTP 2 μl, (dT)₁₇接头(primer 4) 0.5 μl (1 g/L), RNase 抑制剂 1 μl (16 U), AMV 反转录酶 1 μl (25 U), 42 °C 反应 2 h. 反应混和物稀释在 1 ml TE 中, 4 °C 保

存.

取上述反应混和物 1 μ l, primer 1 和接头引物 (primer 5) 各 2 μ l (10 μ mol/L), 2.5 mmol/L dNTP 5 μ l, 10 X PCR buffer 5 μ l, 反应总体积 50 μ l. 预反应参数为 95 2 min, 冷却至 72 , 加入 3 U TaqDNA 聚合酶, 52 2 min, 72 40 min. 然后进 PCR 循环, 反应参数为 94 40 s, 52 2 min, 72 3 min, 反应进行 40 个循环, 最后 72 延伸 15 min.

1.2.5 5'cDNA 末端的扩增

取 1 μ g mRNA 进行 cDNA 第一链的合成, 其方法与 3' cDNA 末端扩增的反转录方法相同, 只是用 primer 2 取代 primer 4. 采用 High pure PCR purification kit (Boehringer Mannheim) 纯化 cDNA. 将纯化的 cDNA 溶解在 23 μ l 水中. 加入 1 μ l dATP (6 mmol/L), 1 μ l 末端 DNA 转移酶 (19 U), 6 μ l 反应缓冲液 (5 X). 37 反应 10 min, 65 温浴 15 min. 用 TE 将反应混和物稀释至 500 μ l. 取其 10 μ l 作为模板, 用 primer 4、primer 3 和 primer 5 进行 PCR 反应, 预反应和 PCR 反应参数与 3' cDNA 扩增的相同.

1.2.6 Mabinlin II 全长 cDNA 的扩增

用 oligo(dT)₁₅ 将 mRNA 反转录合成 cDNA 第一链, 用 primer 6 和 primer 7 进行 PCR 扩增. 反应参数为 94 1 min, 55 1 min, 72 2 min, 最后一循环 72 延伸 10 min.

1.2.7 PCR 产物的纯化、克隆与测序

PCR 产物行普通琼脂糖凝胶电泳, 切下所需的条带, 采用 DNA clean-up system (Promega) 进行纯化, 用 pGEM-T vector system (Promega) 进行克隆^[7], 并在 ABI 377A 测序仪上自动测序.

2 结果与讨论

2.1 3' cDNA 末端序列

在 mabinlin 3' cDNA 末端扩增中, 模板 cDNA 是由 (dT)₁₇ 接头反转录 mRNA 获得的. PCR 反应的上游引物 primer 1 是与 mabinlin cDNA 特异性互补的一段寡聚核苷酸, 下游引物 primer 5 的互补区位于 (dT)₁₇ 接头上. 将扩增产物插入 pGEM-5zf/EcoRV-T 载体上, 用 M13 primer 正向测序, 得到 307 个有效碱基, 包含 mabinlin 3' cDNA 末端的一段 poly(A) 序列.

```
CAGCAGGTGCTCCAACGACAAAT AATCCA GGGTCCACAGCAGTTGAGGCGTCTCTTCGATGCCGCAA 67
GAAATTTGCCCAACATCTGCAACATACCCAACATCGGAACTTGCCCATTCAGAACATGGCCCTAGCC 134
GAACCAACCACTGGCTGACGGAGAGGTGTGTTGTAGAATCCCATGTTGTAGTGTGTTAATAATATAG 201
TTAGCATCGAGGCTAATGTCGAACTAGCACTACTCCTAATAAGAGGTTTCCAAGTCTCTTAACTAA 268
TAAAAAAAAAAAAAAAAAAGTCGACATCGATACGCGTGGTC 307
```

图2 mabinlin II 3'cDNA 末端序列

Fig.2 3'end sequence of mabinlin II cDNA

2.2 5' cDNA 末端序列

用 primer 2 反转录合成 cDNA 第一链. 在末端 DNA 转移酶的作用下, cDNA 第一链的 5' 末端加上了一定数量的碱基 A. 用 (dT)₁₇ 接头合成 cDNA 第二链作为 PCR 模板, primer 5 为上游引物, 下游引物 primer 3 的互补区位于 primer 2 的稍上游. 测序结果表明, 扩增产物

为 464 个有效碱基, 包含 mabinlin 5' cDNA 末端 426 bp 和(dT)₁₇接头的 38 bp.

```

GACCACGCGTATCGATGGCGACTTTTTTTTTTTTTTTT CACCCAAAACCTAGCAATGGCGAAGCTC      67
ATCTTCTCTTCGCGACCTTGGCTCTCTTGGTTCTCCTAGCGAACGCCCCATCCAGACCACCGTTA      134
TCGAGGTCGATGAAGAAGAAGACAACCAA CTGTGGAGATGT CAGAGGCAGTTCCTGCAGCACCCAGCG      201
ACTCCGGGCTTGCCAGCGGTTTCATCCACCGACGAGCCCAGTTCGGCGGA CAGCCGATGAGCTTGAA      268
GACGAAGTCGAGGACGACAACGATGACGAAAAC CAGCCAAGCGGACCGGCGCTCAGACAATGCTGCA      335
ACCAGCTGCGTCAAGTGGACAGACCTTGTGTTTGCCTGTCCTCAGACAAAGCTGCCACGAGGTGCT      402
CCAACGACAAATAATCCAGGGTCCAAGCAGTTGAGGCGTCTCTTCGATGCCGCAAGAAATT      464

```

图3 mabinlin II 5'cDNA 末端序列

Fig.3 5' end sequence of mabinlin II cDNA

2.3 mabinlin 全长 cDNA 序列

以 primer 6 和 primer 7 为 PCR 引物, 以马槟榔种子 mRNA 反转录合成的 cDNA 为模板进行扩增, 其产物行琼脂糖凝胶电泳, DNA 片段大小位于 PCR 分子量标准的第 3、4 条带之间, 约 600 pb(图 4). 测序结果整理的 PCR 产物的核苷酸序列, 和它所编码的氨基酸序列见图 5.

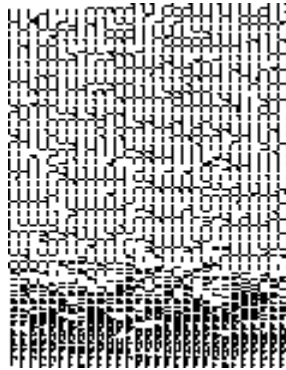


图4 PCR 产物琼脂糖凝胶电泳图

1. mabinlin 全长 cDNA 的 PCR 产物. 2. PCR 分子量标准

Fig.4 Agarose gel electrophoretogram of PCR products

1. PCR products of full-length cDNA encoding mabinlin
2. PCR markers: 1543, 994, 695, 515, 373, 237 bp

本研究克隆得到的 mabinlin 全长 cDNA 序列共 583 bp, 其中编码序列长 465 bp, 共编码 155 个氨基酸. 从 mabinlin A 链与 B 链的氨基酸序列, 可以推导出, mabinlin 在转译后加工剪切的过程中, 在 N 末端剪切掉 35 个氨基酸. 依据信号肽的一般规律, 推测其信号肽酶位点位于 Ala²⁰. 这样, N 端剪切的 35 个氨基酸, 包括 20 个氨基酸组成的信号肽和 15 个氨基酸组成的 N 末端延伸肽. 另外, 在 A、B 二链之间还须剪切 14 个氨基酸组成的连接肽 (Glu⁶⁹至 Asn⁸²) 和一个氨基酸的 C 端延伸 (Pro¹⁵⁵). 由此, mabinlin 前体加工成 33 个氨基酸组成的 A 链 (Gln³⁶至 Asp⁶⁸) 和 72 个氨基酸组成的 B 链 (Gln⁸³至 Tyr¹⁵⁴).

在此基础上,便可进一步开展 mabinlin 基因的植物遗传转化的研究,同时拟构建 mabinlin A、B 二肽链编码序列的大肠杆菌表达载体,使其在大肠杆菌中表达,以期获得含有甜蛋白味的转基因植物和具甜味活性的表达产物。这些研究工作正在进行之中。

CCCTAGCAATGGCGAAGCTCA TCTTCCTCCTCGCGACCTTGGCTCTCTTCGTTCTCCTAGCGAAC	65
M A K L I F L F A T L A L F V L L A N	19
GCCTCCATCCAGACCACCGTTATCGAGGTCGATGAAGAAGAAGACAACCAACTGTGGAGATGT CAG	131
A S I Q T T V I E V D E E E D N Q L W R C Q	41
AGGCAGTTCCTGCAGCACCGACTCCGGGCTTGC CAGCGGTCATCCACCGACGAGCCCA GTTC	197
R Q F L Q H Q R L R A C Q R F I H R R A Q F	63
GGCGGACAGCCCGATGAGCTTGAAGACGAAGTCGAGGACGACAACGATGACGAAAACCGCCAAGG	263
G G Q P D E L E D E V E D D N D D E N Q P R	85
CGACCGGGCTCAGACAATGCTGCAACCAGCTGCGTCAAGTGACAGACCTTGTGTTTGCCTGT C	329
R P A L R Q C C N Q L R Q V D R P C V C P V	107
CTCAGACAAGCTGCCAGCAGGTGCTCCAACGACAAATAATCCAGGGTCCACAGCAGTTGAGGCGT	395
L R Q A A Q Q V L Q R Q I I Q G P Q Q L R R	129
CTCTTCGATGCCGCAAGAAATTTGCCCAACA TCTGCAACATACCCAACATCGAACTTGCCCATTC	461
L F D A A R N L P N I C N I P N I G T C P F	151
AGAACATGGCCCTAGCGGAACCAACCAAGTGGCTGACGGAGAGGTGTGTTGTAGAATCCCATGTTGT	527
R T W P	155
AGTGTGTTAATAATATAGTTAGCATCGAGGCTAATGTCGAACTAGCACTACTCCTA	583

图5 mabinlin II 全长 cDNA 序列及所编码的氨基酸序列

Fig. 5 Full-length cDNA sequence and amino acid sequence of mabinlin II

参 考 文 献

- 1 崔洪志,李敏,郭三堆.植物甜蛋白的研究进展.生物技术通报,1997,(2):10~13
- 2 丁鸣,胡忠.马槟榔甜味蛋白的研究,稳定性和变性.云南植物研究,1986,8(2):181~192
- 3 Liu X, Zheng H. Purification, complete amino acid sequence and structure characterization of the heat stable sweet protein mabinlin. Eur J Biochem. 1993, 211: 281~287
- 4 胡新文,郭建春,郑学勤.植物甜蛋白 mabinlin B 亚基 cDNA 的克隆与序列分析.热带作物学报,1998,(待发表)
- 5 傅荣昭,孙勇如,贾士荣.植物遗传转化技术手册.北京:中国科技出版社,1994.141~144
- 6 卢圣栋.现代分子生物学实验技术.北京:高等教育出版社,1993.151~153
- 7 胡新文,郭建春,郑学勤.Mabinlin 基因克隆中 T-载体技术的应用.热带作物学报,1995,16(增刊):54~58